NATURAL LANGUAGE LEARNING
BY COMPUTER

BY

LAURENT SIKLOSSY

SUBMITTED TO THE CARNEGIE-MELLON UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

PITTSBURGH, PENNSYLVANIA

1968

THIS WORK WAS SUPPORTED PARTIALLY BY THE DANFORTH FOUNDATION AND
PARTIALLY BY THE ADVANCED RESEARCH PROJECTS AGENCY OF THE OFFICE
OF THE SECRETARY OF DEFENSE (SD-146) AND IS MONITORED BY THE AIR
FORCE OFFICE OF SCIENTIFIC RESEARCH.
LEARNING A NATURAL LANGUAGE IS TAKEN AS AN IMPROVEMENT IN A SYSTEM'S ABILITY TO EXPRESS SITUATIONS IN A NATURAL LANGUAGE.

THIS DISSERTATION DESCRIBES A COMPUTER PROGRAM, CALLED ZBIE, WRITTEN IN IPL-V, WHICH ACCEPTS THE DESCRIPTION OF SITUATIONS IN A UNIFORM, STRUCTURED FUNCTIONAL LANGUAGE AND TRIES TO EXPRESS THESE SITUATIONS IN A NATURAL LANGUAGE. EXAMPLES ARE GIVEN FOR GERMAN AND, MOSTLY, RUSSIAN.

AT RUN-TIME, ZBIE BUILDS SIMPLE MEMORY STRUCTURES. PATTERNS AND SETS ARE BUILT ON THE FUNCTIONAL LANGUAGE. THE TRANSLATION RULES OF THE PATTERNS AND AN IN-CONTEXT VOCABULARY PROVIDE THE TRANSITION TO THE NATURAL LANGUAGE. ZBIE IS A CAUTIOUS LEARNER, AND AVOIDS ERRORS BY SEVERAL MECHANISMS. ZBIE IS CAPABLE OF SOME EVOLUTIONARY LEARNING.
ACKNOWLEDGMENTS

AMONG THE MANY STUDENTS WHO HAVE HELPED ME, I WOULD LIKE TO THANK PARTICULARLY G. BERGlassen, R. Fikes, P. Freeman, R. Grove and the LOCAL IPL-V EXPERTS, R. Bushyager and T. Cunningham.

MY WIFE HAS SPENT MANY HOURS PROOFREADING THE THESIS AND HAS GIVEN ME MUCH MORAL SUPPORT.

ABOVE ALL I AM INDEBTED TO PROFESSOR HERBERT A. SIMON FOR HIS GUIDANCE AT ALL STAGES OF THIS WORK.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>TITLE PAGE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>II</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>III</td>
</tr>
<tr>
<td>TABLE OF CONTENTS</td>
<td>IV</td>
</tr>
<tr>
<td>CHAPTER I. INTRODUCTION.</td>
<td>1</td>
</tr>
<tr>
<td>CHAPTER II.</td>
<td></td>
</tr>
<tr>
<td>A. THE FUNCTIONAL LANGUAGE, FL.</td>
<td>5</td>
</tr>
<tr>
<td>B. THE PROGRAM'S INTERNAL REPRESENTATION.</td>
<td>10</td>
</tr>
<tr>
<td>1 - THE PATTERN.</td>
<td>10</td>
</tr>
<tr>
<td>2 - ELEMENTARY DESCRIPTION OF THE MATCHING ROUTINE.</td>
<td>12</td>
</tr>
<tr>
<td>3 - THE TRANSLATION RULE.</td>
<td>13</td>
</tr>
<tr>
<td>4 - THE IN-CONTEXT VOCABULARY.</td>
<td>17</td>
</tr>
<tr>
<td>C. THE PROGRAM'S ORGANIZATION</td>
<td>18</td>
</tr>
<tr>
<td>MODE 1. INITIALIZATION.</td>
<td>18</td>
</tr>
<tr>
<td>MODE 2. SINGLE SENTENCE ANALYSIS.</td>
<td>19</td>
</tr>
<tr>
<td>CHAPTER III. LEARNING RUSSIAN.</td>
<td>34</td>
</tr>
<tr>
<td>CHAPTER IV. A CRITICAL LOOK AT ZBIE.</td>
<td>84</td>
</tr>
<tr>
<td>CHAPTER V. ENVOI</td>
<td>91</td>
</tr>
<tr>
<td>APPENDIX A. CODE FOR CYRILLIC ALPHABET.</td>
<td>93</td>
</tr>
</tbody>
</table>
CHAPTER I.

INTRODUCTION.

WORKERS IN ARTIFICIAL INTELLIGENCE HAVE, AS ONE OF THEIR GOALS, THE WRITING OF SOPHISTICATED COMPUTER PROGRAMS WHICH WILL PERFORM 'INTERESTING' AND 'DIFFICULT' TASKS. PROGRAMS CAN IMPROVE THEIR SOPHISTICATION BY LEARNING, AND LEARNING IS, INDEED, A CENTRAL PROBLEM OF ARTIFICIAL INTELLIGENCE.

ONE OF THE FIRST LEARNING TASKS THAT HUMAN BEINGS PERFORM IS ACQUIRING A NATURAL LANGUAGE (ABBREVIATED NL). THROUGHOUT HISTORY MEN HAVE USED NL'S FOR COMMUNICATING AMONG THEMSELVES AND INVESTIGATING AND INTERACTING WITH THE WORLD. FOR THE PAST DECADE, NATURAL LANGUAGE COMMUNICATION OF HUMANS WITH COMPUTERS HAS BEEN AN ACTIVE AREA OF INTEREST IN ARTIFICIAL INTELLIGENCE.

INTERESTS IN THE FIELDS OF LEARNING AND NATURAL LANGUAGE ARE COMBINED HERE IN A PROGRAM CALLED Z8IE THAT ATTEMPTS TO LEARN NATURAL LANGUAGES AT AN ELEMENTARY LEVEL. THE TASK IS CONSIDERED WORTHY OF INVESTIGATION IN ITS OWN RIGHT; THE PROGRAM DOES NOT TRY TO SIMULATE THE LEARNING BEHAVIOR OF HUMAN BEINGS.

NATURAL LANGUAGE LEARNING PROGRAMS HAVE BEEN FEW. TO THE BEST OF THE AUTHOR'S KNOWLEDGE, CORROBORATED BY A REVIEW OF
RECENT WORK IN ARTIFICIAL INTELLIGENCE (SOLOMONOFF 1966), ONLY ONE WORK CAN QUALIFY: UHR 1964. BY A PROCESS OF STRING MATCHING AND STATISTICAL LEARNING, UHR'S PROGRAMS ATTEMPT TO TRANSLATE STRINGS FROM ONE NL (NL1) INTO STRINGS OF ANOTHER NL (NL2). THE PROGRAMS ARE INSUFICIENTLY DOCUMENTED TO EXPLAIN THEIR STRUCTURE IN DETAIL, BUT FROM THE OUTPUTS EXHIBITED, SEVERAL LIMITS APPEAR: THE IDIOSYNCRASIES OF NL1 GIVE DIFFICULTIES TO THE PROGRAMS AND THE LEARNING PROCESS SEEMS TO CYCLE.

A POSSIBLE CAUSE FOR THE LACK OF INTEREST IN NL LEARNING ProgramS IS THE FEELING, AMONG MANY LINGUISTS, THAT THE LANGUAGE LEARNING TASK IS EXTREMELY ARDROUS. TWO OF THE FOREMOST SCIENTISTS IN THE FIELD OF MODERN LINGUISTICS STATED (CHOMSKY AND MILLER, 1963):

TO IMAGINE THAT AN ADEQUATE GRAMMAR COULD BE SELECTED FROM THE INFINITUDE OF CONCEIVABLE ALTERNATIVES BY SOME PROCESS OF PURE INDUCTION ON A FINITE CORPUS OF UTTERANCES IS TO MISJUDGE COMPLETELY THE MAGNITUDE OF THE PROBLEM.

TWO RELATED AREAS HAVE RECEIVED MUCH MORE ATTENTION: THE INDUCTION OF GRAMMARS OF ABSTRACT LANGUAGES AND THE EXTRAPOLATION OF SEQUENCES. SOLOMONOFF (1958) OFFERED A SKETCH FOR THE MECHANIZATION OF LINGUISTIC LEARNING, WHICH DOES NOT APPEAR TO HAVE BEEN PROGRAMMED, AND, LATER (1964), HIS FORMAL THEORY OF INDUCTIVE INFERENCE PRESENTS VARIOUS MODELS FOR EXTRAPOLATING A LONG SEQUENCE OF SYMBOLS CONTAINING ALL DATA TO BE USED IN THE

BEFORE TRYING TO DEFINE THE LEARNING TASK, LET US CONSIDER THE TECHNIQUE FOR TEACHING LANGUAGES (TO HUMANS) USED BY I. A. RICHARDS AND HIS CO-WORKERS (RICHARDS, 1961). IN THE LANGUAGE-THROUGH-PICTURES SERIES, PICTURES ARE ASSOCIATED WITH SENTENCES IN AN NL TO BE LEARNT. THE PICTURES ARE TO ACT AS A GENERAL REPRESENTATION FOR ALL HUMAN BEINGS (ENGLISH THROUGH PICTURES, BOOK I' IS PREFACED IN 41 LANGUAGES). THE STUDENT IS SUPPOSED TO USE THE PICTURES AS CLUES TO THE MEANING OF THE SENTENCES AND, BY SUCCESSIVE COMPARISONS OF THE SENTENCES, TO INFER THE VOCABULARY AND GRAMMAR OF THE NL STUDIED.

THE STUDENT'S OWN MAIN OR MOTHER TONGUE IS BYPASSED, THEREBY AVOIDING PROBLEMS OF TRANSLATION FROM ONE TONGUE INTO ANOTHER; INSTEAD THE STUDENT LEARNS TO TRANSLATE SITUATIONS DIRECTLY FROM 'REALITY' INTO A NEW NL.

AS AN ASIDE, THE AUTHOR MAY ADD THAT HE TRIED TO LEARN
HEBREW, ABSOLUTELY UNKNOWN BEFOREHAND, FROM 'HEBREW THROUGH PICTURES', HE HAD THE ADVANTAGE OF HAVING READ PREVIOUSLY SEVERAL OTHER 'LANGUAGE THROUGH PICTURES' BOOKS IN KNOWN LANGUAGES; NEVERTHELESS HE HAD GREAT DIFFICULTIES IN TRYING TO DETERMINE THE MEANINGS OF THE PICTURES OR THE CLUES TO BE DERIVED FROM THEM, AND HE FINALLY ABANDONED THE ENDEAVOUR. SEVERAL OTHER PERSONS REPORTED IDENTICAL DIFFICULTIES.)

THE PHILOSOPHIES BEHIND ZBIE AND I. A. RICHARDS'S BOOKLETS ARE SIMILAR. ZBIE USES A FUNCTIONAL LANGUAGE (ABBREVIATED FL) TO REPRESENT SITUATIONS; FL HAS THE SAME FUNCTION IN ZBIE AS THE PICTURES IN RICHARDS. BY SUCCESSIVE COMPARISONS OF SITUATIONS, AS REPRESENTED IN FL AND AS EXPRESSED IN AN NL, ZBIE TRIES TO LEARN HOW TO EXPRESS OTHER SITUATIONS REPRESENTED IN FL AND, FAILING THAT, TO USE ITS PREVIOUS KNOWLEDGE TO TRY TO LEARN HOW TO EXPRESS THE OTHER SITUATIONS. THE LEARNING SEQUENCE USED IS TAKEN FROM 'RUSSIAN THROUGH PICTURES' WITH SLIGHT MODIFICATIONS.

CHAPTER II IS DIVIDED INTO THREE PARTS.

IN PART A, WE DESCRIBE FL BRIEFLY.

IN PART B, WE DESCRIBE THE INTERNAL REPRESENTATIONS USED BY ZBIE: PATTERNS, SETS, TRANSLATION RULES AND IN-CONTEXT VOCABULARY.

IN PART C, WE DESCRIBE THE ORGANIZATION OF ZBIE AND THE MAIN PROCESSOR ROUTINES.
SINCE CHAPTER II IS RATHER DETAILED, THE READER MAY WELL WANT TO COME BACK TO IT AFTER READING THE FOLLOWING CHAPTER.

CHAPTER III COMMENTS ON ZBIE'S LEARNING OF RUSSIAN.

CHAPTER IV COMPARES ZBIE WITH UHR'S PROGRAMS AND DISCUSSES SOME OF ZBIE'S INADEQUACIES.

APPENDIX A SHOWS THE CODE USED TO TRANSLATE THE RUSSIAN CYRILLIC ALPHABET INTO LATIN ALPHANUMERICS.

APPENDIX B GIVES A SIMPLE EXAMPLE, IN GERMAN, OF ZBIE'S 'EVOLUTIONARY LEARNING' CAPABILITIES.

ZBIE IS CODED IN IPL-V (NEWELL, 1964) AND HAS BEEN RUN ON THE CARNEGIE-MELLON UNIVERSITY CDC 6-91 COMPUTER. SINCE IPL-V CODE IS TYPICALLY UNREADABLE, THE PROGRAM IS NOT ENCLOSED, BUT IT IS DESCRIBED SEMI-FORMALLY IN CHAPTER II, PART C.
CHAPTER II.

A. THE FUNCTIONAL LANGUAGE, FL.

The purpose of FL is to represent situations in a fashion somewhat similar to the pictures (and picture language) used in the language - through - pictures series. The main spirit behind FL may be summarized thus: 'similar situations should have similar representations in FL', where an intuitive feeling for similarity is used. For example the sentences:

This is a hat.
This is his hat. (referring to a boy)
This is the boy's hat.

Should have similar representations in FL. To avoid idiosyncrasies of NL's, FL is not inflected and omits articles to improve its descriptive power we have added some semantics for instance the referent of pronouns is specifically mentioned: he=(man) or you=(spoken boy) if the person spoken to is a boy (remember the pictures).

FL is not unlike the language described by Reichenbach (1947) for his analysis of English. Instead of the usual functional notation F(x₁, x₂, ..., xₙ) we use a LISP-like notation
VERBS AND FUNCTION WORDS ARE TREATED AS N-PLACE FUNCTIONS. A FEW EXAMPLES SHOULD MAKE SOME OF THE ELEMENTARY CONSTRUCTIONS CLEAR.

**FL**

<table>
<thead>
<tr>
<th>ENGLISH EQUIVALENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>(BE HAT)</td>
</tr>
<tr>
<td>(BE HAT(OF BOY))</td>
</tr>
<tr>
<td>(BE HAT(OF (BOY)))</td>
</tr>
<tr>
<td>(C BE BOOK HERE)</td>
</tr>
<tr>
<td>(BE (ON HAT TABLE))</td>
</tr>
<tr>
<td>(BE (ON TABLE HAT))</td>
</tr>
<tr>
<td>(BE (ON HAT(OF (BOY)) TABLE))</td>
</tr>
<tr>
<td>(BE (IN (AND HAT BOOK) DRAWER))</td>
</tr>
<tr>
<td>(BE (IN ((AND HAT BOOK)) (DRAWER)))</td>
</tr>
</tbody>
</table>

INPUT TO ZBIE IS IN THE FORM OF IPLV LIST STRUCTURES EQUIVALENT TO THE ABOVE NOTATION. THE MOST IMPORTANT FEATURES OF FL SEEM TO BE ITS UNIFORMITY AND STRUCTURE: FL SENTENCES ARE USUALLY TREES, NOT STRINGS.

TO MAKE EXPLICIT THE STRUCTURE OF THE TREES IN FL, IT IS SUFFICIENT TO DEFINE A RECOGNIZER FOR THE TERMINAL NODES OF THE TREES. THE FOLLOWING ARE TERMINAL NODES:
- AN ATOMIC SYMBOL (I.E. AN IPL-V REGIONAL)
  EXAMPLE: BE, TABLE, BOY, 2.
- (〈ATOMIC SYMBOL〉)
  EXAMPLE: (BOY).
- (SPEAKING 〈ANY FL CONSTRUCT〉)
  EXAMPLE: (SPEAKING BOY).
- (SPOKEN 〈ANY FL CONSTRUCT〉)
  EXAMPLE: (SPOKEN BOY(〈NUMB 2〉)), WHERE NUMB = NUMBER.
- (〈ATOMIC SYMBOL〉〈NUMB ATOMIC SYMBOL〉)
  EXAMPLE: (MAN〈NUMB 2〉).
- 〈ATOMIC SYMBOL〉〈NUMB PLUR〉, WHERE PLUR = PLURAL.
  EXAMPLE: BOY〈NUMB PLUR〉.

THE TERMINAL NODES IN FL ARE CALLED FL UNITS. ALL OTHER CONSTRUCTS IN FL ARE FL COMPLEX STRUCTURES. THE PROGRAM 'UNDERSTANDS' FL TO THE EXTENT THAT IT RECOGNIZES THE FL UNITS OF AN FL STRUCTURE.

(IN THE IPL-V IMPLEMENTATION, THE SQUARE-BRACKETED DESCRIPTION LIST ACTUALLY OCCURS ABOVE THE FL UNIT THAT IS DESCRIBED, SO THAT HAT OF BOY LOOKS LIKE ((OF BOY) HAT) WHEN CONSIDERED PURELY AS A LIST STRUCTURE. FOR AN EXAMPLE, SEE THE IPL-V DESCRIPTION OF SENTENCE 3 IN APPENDIX B. THE ORDER 'DESCRIPTION LIST - DESCRIBED FL UNIT', IS MAINTAINED IN THE PATTERN STRUCTURES, TO BE DESCRIBED. FOR AN EXAMPLE, SEE CHAPTER III, SENTENCE 12.)
AT THIS STAGE FL IS A TOOL, TO BE MODIFIED IF NECESSARY. WE MAKE NO CLAIM THAT IT IS 'THE' REPRESENTATION, OR THAT IT IS UNIVERSAL. IT IS DOUBTFUL THAT THE PICTURE LANGUAGE IS UNIVERSAL. FOR INSTANCE, IN 'GERMAN THROUGH PICTURES', P. 239, A GERMAN BOY PLAYS BASEBALL.
8. THE PROGRAM'S INTERNAL REPRESENTATION.

AT RUN-TIME, ZBIE BUILDS AND THEN USES CERTAIN MEMORY STRUCTURES WHICH WILL NOW BE DESCRIBED.

1 - THE PATTERN.

THE MAIN WORKING STRUCTURE IS THE PATTERN, WHICH IS USED TO MATCH FL STRUCTURES, AND THEN, USING THE PATTERN'S TRANSLATION RULE, TO TRANSLATE THE FL STRUCTURES INTO NL. THERE ARE TWO TYPES OF PATTERNS, DIFFERENTIATED BY A MARKER ON THE DESCRIPTION LIST (D.L.) OF THE PATTERN. A TOP PATTERN IS USED TO MATCH FL SENTENCES; A SUBPATTERN TO MATCH FL COMPLEX SUBSTRUCTURES. A PATTERN IS AN ORDERED LIST OF PAIRS; EACH PAIR CONSISTS OF THE NAME OF A SET AND AN EXTRACTOR. ON THE D.L. OF THE PATTERN IS THE TRANSLATION RULE OF THE PATTERN, AND OTHER INFORMATION.

A MORE FORMAL DESCRIPTION OF A PATTERN CAN BE GIVEN IN B.N.F.1

\[
\begin{align*}
\text{<PATTERN>} & : = \text{<P-LIST><DESCRIPTION LIST>} \\
\text{<P-LIST>} & : = \text{<SET NAME><EXTRACTOR>} \ |
\text{<SET NAME><EXTRACTOR><P-LIST>} \\
\text{<DESCRIPTION LIST>}: = \text{<ATTRIBUTE><VALUE>} | \\
& \hspace{1cm} \text{<ATTRIBUTE><VALUE><DESCRIPTION LIST>} \\
\end{align*}
\]
A pattern is a simple describable list. The elements of a set are FL units or (recursively) patterns, which are they referred to as subpatterns. A set can also be empty.

The translation rule of a pattern is a function of the extractors of the pattern. Example:

```
<P2>  <9-1 = 26178>  <9-2 = 26170>  <9-3 = 2420n>
02 9-1 04 0 04 0 04 0
00 A4 00 T0 00 Y1 00 Y1
00 Y1 02 9-2 00 Y2 00 Y2
00 A3 00 D12
00 Y2 02 9-3
00 D13
04 26180
```

The examples are taken from Chapter III. P2 is the name of the pattern; its P-list is (A4 Y1 A3 Y2); its description list is the list structure 9-1. The sets are A4 and A3; the extractors are Y1 and Y2. The translation rule of the pattern is T0(P2)=9-2. The list (Y1 Y2).

All other attribute - value pairs on the D.L. of the
PATTERNS MAY BE DISREGARDED AT THE PRESENT.

TO UNDERSTAND THE FUNCTION OF THE EXTRACTORS, LET US DESCRIBE THE PROCESS WHICH MATCHES AN FL STRUCTURE TO A PATTERN.

2 - ELEMENTARY DESCRIPTION OF THE MATCHING ROUTINE.


SO THAT THE TOTAL TRANSLATION IS 'E2TO MAL1Y1K'. IF WE CANNOT
FIND A TRANSLATION FOR 'BOY', WE INSERT A 'Z' IN THE TOTAL
TRANSLATION, WHICH WOULD BECOME: 'E2TO Z'.

IT MAY HAPPEN THAT, FOR EXAMPLE, THE SECOND ELEMENT OF FL IS
NOT AN FL UNIT, (SEE SENTENCE 12, CHAPTER III). THE MATCHING
ROUTINE THEN TRIES, RECURSIVELY, TO FIND IN A3 A SUBPATTERN WHICH
CAN MATCH THE SECOND ELEMENT OF FL.

THE MATCHING ROUTINE USES ONLY SET INCLUSION TESTS, BUT USES
THEM RECURSIVELY. IT IS SEEN THAT A NECESSARY CONDITION FOR MATCH
IS THAT THE LENGTH OF THE FL STRUCTURE AT ITS TOP LEVEL BE EQUAL
TO THE NUMBER OF PAIRS (SET, EXTRACTOR) IN THE PATTERN. THE
MATCHING ROUTINE WILL BE CONSIDERED IN DETAIL IN CHAPTER II, PART
C.

3 - THE TRANSLATION RULE.

THE TRANSLATION RULE TO(PATTERN) IS A FUNCTION FROM THE
EXTRACTORS OF THE PATTERN INTO NL AUGMENTED BY Z'S, WHERE A Z
INDICATES THAT SOMETHING WAS NOT TRANSLATED. A FEW EXAMPLES
FOLLOW:

A) LINEAR ARRANGEMENT OF THE EXTRACTORS:

<table>
<thead>
<tr>
<th>&lt;P37&gt;</th>
<th>&lt;9-1 = 26506&gt;</th>
<th>&lt;9-2 = 27334&gt;</th>
<th>&lt;9-3 = 2714&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>02 9-1</td>
<td>04 0</td>
<td>04 0</td>
<td>04 0</td>
</tr>
<tr>
<td>00 A12</td>
<td>00 10</td>
<td>00 Y96</td>
<td>00 Y96</td>
</tr>
<tr>
<td>00 Y97</td>
<td>02 9-2</td>
<td>00 Y97</td>
<td>00 Y97</td>
</tr>
<tr>
<td>00 A2</td>
<td>00 D12</td>
<td>00 Y95</td>
<td>00 Y95</td>
</tr>
</tbody>
</table>
THE TRANSLATION RULE OF P37 IS TO(P37) = {Y96 Y97 Y95}, AND IS TO BE READ AS FOLLOWS: LOOK UP IN THE VOCABULARY, IN THE PROPER CONTEXT (HERE, OF THE SET A2 AND THE PATTERN P37), THE TRANSLATION OF THE PART OF THE FL STRUCTURE THAT WAS MATCHED TO A2; THEN FOLLOW THIS TRANSLATION BY THE TRANSLATION, IN THE PROPER CONTEXT, OF THE PART OF THE FL STRUCTURE THAT WAS MATCHED TO A12, ETC...

B) LINEAR ARRANGEMENT OF SOME OF THE EXTRACTORS.

<table>
<thead>
<tr>
<th>&lt;p1&gt;</th>
<th>&lt;9-1 = 25418&gt;</th>
<th>&lt;9-2 = 25752&gt;</th>
<th>&lt;9-3 = 25274&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>02 9-1</td>
<td>04 0</td>
<td>04 0</td>
<td>04 0</td>
</tr>
<tr>
<td>00 A1</td>
<td>00 P0</td>
<td>00 P0</td>
<td>00 Y2</td>
</tr>
<tr>
<td>00 Y1</td>
<td>02 9-2</td>
<td></td>
<td>00 Y3</td>
</tr>
<tr>
<td>00 A2</td>
<td>00 Y0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>00 Y2</td>
<td>00 J3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>00 A3</td>
<td>00 J4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>00 Y3</td>
<td>00 J3</td>
<td>00 T0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>02 9-3</td>
<td></td>
</tr>
</tbody>
</table>
THE TRANSLATION RULE OF P1 IS TO(P1) = (Y2 Y3). IT IS TO BE READ AS IN CASE A) ABOVE. NOTICE THAT THE EXTRACTOR Y1 IS MISSING IN THE TRANSLATION RULE. SUCH A RULE CAN BE USED WHEN SOME FL PART IS NOT EXPRESSED IN 4L.

C) GROUPING OF SOME EXTRACTORS.

<table>
<thead>
<tr>
<th>(P0)</th>
<th>&lt;9-1 = 25418&gt;</th>
<th>&lt;9-2 = 25228&gt;</th>
<th>&lt;9-3 = 25594&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>02 S-1</td>
<td>04 0</td>
<td>04 0</td>
<td>04 0</td>
</tr>
<tr>
<td>00 A1</td>
<td>00 To</td>
<td>02 S-3</td>
<td>00 Y1</td>
</tr>
<tr>
<td>00 Y1</td>
<td>02 S-2</td>
<td>00 Y3</td>
<td>00 Y2</td>
</tr>
<tr>
<td>00 A2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00 Y2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00 A3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00 Y3</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

THE TRANSLATION RULE IS TO BE READ AS FOLLOWS: TAKE THE FL STRUCTURE THAT WAS MATCHED TO A1, FOLLOW THIS STRUCTURE BY THE FL STRUCTURE THAT WAS MATCHED TO A2; THEN LOOK UP THIS FL COMPLEX STRUCTURE IN THE VOCABULARY IN THE PROPER CONTEXT (HERE, OF THE PATTERN P0). THE NL STRING OBTAINED IS FOLLOWED BY THE TRANSLATION OF THE ELEMENT EXTRACTED BY Y3, AS ABOVE, TO GIVE THE TRANSLATION OF THE STRUCTURE MATCHED TO P0.

EXAMPLE: (CHAPTER III, SENTENCE 1.) IF WE MATCH (BE (MAN) HERE) TO P0, WE SHALL LOOK UP THE FL COMPLEX (BE (MAN)) IN THE VOCABULARY, AND FOLLOW THE TRANSLATION OF (BE (MAN)) BY THE TRANSLATION OF 'HERE' (IN THE PROPER CONTEXT).
THE ABOVE THREE FUNCTIONS ARE NOW USED BY ZBIE. NOTE THAT SINCE THE FL STRUCTURE MATCHED TO A1, SAY, CAN BE A COMPLEX FL PART, A RECURSIVE TRANSLATION LOOK-UP IS IMPLICIT IN THE TRANSLATION RULES. FOR A DETAILED EXAMPLE, SEE CHAPTER III, SENTENCE 12.

THE TRANSLATION RULE FUNCTIONS CAN BE GENERALIZED IMMEDIATELY. TWO EXAMPLES OF DIFFERENT TRANSLATION RULES FOLLOW, AND MANY MORE COULD BE DREAMED UP. (WE ASSUME A PATTERN WITH EXTRACTORS Y20 AND Y21):

D) INTRODUCTION OF CONSTANTS.

\[
92 0 \\
Y21 \\
NLI \\
Y20 0
\]

WHERE NLI IS SOME STRING IN NL. SUCH A RULE COULD BE USED WHEN SOME EXPRESSION IN NL HAS IDIOMATIC FILLERS.

E) DISJOINT PARTS.

\[
92 0 \\
FIRST(Y21) \\
Y20 \\
SECOND(Y21) 0
\]

WHERE FIRST AND SECOND ARE FUNCTIONS (WHICH MUST BE DEFINED) ON THE TRANSLATION OF THE FL STRUCTURE WHICH WAS EXTRACTED BY Y21. SUCH A RULE COULD BE USED TO HANDLE SEPARABLE GERMAN VERBS.
4 - THE IN-CONTEXT VOCABULARY.

THE VOCABULARY OF ZBIE HAS THE TWO FOLLOWING FORMS:

A) FL UNIT = FL(I), SET = A(J), PATTERN = P(K), NL STRING = NL(L).


B) FL COMPLEX = FL(I), PATTERN = P(J), NL STRING = NL(K).

TO BE INTERPRETED: THE TRANSLATION OF THE FL COMPLEX FL(I) IN THE CONTEXT OF THE PATTERN P(J) IS THE STRING NL(K) IN NL.

NOTE THAT WE CANNOT CONCLUDE THAT SOME FL UNIT FL(I) HAS A TRANSLATION IN THE CONTEXT A(J), P(K) FROM THE KNOWLEDGE THAT FL(I) IS A MEMBER OF A(J). WE MAY HAVE A TRANSLATION IN THE CONTEXT OF SOME OTHER PATTERN P(K'), OR, FOR THAT MATTER, NONE AT ALL.
C. THE PROGRAM'S ORGANIZATION.

THE CONTROL STRUCTURE OF ZBIE CONSISTS OF TWO MODES. THERE IS FIRST AN INITIALIZATION OF THE INTERNAL STRUCTURE, WHEN A FIRST PATTERN IS CONSTRUCTED. THE CONTROL THEN PASSES TO A SECOND MODE WHEN SITUATIONS ARE BROUGHT IN ONE BY ONE AND PROCESSED.

MODE 1. INITIALIZATION.

TWO SITUATIONS ARE PRESENTED TO ZBIE, REPRESENTED IN FL AND EXPRESSED IN NL. THE SITUATIONS MUST BE SUFFICIENTLY SIMILAR, SO THAT BY COMPARING THEM ZBIE CAN DEDUCE ITS FIRST PATTERN.

MORE PRECISELY, ZBIE EXPECTS THAT THE TWO FL SENTENCES WILL BE OF DEPTH 0, I.E. HAVE NO FL COMPLEX SUBSTRUCTURE; THAT THEY WILL HAVE EXACTLY ONE ELEMENT DIFFERENT AND IN THE SAME POSITION; AND THAT THE TWO CORRESPONDING NL SENTENCES WILL HAVE EXACTLY ONE ELEMENT DIFFERENT AND IN THE SAME POSITION, WHICH MUST BE AT THE BEGINNING OR AT THE END OF (EITHER) NL SENTENCE. THE DISTINCT ELEMENTS IN FL AND NL ARE THEN ASSUMED TO CORRESPOND TO EACH OTHER; THE COMMON PARTS IN FL AND NL ARE ALSO ASSUMED TO CORRESPOND TO EACH OTHER AND A FIRST PATTERN PO IS SET UP, WITH ITS SETS AND TRANSLATION RULE; THE IN-CONTEXT VOCABULARY IS INITIATED.
MODE 2. SINGLE SENTENCE ANALYSIS.

AFTER INITIALIZATION, 7BIE OPERATES IN THE FOLLOWING MODE. FIRST, THE PREVIOUSLY PROCESSED FL AND NL SENTENCES ARE ERASED. THEN, THE DESCRIPTION IN FL OF A SITUATION IS READ IN, TOGETHER WITH ITS EXPRESSION IN NL WHICH IS STORED FOR LATER USE. THE FL SENTENCE IS THEN PROCESSED FOLLOWING THE BASIC FLOW CHART BELOW (WRITTEN IN AN ALGOL-LIKE LANGUAGE, '<' AND '>' DELIMIT BLOCKS);

SINGLE SENTENCE PROCESSOR.

FOR ALL TOP PATTERNS (LAST CREATED, FIRST CONSIDERED) DO

MATCH FL TO TOP PATTERN;
COMMENT: THE MATCHING ROUTINE IS DESCRIBED IN PART 1 BELOW;
IF TOTAL MATCH THEN
TRANSLATE;
IF TRANSLATION HAS NO UNKNOWNS THEN COMPARE TO INPUT NL;
IF TRANSLATION = INPUT NL THEN EXIT AND READ IN THE NEXT SITUATION> ELSE GO TO ERROR RECOVERY>
ELSE IF TRANSLATION IS CONSISTENT WITH INPUT NL THEN STORE PATTERN LIST IN PATTERN LIST HOLDER TO PROCESS ELSE DO NOTHING>
COMMENT: THE CONSISTENCY TEST IS DESCRIBED IN PART 2;
ELSE STORE PATTERN LIST IN PATTERN LIST HOLDER>
PROCESS ELEMENTS ON THE PATTERN LIST HOLDER
COMMENT: PROCESSING THE PATTERN LISTS IS DESCRIBED IN PART 3;
IF PROCESSING SUCCESSFUL THEN EXIT AND READ IN THE NEXT
SITUATION>

COMMENT 1 EVERYTHING FAILED SO FAR;
CREATE A NEW TOP PATTERN FOR THE SITUATION.
COMMENT 1 THE PATTERN CREATING ROUTINE IS DESCRIBED IN PART 4;

BEFORE DESCRIBING HOW THE PATTERN LISTS ARE PROCESSED, LET US EXPLAIN SOME OF THE TERMS USED.

WE SAW IN CHAPTER II, PART B HOW AN FL SENTENCE WAS MATCHED TO A PATTERN. ONE OF THE MAIN ROUTINES OF ZBIE IS THE MATCHING ROUTINE, WHICH OBTAINS ALL POSSIBLE TOTAL AND PARTIAL MATCHES OF PATTERNS (AND SUBPATTERNS) TO AN FL SENTENCE. LET US CONSIDER THE MATCHING MECHANISM AGAIN.

PART 1 - THE MATCHING ROUTINE.

AN INFINITY OF TREES OF ARBITRARILY LARGE DEPTH. HOWEVER, SINCE ANY FL SENTENCE IS A FINITE TREE, NO PATTERN TREES OF A DEPTH SUPERIOR TO THE DEPTH OF THE FL SENTENCE NEED BE CONSIDERED.

THE PATTERN TREES FORMED BY THE PATTERNS RESEMBLE A DISCRIMINATION NET. A NET OF DEPTH N CANNOT DISTINGUISH BETWEEN TWO TREES OF DEPTH LARGER THAN N WHICH ARE IDENTICAL UP TO AND INCLUDING DEPTH N. HOWEVER, IN SOME CASES, N DIFFERENT PATTERNS COULD DISTINGUISH TWO SUCH TREES THANKS TO THE RECURSIVE FEATURE OF THE PATTERNS. THE PATTERN TREES ARE, THEREFORE, MORE POWERFUL THAN A SIMPLE DISCRIMINATION NET.

WE CAN VIEW THE MATCHING PROCESS AS A TEST OF WHETHER A PATTERN TREE IS CLOSE TO BEING ISOMORPHIC TO THE FL SENTENCE TREE. AN FL SENTENCE TREE IS ISOMORPHIC TO A PATTERN TREE IF THE TREES CAN BE SUPERIMPOSED, AND THE TERMINAL NODES (FL UNITS) OF THE FL SENTENCE ARE RESPECTIVELY IDENTICAL TO THE TERMINAL NODES (FL UNITS OF TERMINAL SETS) OF THE PATTERN TREE. USUALLY WE SHALL HAVE NO ISOMORPHISM; WE THEN LOOK FOR AS GOOD A FIT OF A PATTERN TREE TO THE FL SENTENCE AS WE CAN FIND.

DISREGARDING THE BOOKKEEPING CHORES INVOLVED IN BACKTRACKING AND RECURSION, IT IS ENOUGH TO CONSIDER HOW AN FL SUBTREE IS MATCHED TO AN ELEMENT OF A SET, NODE OF THE PATTERN TREE. BY CONSIDERING THE TOP PATTERNS AS ELEMENTS OF A SET, NO GENERALITY IS LOST. WE TREAT THE SPECIAL CASE OF AN EMPTY SET BY POSTULATING AN EMPTY ELEMENT FOR SUCH A SET.
PART 1.1 - OUTLINE OF THE MATCHING ROUTINE.

MATCH FL (SUB)TREE TO ELEMENT OF SET;
COMMENT: QUICK EXITS;
ELSET := ELEMENT OF SET;
FLTREE := FL SUBTREE;
(IF ELSET IS AN FL WHOLE OR IF ELSET IS EMPTY) THEN <'NO MATCH'; GO TO EXIT>;

COMMENT: ELSET IS A (SUB)PATTERN;
IF LENGTH (P-LIST OF ELSET) = 2 * LENGTH OF FLTREE
THEN <'NO MATCH'; GO TO EXIT>;
COMMENT: INITIALIZE LOOP;
MISTAKE COUNTER := 0;
COMMENT: THE MISTAKE COUNTER IS ASSOCIATED WITH THIS PARTICULAR LEVEL, AS ARE THE OTHER IDENTIFIERS;
I := 0;
LOOP:
IF MISTAKE COUNTER > 1 THEN <'NO MATCH'; GO TO EXIT>;
I := I+1;
IF I-TH SON OF FLTREE DOES NOT EXIST THEN GO TO EXIT;
IFLTREE := I-TH SON OF FLTREE;
IELSET := I-TH SET OF P-IST OF ELSET;
IF IFLTREE IS AN FL WHOLE THEN
<IF IFLTREE IS A MEMBER OF IELSET THEN
GO TO LOOP
ELSE \(<\text{MISTAKE COUNTER} = \text{MISTAKE COUNTER} + 1; \text{GO TO LOOP}>\) 
ELSE \(<\text{MATCH IFLTREE TO THE ELEMENTS OF IELSET}>\) 
COMMENT: THE MATCH IS TRIED SUCCESSIVELY ON ALL THE ELEMENTS OF IELSET. THANKS TO THE HIDDEN BOOKKEEPING, WE ONLY HAVE TO CONSIDER THE MATCH FOR ONE ELEMENT.

IF RETURN 'NO MATCH' THEN 
\(<\text{MISTAKE COUNTER} = \text{MISTAKE COUNTER} + 1; \text{GO TO LOOP}>\) 
ELSE \(\text{GO TO LOOP}>\)

IT IS SEEN THAT, BASICALLY, UP TO ONE 'MISTAKE' (AS DEFINED BY THE PROGRAM) IS ALLOWED AT A GIVEN LEVEL IN A SUBTREE.

PART 1.2 - USE OF THE MATCHING ROUTINE.

THE MATCHING ROUTINE FINDS PATTERN LISTS, WHICH IT STORES IN A PATTERN LIST HOLDER IN DECREASING ORDER OF MATCH-DEPTH. A PATTERN LIST IS A LIST OF PATTERNS, HEADED BY A TOP PATTERN AND CONTAINING OTHER (OR POSSIBLY NO) SUBPATTERNS WHICH WERE MATCHED TO SUBSTRUCTURES OF THE FL SENTENCE. WITH THE PATTERN LIST IS ASSOCIATED A CORRESPONDING LIST OF FL COMPLEX STRUCTURES IN ONE-TO-ONE CORRESPONDENCE WITH THE SUBPATTERNS THEY MATCHED. THE MATCHING ROUTINE ALSO RECORDS THE DEEPEST LEVEL IN THE FL SENTENCE TO WHICH THE MATCH WAS CARRIED, THE MATCH-DEPTH, AND WHETHER THE MATCH WAS TOTAL OR PARTIAL. A MATCH IS PARTIAL IF AT ANY POINT, DURING THE MATCH OF A PATTERN TREE TO THE FL SENTENCE, A MISTAKE COUNTER IS INCREMENTS. THE MATCH-DEPTH IS A MEASURE OF HOW GOOD THE MATCH OF THE FL SENTENCE TO THE PATTERN LIST WAS
BEEN, AND ZBIE LOOKS AT THE BEST MATCHES FIRST.

WHEN THE TRANSLATION OF A PATTERN LIST IS ATTEMPTED, IT MAY HAPPEN THAT SOME FL STRUCTURE CANNOT BE TRANSLATED. FOR EXAMPLE, THE STRUCTURE MAY BE AN FL UNIT WITH NO TRANSLATION IN THE CONTEXT CONSIDERED, OR IT IS AN FL COMPLEX PART WHICH, FOR THE PARTICULAR PATTERN LIST CONSIDERED, HAS NO CORRESPONDING SUBPATTERN TO WHICH IT WAS MATCHED. WE THEN INSERT A Z (Z₀, Z₁, Z₂, ...) IN THE TRANSLATION. THE Z IS CONSIDERED BY ZBIE TO BE THE RESULT OF AN UNFULFILLED EXPECTATION, AND ZBIE CAPITALIZES ON THESE EXPECTATIONS FOR LEARNING. ZBIE CHECKS WHETHER THE TRANSLATION OBTAINED IS CONSISTENT WITH THE INPUT NL SENTENCE.

PART 2 - THE CONSISTENCY TEST.

WE CAN IMagine THAT THE TRANSLATION AND THE INPUT NL SENTENCE ARE PUT SIDE BY SIDE. WE SHALL HAVE CONSISTENCY IF WE CAN REPLACE THE Z'S OF THE TRANSLATION BY NON-EMPTY STRINGS IN NL IN A UNIQUE NON-AMBIGUOUS WAY SO THAT THE TRANSLATION BECOMES IDENTICAL TO THE INPUT. WHETHER THIS CAN BE DONE OR NOT IS OFTEN TRIVIAL IF NO TWO Z'S ARE ADJACENT. WHEN TWO OR MORE Z'S ARE ADJACENT, ZBIE DOES NOT GIVE UP BUT REPLACES THE FIRST Z BY AN APPROPRIATE GOOD GUESS (IF AVAILABLE, SEE BELOW). THE VARIOUS GUESSES THAT WILL BE TRIED ARE PRINTED. IF PROGRESS IS MADE TOWARDS CONSISTENCY, THE GUESS IS ADOPTED (WE DO NOT HAVE COMPLETE BACKTRACKING HERE), AND WE CONTINUE. OTHERWISE, GUESSES ARE TRIED FOR THE SECOND Z. IF THIS LAST RESORT FAILS, THE TRANSLATION AND THE INPUT NL SENTENCE ARE NOT CONSISTENT.
HERE WE SEE AN EXAMPLE OF THE CAUTION ZBIE USES IN LEARNING. IF WE WANT Z Z1 TO CORRESPOND TO THE NL STRING NL1 NL2 NL3, WE CAN MAKE THE CORRESPONDENCE IN TWO WAYS (THE Z'S MUST BE MATCHED TO NON-EMPTY STRINGS IN NL):

Z=NL1, Z1=NL2 NL3;
Z=NL1 NL2, Z1=NL3;

HOWEVER, IF IT IS A GOOD GUESS TO ASSUME THAT Z=NL1, THEN WE CAN LET Z1=NL2 NL3.

AN EXAMPLE FROM CHAPTER III (SENTENCE 4) WILL ILLUSTRATE THE POINT. WE ARE TESTING THE CONSISTENCY OF (Z Z1) AND (T12 ZDES1). (Z Z1) IS THE TRANSLATION OF 'BE (spoken boy) HERE'. 'Z' COMES FROM THE UNKNOWN TRANSLATION OF 'spoken boy' AND 'Z1' FROM THE UNKNOWN TRANSLATION OF 'HERE'. WITH NO ADDITIONAL INFORMATION, ZBIE REFUSES TO MAKE A CORRESPONDENCE BETWEEN THE Z'S AND THE NL WORDS, AND WOULD FIND THE SENTENCE 'TOO HARD'. (STRICTLY, WE SHOULD LET Z=T12, Z1=ZDES1 AS Z'S CORRESPOND TO NON-EMPTY STRINGS, BUT ZBIE IS EVEN MORE CAUTIOUS.) HOWEVER, IT IS A 'GOOD GUESS', IN THIS CONTEXT, TO ASSUME THAT THE TRANSLATION OF 'HERE' IS 'ZDES1', SO THAT, AS A RESULT, Z=T12.

IF SOME PATTERN LIST COMPLETELY MATCHES AN FL SENTENCE AND WE OBTAIN A TOTAL TRANSLATION (WITHOUT ANY Z'S) WHICH IS NOT IDENTICAL TO THE INPUT NL, SEVERAL POSSIBILITIES ARISE:
1) THERE WAS A MISTAKE IN THE INPUT (THAT HAS HAPPENED).
2) THE TRANSLATION AND THE INPUT ARE TWO DIFFERENT WAYS OF
EXPRESSING THE SAME SITUATION. HERE A TEACHER (AND PREFERABLY A TIME SHARING SITUATION) IS NEEDED TO CONVEY THE INFORMATION TO ZBIE.

3) ZBIE MADE SOME ERROR SOMEWHERE AND SHOULD TRY TO RECOVER FROM THIS ERROR. GOOD ERROR RECOVERY IS A VERY HARD PROBLEM WHICH WILL BE MENTIONED LATER. AT THIS STAGE, ZBIE POSSESSES NO ERROR RECOVERY MECHANISM. IT WAS FELT THAT TRYING TO AVOID ERRORS IS A MORE FRUITFUL APPROACH THAN TRYING TO RECOVER FROM THEM. THE ERROR AVOIDING MECHANISM IS POWERFUL ENOUGH SO THAT ZBIE ACTUALLY MAKES NO ERRORS WHEN TESTED ON THE SIMPLE SENTENCES GIVEN AS EXAMPLES.

PART 3 - PROCESSING THE PATTERN LISTS.

WE NOW RETURN TO THE BASIC FLOW-CHART. THE PROCESSING IS SLIGHTLY DIFFERENT FOR FL SENTENCES WITH FL COMPLEX PARTS THAN FOR LINEAR FL SENTENCES (OF DEPTH 0). LET US DESCRIBE PROCESSING THE FORMER SENTENCES.

PART 3.1 - PATTERN LISTS OF FL SENTENCES WITH COMPLEX PARTS.

COMMENT : PROCESS PATTERN LISTS;
FOR ALL PATTERN LISTS OF THE DEEPEST LEVEL (THEN DEEPEST LEVEL - 1, ETC...) DO

<TRANSLATE PATTERN LISTS>
KEEP ONLY PATTERN LISTS WHICH HAVE A TRANSLATION CONSISTENT WITH THE INPUT NL;
PROCESS PATTERN LISTS WITH CONSISTENT TRANSLATION, STARTING
WITH THE PATTERN LISTS THAT HAVE A TRANSLATION WITH THE GREATEST NUMBER OF COMMON ELEMENTS WITH THE INPUT NL SENTENCE.

IN OTHER WORDS, ZBIE DOES A TRANSLATION 'IN PARALLEL' OF ALL PATTERN LISTS THAT MATCHED THE INPUT FL SENTENCE TO A GIVEN DEPTH, STARTING WITH THE MAXIMUM DEPTH (BEST MATCHES) FIRST. ZBIE THEN DISCARDS THE PATTERN LISTS WHICH DID NOT GIVE A CONSISTENT TRANSLATION AND STARTS PROCESSING THE PATTERN LISTS WITH THE BEST FIT TO THE INPUT, AS MEASURED BY A SET INTERSECTION WITH THE INPUT NL SENTENCE.

3.1.1 - PROCESSING THE Z'S.

PROCESSING SUCH A TRANSLATION IS EQUIVALENT TO PROCESSING THE Z'S. EACH Z TAKES THE PLACE OF SOME UNTRANSLATED FL STRUCTURE FLZ AND, THROUGH THE CONSISTENCY TEST (PART 2 ABOVE), THE Z IS TO BE REPLACED BY A NON-EMPTY NL STRING NL7. FROM THE Z WE CAN ALSO OBTAIN INFORMATION SUCH AS: TO WHICH SET, AZ, DID WE TRY TO MATCH FLZ? WHICH WAS THE PATTERN, PZ, TO WHICH THE FATHER OF FLZ (IN THE FL SENTENCE TREE) WAS BEING MATCHED.

IF FLZ IS A UNIT, WE INSERT FLZ IN AZ AND SET UP THE IN-CONTEXT VOCABULARY FLZ AZ PZ NLZ.

IF FLZ IS A COMPLEX FL, WE TRY TO CREATE A LIST OF SUBPATTERNS TO MATCH FLZ TO NLZ (SEE BELOW). IF SUCCESSFUL, WE WOULD LIKE TO INSERT THE TOP SUBPATTERN OF THE LIST AT THE TOP OF AZ.
BEFORE INSERTING, WE CHECK TO MAKE SURE THAT THE SURPATTERN
LIST WILL NOT CAUSE AMBIGUITIES. THIS CAN HAPPEN IF THERE IS
ALREADY IN AZ A SUBPATTERN LIST SP1 OF WHICH THE NEW SUBPATTERN
LIST SP2 IS A HOMOMORPHIC IMAGE, I.E. SETS OF SP2 ARE SUBSETS OF
THE CORRESPONDING SETS OF SP1 AND THE TRANSLATIONS OF THE
SUBPATTERNS IN SP1 AND SP2 ARE APPROPRIATELY IDENTICAL. IF SUCH A
CONDITION IS SATISFIED (WITH SOME MINOR ADDITIONS), NO INSERTION
TAKES PLACE. 'NOT INSERTED' IS PRINTED, AND PROCESSING OF THE
PATTERN LIST IS ENDED.

IF AN FL COMPLEX PART IS TO BE MATCHED TO A SET, THE
BACK-TRACKING MATCHING ROUTINE WILL SEARCH FOR ALL THE
SUBPATTERNS (OF THE SET) THAT WILL MATCH THE FL COMPLEX PART. IF
TWO SUBPATTERNS CAN TRANSLATE THE FL COMPLEX PART (IN DIFFERENT
WAYS), THEN A POTENTIAL AMBIGUITY IS INTRODUCED IN THE SYSTEM.
EXAMPLE: IN CHAPTER III, SENTENCE 27, THE FL COMPLEX (MOD THIS)
IS TRANSLATED AS 'E2TOT' BY SUBPATTERN P33 AND AS 'E2TA' BY
SUBPATTERN P32. THE TWO SUBPATTERNS WILL CAUSE AMBIGUITY IF THEY
BELONG TO THE SAME SET.

HERE IS ANOTHER EXAMPLE OF HOW ZRJE TRIES TO AVOID ERRORS
THE CONTEXT IN WHICH THE SUBPATTERNS WERE ORIGINALLY BUILT WAS
TOO SMALL, AND POSSIBLY LATER, BY WIDENING THE CONTEXT,
SUBPATTERNS CAN BE BUILT WITHOUT RISKING AMBIGUITIES AT A LATER
TIME.

NOTE THAT WE INSERT THE HEAD OF A SUBPATTERN LIST AT THE TOP
OF AZ SO THAT IF THE SAME FL SENTENCE IS PRESENTED ONCE MORE IMMEDIATELY AFTERWARDS, IT WILL BE MATCHED WITHOUT ANY BACKTRACKING USING THE SUBPATTERN(S) JUST CREATED.

PART 3.2 - PATTERN LISTS OF LINEAR FL SENTENCES.

WHEN PROCESSING LINEAR FL SENTENCES, WE ARE ONLY CONCERNED WITH TOP PATTERNS. PROCESSING ALL TOP PATTERNS WHICH GAVE A PARTIAL MATCH WOULD BE WASTEFUL, SO WE PROCESS THEM AS A STACK; LAST CREATED, FIRST CONSIDERED. HOWEVER IF A PATTERN IS FIRST CONSIDERED, AND ITS TRANSLATION IS JUST Z (NOTHING IN NL), THEN 'WAIT' IS PRINTED AND THE PATTERN IS INSERTED AT THE BOTTOM OF THE STACK FOR RECONSIDERATION LATER, IF NECESSARY. WE DO THIS BECAUSE WE MAY WELL FIND ANOTHER PATTERN WITH A TRANSLATION THAT CONTAINS SOME NL ELEMENTS AND IS CONSISTENT WITH THE INPUT.

PART 3.2.1 - MATCH-BACK.

THE MAJOR DIFFERENCE IN PROCESSING LINEAR OR NON-LINEAR FL SENTENCES IS THAT IF WE UPDATE THE VOCABULARY BY TRANSLATING AN FL COMPLEX PART IN THE CONTEXT OF A PATTERN ONLY, I.E. THE TRANSLATION RULE OF THE PATTERN WAS OF TYPE C (SEE ABOVE, CHAPTER II-9), THEN WE TRY TO SPLIT THE TRANSLATION USING EXACTLY THE SAME ROUTINES THAT WERE USED FOR THE INITIALIZATION OF THE SYSTEM. IF SUCCESSFUL, WE CREATE A NEW PATTERN WHICH WILL HAVE THE SAME SETS AND EXTRACTORS AS THE PATTERN CONSIDERED, BUT WHICH WILL HAVE A DIFFERENT (AND SIMPLER) TRANSLATION RULE.
THIS PROCESS MATCH-BACK. FOR AN EXAMPLE, SEE CHAPTER III. SENTENCE 31 PATTERN P1 IS OBTAINED FROM PATTERN P0 BY MATCH-BACK.

SINCE WE TRY TO MATCH AN FL SENTENCE TO PATTERNS IN THE REVERSE ORDER IN WHICH THEY WERE CREATED (LAST PATTERN CREATED, FIRST TRIED), AN FL SENTENCE CAN BE MATCHED AND TRANSLATED BY A PATTERN (OBTAINED BY MATCH-BACK) WHICH IS NEWER THAN THE PATTERN THAT WOULD OTHERWISE HAVE MATCHED THE FL SENTENCE. AS A RESULT, IT CAN HAPPEN THAT, EFFECTIVELY, SOME OF THE OLDER PATTERNS ARE NEVER REACHED ANY MORE. THIS RESULT CAN BE BENEFICIAL AS SOME OF THE EARLIER PATTERNS MAY HAVE INCORPORATED MISTAKES WHICH, THEREFORE, WILL NOT BE MADE ANY MORE. WE HAVE HERE A RATHER ELEMENTARY EXAMPLE OF WHAT MAY BE CALLED 'EVOLUTIONARY LEARNING'. AN EXAMPLE IN GERMAN APPEARS IN APPENDIX R.

PART 3.2.2 - 'TRY LEARN MORE'.

IF A LINEAR FL SENTENCE HAS BEEN TRANSLATED BY A PATTERN P1, AND IF THE SENTENCE WAS COMPLETELY MATCHED BY SOME OTHER PATTERN PJ WHICH HAD BEEN CREATED AFTER P1, THEN ZBIE TAKES THE FIRST SUCH PATTERN PJ THAT WAS CONSIDERED AND TRIES TO LEARN FROM PJ, AS IF THE SENTENCE HAD NOT BEEN TRANSLATED. ZBIE PRINTS 'TRY LEARN MORE' WHEN SUCH A CASE OCCURS. IT IS THE SIMPLE 'TRY LEARN MORE' PROCESS WHICH MAKES 'EVOLUTIONARY LEARNING' POSSIBLE. WE ONLY USE THIS PROCESS FOR LINEAR FL SENTENCES. FOR COMPLEX FL SENTENCES, CONTEXT IS ESSENTIAL, AND THE CONTEXT EXISTS MORE AT THE LEVEL OF THE IN-CONTEXT TRANSLATION THAN AT THE LEVEL OF THE
SET INCLUSION TESTS.

PART 4 - THE PATTERN CREATING ROUTINE.

FINALLY, WE MUST CONSIDER ANOTHER VERY IMPORTANT ROUTINE IN ZBIE: THE PATTERN CREATING ROUTINE. DEPENDING ON AN INPUT PARAMETER, THIS ROUTINE CREATES A TOP PATTERN OR A SUBPATTERN. WE SAW ABOVE HOW THE NEED FOR NEW SUBPATTERNS ARISES. ZBIE ATTEMPTS TO CREATE A TOP PATTERN WHEN ALL PREVIOUSLY DESCRIBED PROCESSING HAS FAILED.

THE PATTERN CREATING ROUTINE IS ONE OF THE LONGEST AND MOST COMPLICATED IN THE SYSTEM. ONLY ITS MAIN PARTS WILL BE DESCRIBED. IT TRIES TO CREATE A LIST OF PATTERNS WHICH WILL MATCH THE INPUT FL COMPLEX STRUCTURE FLI TOTALLY AND GIVE AS A TRANSLATION (AFTER SUCH A MATCH) THE ORIGINAL NL STRING NLI, MAKING USE OF ALREADY KNOWN INFORMATION.

THE ROUTINE MAKES FROM ONE TO FOUR PASSES ON FLI. EACH PASS IS FIRST PERFORMED ON ALL APPROPRIATE ELEMENTS OF FLI BEFORE THE NEXT PASS IS BEGUN. THE MAIN FEATURES OF EACH PASS ARE:

PART 4.1 - OUTLINE OF PATTERN CREATING ROUTINE.

1) FIND WHETHER AN FL UNIT (NOT A VERB) IN FLI HAS A TRANSLATION (IN SOME CONTEXT A, P) WHICH IS FOUND EXACTLY IN NLI. IF THAT IS THE CASE, AND FINALLY THE PATTERNS ARE CREATED, THEN THE SAME SET A WILL BE USED IN SAY SOME (SUB)PATTERN PJ, AND ON THE SET A WE PUT THE INFORMATION: IF WE WANT THE TRANSLATION OF AN ELEMENT IN
A in the context P, it is a good guess to assume that the translation is the same as in the context A, Pj and vice-versa. This is a case of the 'good guess' mentioned when discussing the consistency of a translation with an NL sentence.

2) Find whether an FL whole (not a verb) in FLI has a translation which 'looks like' some string in NLI. A test for such a similarity is made using the first few characters of the print names of the translation and of the elements in NLI. For instance, in Russian, the genitive of boy looks like the nominative of boy, (remember that, because of the FL statement, we expect something which has to do with 'boy'). Such a test would have to be improved to work for languages (such as Hebrew) where prefixes are added to words.

3) Treat verbs as in 1). (Verbs vary too much to be reliable at first).

4) If some FL whole in FLI had previously been found not to be translated (we had a partial translation rule, of type B, for instance), then assume that again it will not have to be translated.

After the application of a pass to an element of FLI, a check is made for certain terminating conditions: all FL wholes in FLI used up, all NL elements in NLI used up, or only one FL whole left unaccounted for. The system also checks for possible ambiguities. For example, suppose two different FL wholes in FLI
HAVE HAD IDENTICAL TRANSLATIONS, TO BE FOUND IN NLI; THERE IS NO WAY TO MAKE A CORRESPONDENCE, AND THE PATTERN CREATING ROUTINE EXITS WITH 'TOO HARD'.

THE NEXT STEP IS TO BUILD TRANSLATION RULES FOR THE SUBPATTERNS CREATED. AT THIS STAGE, ONLY TRANSLATION RULES OF TYPES A) AND B) ARE ALLOWED. THE RULES ARE BUILT BY CONSIDERING THE NL STRING INPUT TO THE ROUTINE AND MAKING SURE THAT THE TRANSLATION OF THE PATTERNS BUILT WHEN MATCHED TO THE INPUT FL
CHAPTER III.

LEARNING RUSSIAN.

We are presenting here the output of a computer run during which ZBIE attempts to learn Russian. The output is reproduced as obtained from the printer. Inputs in IPL-V form have been deleted except in one illustrative case.

The important attributes and values of a pattern PJ have the following meaning:

TO(PJ) = translation rule (a list structure).
PO(PJ) = a list of patterns with a p-list identical to PJ.
VO(PJ) = J4 if PJ is a subpattern.

The other attributes are used for bookkeeping and may be disregarded.

Comments are given at the right of inputs, or at the end of the processing of a sentence.

Looking at sentence comments

(be (man )here )
on zdes1
(be (man )there )
ON TAM
PROCESS START
PUT INTO VOCABULARY
(RE (MAN))
ON
P
PUT INTO VOCABULARY
HERE
ZDES1
A3
P
PUT INTO VOCABULARY
THERE
TAM
A3
P

NEW PATTERN

<table>
<thead>
<tr>
<th>&lt;PO&gt;</th>
<th>&lt;9-1 = 25418&gt;</th>
<th>&lt;9-2 = 25228&gt;</th>
<th>&lt;9-3 = 25524&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>02 9-1</td>
<td>04 0</td>
<td>04 0</td>
<td>04 0</td>
</tr>
<tr>
<td>00 A1</td>
<td>00 Y0</td>
<td>02 9-3</td>
<td>00 Y1</td>
</tr>
<tr>
<td>00 Y1</td>
<td>02 9-2</td>
<td>00 Y3</td>
<td>00 Y3</td>
</tr>
<tr>
<td>00 A2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00 Y2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00 A3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00 Y3</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
COMMENT: THE INITIALIZATION PHASE IS OVER. FOR THE NEXT SENTENCE, WE SHOW THE IPL-V INPUT.

COMMENT: 'PUT INTO VOCABULARY' HAS THREE OR FOUR ARGUMENTS. EXAMPLES OF BOTH CASES ARE ILLUSTRATED ABOVE:


ALL ADDITIONS TO THE VOCABULARY ARE PRINTED AS DESCRIBED ABOVE.

IP INPUT DATA

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>01</td>
</tr>
<tr>
<td>X4</td>
<td>95</td>
</tr>
<tr>
<td>95</td>
<td>0</td>
</tr>
</tbody>
</table>

IP BE

E0

IP (SPEAKING MAN)

91
IP HERE

IP SPEAKING

IP MAN

IP

IP

IP A1

IP ZDES1

IP

H2 = 2357 CELLS

LOOKING AT SENTENCE 3

(BE (SPEAKING MAN) HERE)

A1 ZDES1

PROCESS START

TRY PATTERN

P

NOT MATCHED.

TRY LEARN FROM PATTERN

P

(Z ZDES1)

PUT INTO VOCABULARY

(RE (SPEAKING MAN))

A1

P

PUT INTO VOCABULARY

(SPEAKING MAN)

THE INPUTS ARE FIRST PRINTED

THEN ZBIE STARTS PROCESSING.

(SPEAKING MAN) IS NOT IN A2.

Z STANDS FOR THE UNTRANSLATED

(BE(SPEAKING MAN)) FL COMPLEX.

THE RUSSIAN LETTER (=I).

PATTERN P1 IS BEING BUILT.
THE RUSSIAN LETTER (=I).

A2

A SET NAME.

P1

NEW PATTERN

<table>
<thead>
<tr>
<th>&lt;\textsuperscript{\textast}\textsuperscript{*} &gt;</th>
<th>&lt;9-1 = 25418&gt;</th>
<th>&lt;9-2 = 25752&gt;</th>
<th>&lt;9-3 = 25200&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>02 9-1</td>
<td>04 J</td>
<td>04 0</td>
<td>04 0</td>
</tr>
<tr>
<td>00 A1</td>
<td>00 P0</td>
<td>00 P0</td>
<td>00 Y2</td>
</tr>
<tr>
<td>00 Y1</td>
<td>02 9-2</td>
<td>00 Y3</td>
<td></td>
</tr>
<tr>
<td>00 A2</td>
<td>00 Y0</td>
<td>00 J3</td>
<td></td>
</tr>
<tr>
<td>00 Y2</td>
<td>00 J3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>00 A3</td>
<td>00 J4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>00 Y3</td>
<td>00 T0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>02 9-3</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

COMMENT: PATTERN P1 HAS A P-LIST IDENTICAL TO THE P-LIST OF P0. HOWEVER, THE TRANSLATION RULES ARE DIFFERENT. ZBIE USED (BE(MAN)) \rightarrow ON, AND (BE (SPEAKING MAN)) \rightarrow A1, BOTH TRANSLATIONS IN THE CONTEXT OF THE PATTERN P0, FOR MATCH BACK. THE COMPARISON OF THE NON-COMMON PARTS IN FL AND NL GAVE (SPEAKING MAN) \rightarrow A1 (IN THE CONTEXT A2, P1). SINCE THE NL STRINGS USED FOR MATCH-BACK HAVE NOTHING IN COMMON, '3E' IS ASSUMED NOT TO BE TRANSLATED, AND THE EXTRACTOR Y1 OF THE SET A1 (WHICH CONTAINS 'BE') IS NOT USED IN THE TRANSLATION RULE TO(P1). IT IS WELL WORTH COMPARING THIS PATTERN P1 TO THE PATTERN P1 CREATED BY ZBIE WHEN LEARNING SOME GERMAN (SEE APPENDIX B).
Z81E ALSO NOTES THAT IT IS A 'GOOD GUESS' TO ASSUME THAT THE
TRANSLATIONS OF A MEMBER OF A3 IN THE CONTEXTS (A3,P0) AND (A3,
P1) ARE THE SAME.

LOOKING AT SENTENCE 4
(BE (SPOKEN BOY )HERE )
T12 ZDES1
PROCESS START
TRY PATTERN
P1
NOT MATCHED.
TRY PATTERN
P
NOT MATCHED.
TRY LEARN FROM PATTERN
P1
(Z Z1)
GUESS
Z1
ZDES1
PUT INTO VOCABULARY
(SPOKEN BOY )
T12
A2
P1
(SPOKEN BOY) IS NOT IN SET A2.
'HERE' IS NOT KNOWN IN THE CONTEXT
(A3,P1), BUT IS GUESSED TO BE
THE SAME AS IN THE CONTEXT (A3,P0).
THE TRANSLATION IS CONSISTENT.
PUT INTO VOCABULARY
HERE
ZDES1
A3
P1

NOW 'HERE' IS ALSO KNOWN IN
THE CONTEXT (A3, P1).

A3 NOW 'HERE' IS ALSO KNOWN IN
THE CONTEXT (A3, P1).

LOOKING AT SENTENCE 5
(BE (MAN) HERE)
ON ZDES1
PROCESS START
TRY PATTERN
P1
PATTERN MATCHED. RESULT:
(Z ZDES1)
MAN IS NOT KNOWN IN
THE CONTEXT (A2, P1).

PATTERN MATCHED. RESULT:
(ON ZDES1)
TRY LEARN MORE
PATTERN P1 WAS MATCHED ENTIRELY.

TRY LEARN FROM PATTERN
P1
(Z ZDES1)
P1: "T INTO VOCABULARY
(MAN)
ON
A2
THE SENTENCE WILL BE TRANSLATED
LOOKING AT SENTENCE 6
(BE (SPEAKING MAN )MAN )
A1 MUGYINA
PROCESS START
TRY PATTERN
P1
'MAN' IS NOT IN THE SET A3.
'SINCE PO HAS THE SAME P-LIST AS
P1, THE MATCH ON PO IS NOT
ATTEMPTED.
NOT MATCHED.
TRY PATTERN
P
NOT MATCHED.
TRY LEARN FROM PATTERN
P1
(A1 Z )
PUT INTO VOCABULARY
MAN
MUGYINA
A3
P1
LOOKING AT SENTENCE 7
(BE (MAN )MAN )
ON MUGYINA
A NEW SENTENCE.
PROCESS START
TRY PATTERN
P1
PATTERN MATCHED. RESULT:
(ON HUGYINA) TRANSLATED COMPLETELY.

LOOKING AT SENTENCE 8
(BE (SPOKEN BOY)BOY)
T\A MAL\YIK
PROCESS START
TRY PATTERN
P1
NOT MATCHED.
TRY PATTERN
P
NOT MATCHED.
TRY LEARN FROM PATTERN
P1
(TI2 Z)
PUT INTO VOCABULARY
BOY
MAL\YIK
A3
P1

BUILDING UP THE VOCABULARY.

LOOKING AT SENTENCE 9
(BE (SPOKEN GIRL)GIRL)
T\A DEVYKA
PROCESS START
TRY PATTERN
(SPOKEN GIRL) IS NOT IN SET A2,
'GIRL' IS NOT IN SET A3.
P1
NOT MATCHED.
TRY PATTERN
P
NOT MATCHED.
TOO HARD

TWO MISTAKES AT ONE LEVEL WERE OBTAINED DURING PATTERN MATCHING, SO THAT NO PATTERN IS CLOSE TO THE FL SENTENCE. THE TWO FL UNITS (SPOKEN GIRL) AND 'GIRL' CANNOT BE TRANSLATED IN ANY CONTEXT (AT THIS STAGE), AND NO NEW PATTERN CAN BE BUILT TO TRANSLATE THE SENTENCE. NOTE THAT IF ZRIE HAD USED SOME OF THE SEMANTICS BUILT INTO FL TO GUESS THAT (SPOKEN GIRL) MAY WELL HAVE A TRANSLATION IDENTICAL TO (SPOKEN BOY), IN THE SAME CONTEXTS, THEN THE SENTENCE COULD HAVE BEEN PROCESSED SUCCESSFULLY AT THIS STAGE. THE GUESSED TRANSLATION WOULD HAVE BEEN 'T12 7', WHICH IS CONSISTENT WITH THE INPUT.

WE SHALL HEREAFTER LEAVE OUT THE MONOTONOUS 'TRY PATTERN ...
PJ ... NOT MATCHED ...'.

LOOKING AT SENTENCE 10
(BE BOY )
E2TO MAIYIK
PUT INTO VOCABULARY
BE
E2TO
A NEW SET.

PUT INTO VOCABULARY

MALIYIK

THE SAME TRANSLATION AS

IN THE CONTEXT (A3,P1).

NEW PATTERN

<table>
<thead>
<tr>
<th>&lt;P2&gt;</th>
<th>&lt;9-1 = 26178&gt;</th>
<th>&lt;9-2 = 26170&gt;</th>
<th>&lt;9-3 = 2618&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>02 9-1</td>
<td>04 0</td>
<td>04 0</td>
<td>04 0</td>
</tr>
<tr>
<td>00 A4</td>
<td>00 Y1</td>
<td>00 Y1</td>
<td></td>
</tr>
<tr>
<td>00 Y1</td>
<td>02 9-2</td>
<td>00 Y2</td>
<td>00 Y2</td>
</tr>
<tr>
<td>00 A3</td>
<td>00 D12</td>
<td>00 Y2</td>
<td></td>
</tr>
<tr>
<td>00 Y2</td>
<td>02 9-3</td>
<td>00 D13</td>
<td>04 2618</td>
</tr>
</tbody>
</table>

COMMENTS: THE SENTENCE IN FL HAS A LENGTH OF TWO. PREVIOUS PATTERNS EXPECT AN FL SENTENCE OF LENGTH THREE. TO BUILD THE NEW PATTERN, ZBIE TRIED TO FIND TRANSLATIONS, FOR THE FL UNITS IN THE FL SENTENCE, WHICH WERE IDENTICAL OR CLOSE TO NL STRINGS IN THE INPUT NL. 'BE' HAD NO TRANSLATION; HOWEVER, 'BOY' HAD A TRANSLATION 'MALLIK', IN THE CONTEXT (A3, P1) (SEE SENTENCE A), WHICH IS FOUND IN THE INPUT. THE FL UNIT 'BE' WHICH HAD NOT BEEN ACCOUNTED FOR IS MATCHED TO THE REMAINING PARTS OF THE NL INPUT, NAMELY 'E2TO'. ZBIE ALSO MAKES A NOTE THAT IT IS A 'GOOD GUESS'
TO ASSUME THAT THE TRANSLATIONS OF AN FL UNIT IN THE CONTEXTS (A3,P2) AND (A3,P1) ARE IDENTICAL.

LOOKING AT SENTENCE 11
(BE FOOT)
E2TO NOG1A
BUILDING UP THE VOCABULARY.

PROCESS START
TRY LEARN FROM PATTERN
P2
(E2TO Z)
PUT INTO VOCABULARY
FOOT
NOG1A
A3
P2

LOOKING AT SENTENCE 12
(BE FOOT (OF BOY))
NOT A LINEAR FL SENTENCE.

E2TO NOG1A MAILYIKA

PROCESS START
TRANSLATE
P2
RESULT:
(E2TO Z)
Z STANDS FOR THE TRANSLATION OF
THE FL COMPLEX FOOT (OF BOY), AND,
TRY
BY THE CONSISTENCY TEST, 7 IS TO BE
P2
REPLACED BY 'NOG1A MAILYIKA'.
PUT INTO VOCABULARY BOY MALIYIKA A6 A DIFFERENT TRANSLATION, P38 USING A DIFFERENT SET.

PUT INTO VOCABULARY FOOT NOGIA A3 THE SAME TRANSLATION AS IN THE CONTEXT (A3,P2).


NEW PATTERN

<table>
<thead>
<tr>
<th>&lt;P38&gt;</th>
<th>&lt;9-1 = 26474&gt;</th>
<th>&lt;9-2 = 26766&gt;</th>
<th>&lt;9-3 = 26646&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>02 9-1</td>
<td>04 0</td>
<td>04 0</td>
<td>04 0</td>
</tr>
<tr>
<td>00 A5</td>
<td>00 T0</td>
<td>00 Y100</td>
<td>00 Y100</td>
</tr>
<tr>
<td>00 Y101</td>
<td>02 9-2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>00 A6</td>
<td>00 D12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>00 Y100</td>
<td>02 9-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>00 D13</td>
<td>04 26646</td>
<td></td>
<td></td>
</tr>
<tr>
<td>00 D9</td>
<td>00 A7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>00 V0</td>
<td>00 J4</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
<P39>  <9-1 = 26454>  <9-2 = 26788>  <9-3 = 24899>

02 9-1  04 0  04 0  04 0
00 A7   00 T0  00 Y98 00 Y98
00 Y99  02 9-2 00 Y99 00 Y99
00 A3   00 D12
00 Y98  02 9-3
00 D13  04 26678
00 V0   00 J4

PUT INTO SET
P39
A3


THE TRANSLATIONS IN PARALLEL OF PATTERN LISTS HAVING MATCHED THE FL SENTENCE TO THE DEEPEST MATCH-DEPTH ARE TRIED. THIS ACTION FOLLOWS 'TRANSLATE'. HERE, ONLY ONE PATTERN, P2, IS FOUND. THEN, THE PATTERN LISTS WITH A CONSISTENT TRANSLATION ARE TRIED. THIS ACTION FOLLOWS 'TRY'. HERE WE ONLY 'TRY' THE PATTERN P2. BY THE CONSISTENCY TEST, THE FL COMPLEX FOOT (OF BOY) CORRESPONDS TO THE NL STRING 'NOG1A MAL1YIKA', AND THE PATTERN CREATING ROUTINE
MANAGES TO BUILD SURPATTERNS TO MATCH THE FL COMPLEX. NOTE THAT PREVIOUS KNOWLEDGE HAS BEEN USED TO AVOID BUILDING A WHOLE NEW TUP PATTERN.

WHEN BUILDING THE SURPATTERNS, THE TRANSLATION OF 'FOOT' IN THE CONTEXT (A3,P2) IS USED. THE PREVIOUSLY ENCOUNTERED TRANSLATION OF 'BOY', 'MAL1YIK', IS NOT FOUND IN THE NL STRING. HOWEVER, 'MAL1YIK' IS VERY CLOSE TO 'MAL1YIKA', SO THAT THE TRANSLATION OF 'BOY', IN A NEW CONTEXT, IS ASSUMED TO BE 'MAL1YIKA'. SINCE ALL THE WORDS IN THE NL STRING 'NOGIA MAL1YIKA' HAVE BEEN ACCOUNTED FOR, 'OF' IS ASSUMED NOT TO BE TRANSLATED.


FOLLOWED BY THE TRANSLATION OF THE FL COMPLEX (OF BOY) MATCHED TO P38, (Y99). THE TRANSLATION RULE OF P38 CALLS FOR THE TRANSLATION OF THE ELEMENT MATCHED TO A6, 'BOY', IN THE CONTEXT (A6,P39), NAMELY 'MALIYIKA', (Y130). THE TRANSLATION ROUTINE IS SEEN TO CALL ON ITSELF RECURSIVELY.


LOOKING AT SENTENCE 13
(BE HAND (OF BOY)
E2 TO RUKA MALIYIKA
PROCESS START
TRANSLATE
P2
P39
P38
RESULT
'HAND' IS NOT IN SET A3.

TRY P39 P36
PUT INTO VOCABULARY
HAND RUKA A3 P39

LOOKING AT SENTENCE 14
(BE HAND )
E2TO RUKA
PROCESS START
TRY PATTERN
P2
PATTERN MATCHED. RESULT:
(E2TO Z )
TRY LEARN FROM PATTERN
P2
(E2TO 7 )
GUESS BUT 'HAND' CAN BE GUESSED, USING
Z
RUKA

LOOKING AT SENTENCE 15
(BE BOOK )
E2TO KNIGIA
PROCESS START
TRY LEARN FROM PATTERN P2
(E2TO Z )
PUT INTO VOCABULARY BOOK
KNIGIA A3 P2
LOOKING AT SENTENCE 16 (Q de BOOK WHERE )
G1DE KNIGIA PROCESS START
PUT INTO VOCABULARY WHERE G1DE A10 P3
PUT INTO VOCABULARY BOOK
KNIGIA A3 P3
MORE VOCABULARY.
Q MEANS QUESTION; IT IS A MARKER, AND IS NOT TO BE TRANSLATED. BUILDING PATTERN P3.
COMMENT: NOTICE THE TRANSLATION RULE.

LOOKING AT SENTENCE 17
(BE (IN (BOOK) HAND))

ONA V RUKE
PROCESS START
TRANSLATE
P2
RESULT:
(E2 TO Z) NOT CONSISTENT.
TOO HARD

BOTH (BOOK) AND 'IN' ARE UNKNOWN, AND A NEW TOP PATTERN COULD NOT BE CREATED. ZBIE WILL NOW LEARN (BOOK), THEN RETURN TO THE SAME SENTENCE. 'HAND' COULD BE GUESSED FROM 'RUKA' TO 'RUKE'
LOOKING AT SENTENCE 18
(BE (BOOK) HERE)
ONA ZDES1
PROCESS START
TRY LEARN FROM PATTERN
P1
(Z ZDES1)
PUT INTO VOCABULARY
(BOOK)
ONA
A2
P1

NOW (BOOK) IS KNOWN AS 'ONA'
IN THE CONTEXT (A2,P1).

LOOKING AT SENTENCE 19
(BE (IN (BOOK) HAND))
ONA V RUKE
PROCESS START
TRANSLATE
P2
RESULT
(E2 TO Z)
PUT INTO VOCABULARY
HAND
RUKE
A13

IDENTICAL TO SENTENCE 17.

NOT CONSISTENT.
BUILDING PATTERN P4.

A NEW SET.
P37
PUT INTO VOCABULARY
IN
V
A12
P37
PUT INTO VOCABULARY
(BOOK)
ONA
A2
P37
PUT INTO SET
P37
A14
NEW PATTERN

<table>
<thead>
<tr>
<th>&lt;P37&gt;</th>
<th>&lt;9-1 = 26506&gt;</th>
<th>&lt;9-2 = 27334&gt;</th>
<th>&lt;9-3 = 271A3&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>02 9-1</td>
<td>04 0</td>
<td>04 0</td>
<td>04 0</td>
</tr>
<tr>
<td>00 A12</td>
<td>00 0</td>
<td>00 Y96</td>
<td>00 Y96</td>
</tr>
<tr>
<td>00 Y97</td>
<td>02 9-2</td>
<td>00 Y97</td>
<td>00 Y97</td>
</tr>
<tr>
<td>00 A2</td>
<td>00 D12</td>
<td>00 Y95</td>
<td>00 Y95</td>
</tr>
<tr>
<td>00 Y96</td>
<td>02 9-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>00 A13</td>
<td>00 D13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>00 Y95</td>
<td>04 27144</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>00 D9</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>00 A14</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>00 V0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Comment: Patterns P2 and P4 both have two sets on their P-lists, but these sets and the translation rules of the patterns are quite different. The translation rule of P37 is worth noting.

Looking at sentence 20
(BE (IN (BOOK) HAND (OF BOY))

ONA V RUKE MALYIKA

Process start
Translate
P4
P37

Result:
(ONA V Z)

A consistent translation.

Try
P4
P37

RUKE MALYIKA.
PUT INTO VOCABULARY HAND
PUT INTO VOCABULARY RUKE
PUT INTO VOCABULARY A13
PUT INTO VOCABULARY P36
PUT INTO VOCABULARY BOY
PUT INTO VOCABULARY MALIYIKA
PUT INTO VOCABULARY A6
PUT INTO SET P35
PUT INTO SET P35
PUT INTO SET A16

NEW PATTERN

<table>
<thead>
<tr>
<th>&lt;P35&gt;</th>
<th>&lt;9-1 = 27532&gt;</th>
<th>&lt;9-2 = 27776&gt;</th>
<th>&lt;9-3 = 2764&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>02 9-1</td>
<td>04 0</td>
<td>04 0</td>
<td>04 0</td>
</tr>
<tr>
<td>00 A15</td>
<td>00 T0</td>
<td>00 Y93</td>
<td>00 Y93</td>
</tr>
<tr>
<td>00 Y94</td>
<td>02 9-2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>00 A6</td>
<td>00 D12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>00 Y93</td>
<td>02 9-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>00 D13</td>
<td>04 27674</td>
<td></td>
</tr>
<tr>
<td></td>
<td>00 D9</td>
<td>00 A16</td>
<td>00 V0</td>
</tr>
<tr>
<td></td>
<td>00 J4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P36</td>
<td>9-1 = 2665A</td>
<td>9-2 = 877A6</td>
<td>9-3 = 2772D</td>
</tr>
<tr>
<td>-------</td>
<td>-------------</td>
<td>-------------</td>
<td>-------------</td>
</tr>
<tr>
<td>02 9-1</td>
<td>04 0</td>
<td>04 0</td>
<td>04 0</td>
</tr>
<tr>
<td>00 A16</td>
<td>00 T0</td>
<td>00 Y91</td>
<td>00 Y91</td>
</tr>
<tr>
<td>00 Y92</td>
<td>02 9-2</td>
<td>00 Y92</td>
<td>00 Y92</td>
</tr>
<tr>
<td>00 A13</td>
<td>00 D12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>00 Y91</td>
<td>02 9-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>00 D13</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>04 27706</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>00 V0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>00 J4</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Looking at sentence 21

(Be pencil)

E2 to Karandaw

Process start

Try learn from pattern

P4

When the translation is just Z,

(Z)

We wait the first time around.

Wait

Try learn from pattern

P2

(E2 to Z)

Consistent translation.

Put into vocabulary

Pencil

Karandaw

A3
IF WE HAD PROCESSED P4 FIRST, WE WOULD HAVE OBTAINED:

PUT INTO VOCABULARY
PENCIL
E2TO KARANDAW
A14
P4

LOOKING AT SENTENCE 22
(Q WE PENCIL WHERE)
G1IDE KARANDAW
PROCESS START
TRY PATTERN 'PENCIL' IS NOT KNOWN IN THE CONTEXT (A3,P3).
P3
(G1IDE Z )
TRY LEARN FROM PATTERN 'PENCIL' CAN BE GUESSED FROM THE CONTEXT (A3,P2) TO THIS CONTEXT (A3,P3).
P3
(G1IDE Z )
GUESS Z
KARANDAW
'PUT INTO VOCABULARY
PENCIL
KARANDAW
A3
P3
IF 'GOOD GUESSES' HAD BEEN USED, THIS SENTENCE WOULD HAVE BEEN TRANSLATED.

LOOKING AT SENTENCE 23
(BE (PENCIL THERE )
ON TAM
PROCESS START
TRY LEARN FROM PATTERN
P1
(Z Z1 )
THE GUESS MAKES THE TRANSLATION CONSISTENT WITH THE INPUT.
Z1
TAM
PUT INTO VOCABULARY
(PENCIL )
ON
A2
P1
MORE VOCABULARY.
PUT INTO VOCABULARY
THERE
TAM
A3
P1
LOOKING AT SENTENCE 24
(BE (IN (PENCIL )DRAWER ))
ON V A1W2IKE
PROCESS START
TRANSLATE
P4
P37
RESULT:
(Z V Z1 )
GUESS
Z
ON
TRY
P4
P37
PUT INTO VOCABULARY
(PENCIL )
ON
A2
P37
PUT INTO VOCABULARY
DRAWER
A1W2IKE
A13
P37

LOOKING AT SENTENCE 25
(BE BOY (MOD THIS )HERE )

EVEN WITHOUT THE GUESS, THE
TRANSLATION IS CONSISTENT WITH
THE NL INPUT.
**EDIT HALYIK ZDES1**

**PROCESS START**

**TRANSLATE**

P1

RESULT1

(Z ZDES1 )

**TRY**

P1

**PUT INTO VOCABULARY**

BUILDING SUBPATTERNS P34 AND P33.

**PUT INTO VOCABULARY**

**BOY**

**HALYIK**

A3

P34

**PUT INTO SET**

P33

A19

**NEW PATTERN**

<table>
<thead>
<tr>
<th>P33</th>
<th>&lt;9-1 = 27360&gt;</th>
<th>&lt;9-2 = 27812&gt;</th>
<th>&lt;9-3 = 27800&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>02</td>
<td>04 0</td>
<td>04 0</td>
<td>04 0</td>
</tr>
<tr>
<td>00</td>
<td>Y90</td>
<td>00 Y90</td>
<td>00 Y85</td>
</tr>
<tr>
<td>00</td>
<td>04 0</td>
<td>04 0</td>
<td>04 0</td>
</tr>
</tbody>
</table>
PUT INTO SET
P34
A2

LOOKING AT SENTENCE 26
(BE GIRL )
E2 TO DEVOYKA
PROCESS START
TRY LEARN FROM PATTERN
P4
(Z )
WAIT
TRY LEARN FROM PATTERN
P2
(E2TO Z )
PUT INTO VOCABULARY
GIRL
DEVOYKA
A3
P2
MORE VOCABULARY.

LOOKING AT SENTENCE 27
(BE (SPOKEN GIRL )GIRL )
T12 DEVOYKA
THIS IS SENTENCE 9, WHICH HAD
PROCESS START
BEEN FOUND 'TOO HARD' PREVIOUSLY.
TRY LEARN FROM PATTERN
P1
(Z Z1 )
GUESS
Z1
DEVOYKA
PUT INTO VOCABULARY
(SPOKEN GIRL )
T12
PUT INTO VOCABULARY
GIRL
DEVOYKA
A3
P1

LOOKING AT SENTENCE 28
(BE GIRL (MOD THIS) HERE)
E2TA DEVOYKA ZDES1
PROCESS START
MATCHED PATTERN-LIST
P1
P34
P33
RESULTI
(E2TOT Z ZDES1 )
TRANSLATE
P1
P34
RESULTI
(Z Z1 ZDES1 )
GUESS
WITH THE GUESS, THE TRANSLATION
Z1
DEVOYKA
IS CONSISTENT WITH THE NL INPUT.
TRY
P1
P34
PUT INTO VOCABULARY
THIS
E2TA
A21
P32
NEW PATTERN
NOT INSERTED
TRANSLATE
P1
RESULT:
(Z ZDES1 )
TRY
P1
PUT INTO VOCABULARY
GIRL
DEVOYKA
A3
P31
PUT INTO VOCABULARY
THIS
E2TA
A21
P30
SEE COMMENTS BELOW.
PATTERN LIST MATCHING THE FL
SENTENCE TO A MATCH-DEPTH OF 0.
CONSISTENT WITH INPUT. Z TAKES
THE PLACE OF GIRL(MOD THIS).
BUILDING SUBPATTERNS P30 AND P31.
PUT INTO SET
P30
A23
NEW PATTERN

<table>
<thead>
<tr>
<th>&lt;P30&gt;</th>
<th>&lt;9-1 = 28520&gt;</th>
<th>&lt;9-2 = 28752&gt;</th>
<th>&lt;9-3 = 28650&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>02 9-1</td>
<td>04 0</td>
<td>04 0</td>
<td>04 0</td>
</tr>
<tr>
<td>00 A22</td>
<td>00 T0</td>
<td>00 YA3</td>
<td>00 YB3</td>
</tr>
<tr>
<td>00 Y84</td>
<td>02 9-2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>00 A21</td>
<td>00 D12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>00 Y83</td>
<td>02 9-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>00 D13</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>04 28636</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>00 D9</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>00 A23</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>00 V0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>00 J4</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>&lt;P31&gt;</th>
<th>&lt;9-1 = 28392&gt;</th>
<th>&lt;9-2 = 28762&gt;</th>
<th>&lt;9-3 = 28682&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>02 9-1</td>
<td>04 0</td>
<td>04 0</td>
<td>04 0</td>
</tr>
<tr>
<td>00 A23</td>
<td>00 T0</td>
<td>00 YA2</td>
<td>00 YB2</td>
</tr>
<tr>
<td>00 Y82</td>
<td>02 9-2</td>
<td>00 YB1</td>
<td>00 YB1</td>
</tr>
<tr>
<td>00 A3</td>
<td>00 D12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>00 Y81</td>
<td>02 9-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>00 D13</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>04 28668</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>00 V0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>00 J4</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Put into set
P31
A2

Here is our first encounter with Zbie's look-ahead capabilities. The FL sentence if first matched to a match-depth of 2 by the pattern-list (P1 P34 P33) but the translation is not consistent with the NL input, since the first NL words of the translation and input are different. Next, the FL sentence is matched to a match-depth of 1 by the pattern-list (P1 P34). The FL structure \[\text{mod this}\] has no corresponding subpattern to which it is then matched and contributes the unknown \(z_i = z_0\) to the translation of the FL sentence. 'Girl' is not known in the context (A3, P34), and contributes the unknown \(z_1\). Since the translation of 'Girl' can be correctly guessed, the translation of the FL sentence is consistent with the input. The unknown \(z\) takes the place of the FL structure \[\text{mod this}\] and, by the consistency test, is to be replaced by the FL sentence is first matched to a match-depth of 2 by the pattern-list (P1 P34 P33), but the translation is not consistent with the NL input since the first NL words of the translation and input are isomorphic to P32. (We would only require that P32 be a homomorphic image of a subpattern in A19). P32 is 'not inserted' in A19 and the next pattern list is tried. Sentence 3a will provide another example of the same techniques.
LOOKING AT SENTENCE 29
(BE TREE )
E2TO DEREVO
PROCESS START
TRY LEARN FROM PATTERN
P4
(Z )
WAIT
TRY LEARN FROM PATTERN
P2
(E2TO Z )
PUT INTO VOCABULARY
TREE
DEREVO
A3

MORE VOCABULARY BEFORE
THE NEXT SENTENCE.

LOOKING AT SENTENCE 30
(BE TREE (MOD THIS )HERE )
E2TO DEREVO ZDES1
PROCESS START
MATCHED PATTERN-LIST
PUT INTO VOCABULARY
TREE
DEREVO
A3

MORE VOCABULARY BEFORE
P2

THE NEXT SENTENCE.

LOOKING AT SENTENCE 30
(BE TREE (MOD THIS ITEM))
E2TO DEREVO ZDES1

PROCESS START

MATCHED PATTERN-LIST

P1

'TREE' IS NOT KNOWN IN THE
CONTEXT (A3,P31), BUT CAN BE
GUESSED FROM THE CONTEXT
RESULTI

(A3,P2).

(E2TA Z ZDES1 )

NOT CONSISTENT.

MATCHED PATTERN-LIST

P1

'TREE' IS NOT KNOWN IN THE
CONTEXT (A3,P34), BUT CAN BE
GUESSED FROM THE CONTEXT
RESULTI

(A3,P2).

(E2TO Z ZDES1 )

NOT CONSISTENT.

TRANSLATE

PATTERN-LIST WITH

MATCH-DEPTH OF 1.

P31

RESULTI

(Z Z1 ZDES1 )

CONSISTENT AFTER GUESS.

GUESS

Z1

DEREVO
TRANSLATE
P1
P34
RESULT:
(Z2 Z3 ZDES1)
GUESS
Z3
DEREVO
TRY
P1
P31
PUT INTO VOCABULARY
THIS
E2TO
A25
P29
NEW PATTERN

\[
\begin{array}{c}
\langle P29 \rangle \\
02 \, 9-1 \\
00 \, A24 \\
00 \, Y80 \\
00 \, A25 \\
00 \, Y79 \\
\langle 9-1 = 29028 \rangle \\
04 \, 0 \\
00 \, T0 \\
02 \, 9-2 \\
00 \, D12 \\
02 \, 9-3 \\
00 \, D13 \\
04 \, 29142 \\
00 \, V0 \\
\langle 9-2 = 29154 \rangle \\
04 \, 0 \\
00 \, Y79 \\
\langle 9-3 = 29154 \rangle \\
04 \, 0 \\
00 \, Y79 \\
\end{array}
\]
00 J4

NOT INSERTED

TRY

P1

P34

PUT INTO VOCABULARY

THIS

E2T0

A25

P28

NEW PATTERN

\[ \langle P28 \rangle = \langle 9-1 = 29148 \rangle, \langle 9-2 = 29258 \rangle, \langle 9-3 = 29272 \rangle \]

\[ 02 9-1 04 0 04 0 04 0 \]

\[ 00 A26 00 T0 00 Y77 00 Y77 \]

\[ 00 Y78 02 9-2 00 Y77 00 Y77 \]

\[ 00 A25 00 D12 \]

\[ 00 Y77 02 9-3 00 D13 \]

\[ 04 D13 04 29122 \]

\[ 00 V0 \]

\[ 00 J4 \]

NOT INSERTED

P28 IS ISOMORPHIC TO P33.

THE NEXT PATTERN-LIST.

BUILDING SURPATTERN P28.

TRANSLATE PATTERN-LIST WITH MATCH-DEPTH OF 0.

RESULT:

\( (Z \ ZDES1 \ ) \)
TRY
P1
TREE DEREVO A3 P27
PUT INTO VOCABULARY THIS E2TO 'E2TO' WAS ALSO THE TRANSLATION A25 OF 'BE' IN THE CONTEXT (A4, P2).
P26
PUT INTO SET P26 A28
NEW PATTERN

<table>
<thead>
<tr>
<th>&lt;P26&gt;</th>
<th>&lt;9-1 = 29082&gt;</th>
<th>&lt;9-2 = 29476&gt;</th>
<th>&lt;9-3 = 2941&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>02</td>
<td>04 0</td>
<td>04 0</td>
<td>04 0</td>
</tr>
<tr>
<td>00</td>
<td>A27</td>
<td>00 T0</td>
<td>00 Y75</td>
</tr>
<tr>
<td>00</td>
<td>Y76</td>
<td>02 9-2</td>
<td></td>
</tr>
<tr>
<td>00</td>
<td>A25</td>
<td>00 D12</td>
<td></td>
</tr>
<tr>
<td>00</td>
<td>Y75</td>
<td>02 9-3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>00 D13</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>04 29398</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>00 D9</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>00 A28</td>
<td></td>
</tr>
</tbody>
</table>
PUT INTO SET
P27
A2

LOOKING AT SENTENCE 31
(BE BOY [MOD THAT ]THERE )
TOT MALIYIK TAM

PROCESS START

TRANSLATE PATTERN-LISTS (P1 P27 P26),
(P1 P31 P30) AND (P1 P34 P33) ALL
MATCH THE FL SENTENCE TO A
MATCH-DEPTH OF 2, AND ARE
RESULT: TRANSLATED IN PARALLEL.
(Z Z1 TAM )
CONSISTENT AFTER GUESS.
GUESS
Z1
MAL1YIK
TRANSLATE
P1
P31
P30
RESULT
(Z2 Z3 TAM )
GUESS
Z3
MAL1YIK
TRANSLATE
P1
P34
P33
RESULT
(Z4 MAL1YIK TAM )
CONSISTENT AFTER GUESS.
CONSISTENT WITH INPUT.
THE TRANSLATION OF THE PATTERN-
LIST (P1 P34 P33) IS CLOSEST TO
THE NL INPUT, AND (P1 P34 P33)
IS TRIED FIRST.
PUT INTO VOCABULARY
THAT
TOT
A18
IF ZBIE HAD PROCESSED THE PATTERN-LISTS, STARTING WITH THOSE OF DEEPEST MATCH-DEPTH, IN THE ORDER IN WHICH THEY ARE CONSIDERED, THEN ZBIE WOULD HAVE STARTED PROCESSING THE PATTERN-LIST (P1 P27 P26). ZRJE WOULD HAVE CREATED THE FOLLOWING VOCABULARY ENTRIES:

THAT TOT A25 P26
BOY MAL1YIK A3 P27

THE LAST ENTRY WOULD NOW ASSURE THE (INCORRECT) TRANSLATION OF (BE BOY|MOD THIS) HERE) AS 'E2TO MAL1YIK ZDES1'. THE NEXT EXAMPLE IS A SLIGHT VARIATION OF THIS ONE.

LOOKING AT SENTENCE 32
(BE GIRL (MOD THAT )THERE )
TA GEVOYKA TAM
PROCESS START
MATCHED PATTERN LIST
P1 'GIRL' IS NOT KNOWN IN THE CONTEXT (A3,P34).
P34
P33
RESULTI
(TOT Z TAM ) NOT CONSISTENT.
TRANSLATE PATTERN-LISTS
P1 (P1 P27 P28) AND (P1 P31 P30)
P27 IN PARALLEL.
RESULT1
(Z Z1 TAM ) CONSISTENT AFTER GUESS.
GUESS
Z1
DEVOYKA
TRANSLATE
P1
P31
P30
RESULT1
(Z2 DEVOYKA TAM ) CONSISTENT.
TRY
PATTERN-LIST (P1 P31 P30) GIVES
P1
A TRANSLATION CLOSEST TO THE
P31
NL INPUT.
P30
PUT INTO VOCABULARY
THAT
TA
A21
P30

LOOKING AT SENTENCE 33
(BE HAND (OF (SPEAKING MAN )))
E2TO MOA1 RUKA
PROCESS START
TRANSLATE
P2 MATCH-DEPTH OF 2.
P39
P38
RESULT:
(E2TO RUKA Z )
TRANSLATE
P2 MATCH-DEPTH OF 1.
P39
RESULT:
(E2TO RUKA Z )
TRANSLATE
P4 AND P2 (MATCH-DEPTH 0).
RESULT:
(Z )
TRANSLATE
P2
RESULT:
(E2TO Z1 )
TRY
P2
PUT INTO VOCABULARY
(SPEAKING MAN )
HOA1
A30
P24
PUT INTO VOCABULARY
THE TRANSLATION DUE TO P2 IS
CLOSEST TO THE NL INPUT.
<table>
<thead>
<tr>
<th>&lt;P24&gt;</th>
<th>&lt;9-1 = 29396&gt;</th>
<th>&lt;9-2 = 27938&gt;</th>
<th>&lt;9-3 = 281Ny&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>02 9-1</td>
<td>04 0</td>
<td>04 0</td>
<td>04 0</td>
</tr>
<tr>
<td>00 A29</td>
<td>00 T0</td>
<td>00 Y71</td>
<td>00 Y71</td>
</tr>
<tr>
<td>00 Y72</td>
<td>02 9-2</td>
<td>00 D12</td>
<td></td>
</tr>
<tr>
<td>00 A30</td>
<td>00 D13</td>
<td>04 29418</td>
<td></td>
</tr>
<tr>
<td>00 Y71</td>
<td>02 9-3</td>
<td>00 D9</td>
<td>00 A31</td>
</tr>
<tr>
<td></td>
<td>00 V0</td>
<td>00 J4</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>&lt;P25&gt;</th>
<th>&lt;9-1 = 27180&gt;</th>
<th>&lt;9-2 = 29464&gt;</th>
<th>&lt;9-3 = 2861~&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>02 9-1</td>
<td>04 0</td>
<td>04 0</td>
<td>04 0</td>
</tr>
<tr>
<td>00 A31</td>
<td>00 T0</td>
<td>00 Y70</td>
<td>00 Y70</td>
</tr>
<tr>
<td>00 Y70</td>
<td>02 9-2</td>
<td>00 Y69</td>
<td>00 Y69</td>
</tr>
<tr>
<td>00 A3</td>
<td>00 D12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>00 Y69</td>
<td>02 9-3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
PUT INTO SET
P25
A3

LOOKING AT SENTENCE 34
(BE HAND [OF (SPOKEN BOY)])
E2TO TVOA1 RUKA

PROCESS START
TRANSLATE

TRANSLATE IN PARALLEL
PATERN LISTS (P2 P25 P24) AND
(P2 P39 P38) WHICH MATCHED
THE FL SENTENCE TO A DEPTH OF 2.
RESULT:
(E2TO Z RUKA )
CONSISTENT.

TRANSLATE
P2
P39
P38
RESULT:
(E2TO RUKA Z1 )
NOT CONSISTENT.

TRY
P2
LOOKING AT SENTENCE 35
(BE (ON (BOOK )HAND ))
ONA NA RUKE
PROCESS START
TRANSLATE
P4
P37
RESULT:
(ONA Z RUKE )
TRY
P4
P37
PUT INTO VOCABULARY
ON
NA
A12
P37

COMMENT: THIS SLIGHTLY UNNATURAL SENTENCE WAS ADDED TO
AFFORD A SMOOTH TRANSITION TOWARDS THE NEXT SENTENCE WITHOUT BUILDING A NEW SUBPATTERN. SEE THE COMMENTS AT THE END OF SENTENCE 36.

LOOKING AT SENTENCE 36

(BE (ON (BOOK )TABLE ))

ONA NA STOLE

PROCESS START

TRANSLATE

P4

P37

RESULT:

(ONA NA Z )

'TABLE' IS NOT KNOWN IN THE

TRY

CONTEXT (A13, P37).

P37

PUT INTO VOCABULARY

TABLE

STOLE

A13

P37

ASSUME THAT: A) THIS SENTENCE HAD BEEN PRESENTED BEFORE SENTENCE 35; B) 'TABLE' WAS KNOWN IN SOME CONTEXT, FOR INSTANCE AS 'STOL' (NOMINATIVE); AND C) 'ON' WAS NOT KNOWN (WHICH WAS THE CASE HERE). THEN ZBIE WOULD HAVE BUILT A NEW SUBPATTERN BY: A) GUESSING 'TABLE' FROM THE PREVIOUS PRINT-NAME; B) PAIRING (BOOK)
WITH ONE ACCEPTABLE TRANSLATION, 'ONA', OF (BOOK); C) ASSIGNING 'ON' TO THE STILL UNACCOUNTED FOR NL STRING 'NA'. THE NEW SUBPATTERN WOULD HAVE BEEN INSERTED IN THE SET A11, ON THE P-LIST OF P4.

ACTUALLY, IT IS NOT NECESSARY TO BUILD A NEW SUBPATTERN. THE NEW FL SENTENCE FITS INTO THE PREVIOUS PATTERN STRUCTURE IF WE ARE CAREFUL TO GIVE ENOUGH INTERMEDIARY SENTENCES, SUCH AS SENTENCE 35.

ZBIE TENDS TO GENERATE TOO MANY SETS AND PATTERNS. IT WOULD BE INTERESTING TO GIVE ZBIE THE CAPABILITY TO RE-ORGANIZE ITS STRUCTURES, FOR EXAMPLE BY MERGING SOME SETS OR SOME PATTERNS.

LOOKING AT SENTENCE 37
(FUTURE TAKE (SPEAKING MAN) BOOK)
A1 VOZIMU KNIG1U 'FUTURE' IS A MARKER IN FL, AND
PROCESS START IS NOT TO BE TRANSLATED.
PUT INTO VOCABULARY BUILDING PATTERN P5.
BOOK
KNIG1U
A34
P5
PUT INTO VOCABULARY
TAKE
VOZIMU
A33
**PUT INTO VOCABULARY**

(SPEAKING MAN)

A1 THE RUSSIAN LETTER.

A2 THE SET NAME.

P5 NEW PATTERN

<table>
<thead>
<tr>
<th>&lt;P5&gt;</th>
<th>&lt;9-1 = 28970&gt;</th>
<th>&lt;9-2 = 28212&gt;</th>
<th>&lt;9-3 = 27290&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>02 9-1</td>
<td>04 0</td>
<td>04 0</td>
<td>04 0</td>
</tr>
<tr>
<td>00 A32</td>
<td>00 T0</td>
<td>00 Y3</td>
<td>00 Y3</td>
</tr>
<tr>
<td>00 Y1</td>
<td>02 9-2</td>
<td>00 Y2</td>
<td>00 Y2</td>
</tr>
<tr>
<td>00 A33</td>
<td>00 D12</td>
<td>00 Y4</td>
<td>00 Y4</td>
</tr>
<tr>
<td>00 Y2</td>
<td>02 9-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>00 A2</td>
<td>00 D13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>00 Y3</td>
<td>04 27858</td>
<td></td>
<td></td>
</tr>
<tr>
<td>00 A34</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00 Y4</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

END OF RUN
CHAPTER IV.

A CRITICAL LOOK AT ZBIE.

UHR’S PROGRAMS (1964) MAY WELL GIVE AN ILLUMINATING CONTRAST TO SOME OF THE MECHANISMS USED BY ZBIE. UHR SHOWS OUTPUTS FOR TWO PROGRAMS AND DESCRIBES A PLANNED THIRD PROGRAM. ALL HIS PROGRAMS LEARN TO TRANSLATE A NATURAL LANGUAGE, NL1, INTO ANOTHER NATURAL LANGUAGE, NL2.

THE FIRST PROGRAM WAS APPLIED TO TRANSLATE GERMAN INTO ENGLISH. WORDS ARE SEPARATED BY BLANKS. THE PROGRAM DOES NOT USE CONTEXT AND RUNS INTO DIFFICULTIES ON THAT ACCOUNT. FOR EXAMPLE, ‘DAS’ IN THE GERMAN ‘DAS IST’ OR ‘DAS GLAS’ BECOMES RESPECTIVELY ‘THIS’ OR ‘THE’. UHR’S FIRST PROGRAM CANNOT HANDLE SUCH A DIFFICULTY.

THE SECOND PROGRAM, WHICH WAS APPLIED TO TRANSLATE ENGLISH INTO FRENCH, IS MORE POWERFUL IN SEVERAL WAYS. WORDS ARE NOT SEPARATED, AND PART OF THE FUNCTION OF THE PROGRAM IS TO SEPARATE IMPORTANT ELEMENTS IN A CONTINUOUS FLOW. FOR INSTANCE, THE ‘S’ MARKING A PLURAL CAN BE SEPARATED. IT IS ALSO MORE DIFFICULT TO TRANSLATE A LANGUAGE WITH LITTLE INFLECTION, SUCH AS ENGLISH, INTO A MORE INFLECTED LANGUAGE, SUCH AS FRENCH, THAN TO TRANSLATE
AN INFLECTED LANGUAGE, GERMAN, INTO ENGLISH. IN THE LATTER CASE, AN INCREASED DICTIONARY OFTEN SUFFICES, WHILE IN THE FORMER CASE, CONTEXT IS ESSENTIAL.

TO OBTAIN CONTEXT, CLASSES ARE BUILT ON ELEMENTS OF NL2 CLASS CLUSTERS ARE USED TO RESOLVE AMBIGUITIES AND GENERALIZE PERMUTATIONS. IN THE GIVEN OUTPUT, THOUGH, THE PROGRAM DOES NOT SEEM TO LEARN THE GENERALIZED PERMUTATION: ADJECTIVE + NOUN + NOUN + ADJECTIVE.

THE THIRD (PROJECTED) PROGRAM IS A STRENGTHENING OF PROGRAM 2. IT CAN GENERATE CLASSES OF CLASSES, I.E. ENCOMPASS GROUPS OF WORDS.


LET US APPLY SOME OF THE ABOVE CRITERIA TO ZBIE. INPUTS TO ZBIE ARE WELL SEPARATED WORDS AND ZBIE'S STRING MANIPULATING CAPABILITIES ARE LIMITED TO CHECKING WHETHER THE FIRST FEW CHARACTERS OF TWO WORDS ARE IDENTICAL OR NOT. ZBIE TRIES TO TRANSLATE FROM A STRONGLY NON-INFLECTED LANGUAGE (FL) INTO ANOTHER LANGUAGE, WHICH WILL USUALLY BE INFLECTED. THE CLASSES
USED BY UHR'S PROGRAMS ARE BUILT ON NL2; THE CORRESPONDING SETS ARE BUILT ON FL, EQUIVALENT TO NL1, BY ZBIE. THE CLASSES OF CLASSES DESCRIBED BY UHR FOR THE PROJECTED PROGRAM MAY CORRESPOND TO THE RECURSIVE HIERARCHY (SUB)PATTERN - SET - SUBPATTERN. ZBIE MAKES FEW ASSUMPTIONS ON WORD ORDER, AND HENCE WAS ABLE TO LEARN THE SENTENCES EXHIBITED IN CHAPTER 3 WHEN THE RUSSIAN SENTENCES HAD ALL BEEN PREVIOUSLY INVERTED.

ALL THE PROGRAMS CONSIDERED BUILD STRUCTURES AT RUN-TIME AND USE THESE STRUCTURES TO LEARN FURTHER, WHICH MAY IMPLY BUILDING NEW STRUCTURES.

WE FEEL THAT USING A UNIFORM, STRUCTURED REPRESENTATION, SUCH AS FL, WHICH IS IN SOME SENSE 'UNDERSTOOD', GIVES ZBIE A GREAT ADVANTAGE OVER A FLAT STRING INPUT. IT DESTROYS THE SYMMETRY THAT UHR'S PROGRAMS POSSESS. IT MAY BE THAT STRUCTURES AS POWERFUL AS (OR EVEN MORE POWERFUL THAN) THOSE BUILT IN FL CAN BE CONSTRUCTED EVENTUALLY BY PROGRAMS STARTING FROM (UNSTRUCTURED) STRINGS.

WE HAVE ALREADY ENCOUNTERED VARIOUS DEFICIENCIES IN ZBIE SOME OF THESE COULD BE SOLVED BY MORE PROGRAMMING; OTHERS NECESSITATE A THOROUGH RETHINKING OF THE LEARNING TASK.

AMONG THE FIRST WE SHOULD MENTION:

- THE INITIALIZATION PHASE IS INFLEXIBLE. SENTENCES COULD BE STORED UNTIL TWO SENTENCES MEET THE CONDITIONS TO INITIALIZE THE
FIRST PATTERN.

- The translation rules should be more general. We gave in Chapter II, Part 8, two kinds of desirable translation rules which are presently not used by Zbie.

- The methods to obtain 'good guesses' should be improved. These methods could use, for instance, some of the semantics built into FL (see below).

- The string manipulating capabilities should be more versatile. In particular, prefixes and suffixes should be discovered, thereby allowing generalizations of translations. For example: the translation of an element FL1 in the context A1, P1 is obtained by adding some string in NL to the translation of the same element FL1 in some other context A1, P1.

- A pattern should have several translation rules to reflect different ways in which the same situation can be expressed in NL.

- As the number of (sub)patterns increases, heuristics (maybe of a semantic nature) must be used to speed up search during matching.

- Presently, the context of the (main) verb is the whole pattern, so that, as the same verb (in FL) changes (in NL), a new top pattern is created. In this way, the number of patterns increases much too rapidly.
Besides structure, we tried to build some semantics into FL. We can specify the referent of a pronoun. We can index, if necessary, several persons appearing in a situation to differentiate among them, just as they may be differentiated in a picture. Until now, though, no use has been made of the content of the situations.

A first step, which was nearly implemented, is to use findings such as: the subject pronouns (man) and (boy) have identical translations in Russian. We could use this finding to make 'good guesses' on how 'man' and 'boy' will behave in similar situations. We could also use the functional language description of: (speaking man number 21), (speaking (and man woman)), etc... to obtain the notion of (speaking plural).

A second step would be to process several situations which are dynamically related. Sequences such as: the book is on the table, I shall take the book in my hand, I am taking the book, the book is in my hand, it was on the table, etc... are used by I. A. Richards to teach tenses.

A simple scheme of set inclusion tests alone, as used by Zbie, is not powerful enough to learn, for example, the Russian reflexive possessive svoi. A more powerful scheme which looks at correlations among elements in the functional language is needed: in the FL sentence
(PUT (PERSON1) (IN HAT) IF (PERSON2)) DRAWER)

ARE PERSON1 AND PERSON2 THE SAME.

THE ABOVE DEFICIENCIES CAN BE OVERCOME BY MAKING BETTER USE OF FL. WE MUST ALSO QUESTION THE USE OF FL.

THE PURPOSE OF FL WAS DESCRIBED IN CHAPTER II. UNDOUBTEDLY, THE 'VISION OF THE WORLD' AS REPRESENTED BY FL IS VERY CLOSE TO AN INDO-EUROPEAN'S 'VISION OF THE WORLD'. IT MAY BE POSSIBLE THAT A SYSTEM COULD DISCARD THE FUNCTIONAL LANGUAGE AND BOOTSTRAP ITSELF TO LOOK AT THE WORLD IN THE LEARNED NATURAL LANGUAGE.

THE RUNNING SYSTEM.

PROGRAMMING ZBIE HAS BEEN A WORTHWHILE EXPERIENCE. THE PROGRAM GENERATED SOME SURPRISES. THE MECHANISMS FOR AVOIDING ERRORS WERE NOT VERY CLEAR AT THE START. THE ADVANTAGE IN TRANSLATING IN PARALLEL PATTERN LISTS OF A GIVEN DEPTH HAD NOT BEEN FORESEEN; IT WAS THOUGHT THAT A SIMPLE HISTORIC MONITOR: LAST PATTERN LIST OBTAINED, FIRST TRIED, WOULD BE ADEQUATE.

AT PRESENT, THE PATTERN BUILDING ROUTINES CAN HANDLE ONLY FL ELEMENTS WHICH HAVE SINGLE NL WORDS AS TRANSLATIONS. (THIS LIMITATION CAN BE OVERCOME BY ADDITIONAL PROGRAMMING.) CONSEQUENTLY, ZBIE WOULD BRANCH TO AN EMPTY EXIT WHEN TRYING TO PROCESS, IN FRENCH, PENCIL → LE CRAYON (TWO NL WORDS). DEALING WITH ARTICLES, MOSTLY UNNECESSARY BUT CUMBERSOME OBJECTS, POSES SOME PROBLEMS. IF PENCIL → LE CRAYON, AND IF WE ALSO KNOW THAT
BED ← LE LIT AND WALL ← LE MUR, WE COULD CONCLUDE THAT 'LE' IS AN UNNECESSARY OBJECT. THEN IF WE MEET PENCIL ↔ UN CRAYON (POSSIBLY), WE CAN ASSUME THAT 'UN' IS ALSO AN UNNECESSARY OBJECT. THE SCHEME COLLAPSES NEXT, THOUGH, WHEN WE MEET PENCIL ↔ SON CRAYON (HIS OR HER PENCIL). HERE A TEACHER MAY WELL BE NEEDED.

A 'GOOD' TEACHING SEQUENCE IS PROBABLY VERY IMPORTANT. EVEN WITH A GOOD SEQUENCE, THOUGH, A TIME WILL COME WHEN SOME BAD GENERALIZATIONS HAVE BEEN MADE, AND THESE MUST BE UNLEARNED. ZBIE IS CAREFUL ENOUGH TO AVOID MISTAKES SO THAT, AT THE ELEMENTARY LEVEL CONSIDERED, ERROR RECOVERY IS NOT NEEDED. WE HAVE NO EXCITING SCHEME TO PROPOSE. FOR ERROR RECOVERY, IT APPEARS THAT SUITABLE CHANGES IN SET MEMBERSHIPS MAY BE SUFFICIENT IN SOME CASES, BUT ADDITIONAL WORK AT A MORE ADVANCED LEVEL OF LANGUAGE PROFICIENCY IS IN ORDER.
ZBIE is a program of medium length (about 5000 lines of IPL-V code) which tries to solve some aspects of a task that may or may not be very difficult: natural language learning. ZBIE interprets the task as learning to express in a natural language situations as described in a uniform, structured functional language, which may be considered as giving some content to the situations, although actually very little of the power of the functional language is used.

To learn a language, ZBIE builds elementary structures at run-time: sets, patterns, simple translation rules, an in-context vocabulary. It is capable of somewhat improving its structure to the effect that some previously learned instance can be learned in a better way. It does not use statistical learning schemes. It has some error-avoiding capability, and has potential error-recovery possibilities which were actually not needed at the modest level considered here. As could be expected, ZBIE uses context both to learn and to avoid errors.

Given situations are parsed in the functional language, not in the natural language; the parsing gives a measure of how close
A new situation is to certain classes of previously encountered situations, Zbie tries to minimize its learning at each stage by trying to capitalize on the maximum amount of information available from previous situations, measured, quite simply, by the depth to which a new situation is parsed by a collection of patterns. In fact, Zbie abandons the learning task when faced with situations that it considers too difficult to handle at a given stage. Since it can diagnose the particular difficulties encountered, Zbie could ask appropriate questions in a time-sharing environment.

It would appear that the functional language should possess in its description all the semantic subtleties of the natural language learnt, an unappealing fact. A much more powerful system may be able to bootstrap itself and use the natural language it has started to learn as its main representation, with possible references from time to time to the functional language. Semantic subtleties could then be described in the natural language itself.

It appears to be fruitful to look at natural languages as ways to express situations with structure and content. Much valuable research can be done in areas touched upon by Zbie, among which the design of more powerful, semantically oriented, functional languages; the problem of error recovery, and the evolutionary reorganization and improvement of structures may well be the most interesting and challenging.
APPENDIX A.

CODE FOR CYRILLIC ALPHABET.

<table>
<thead>
<tr>
<th>A</th>
<th>A</th>
<th>P</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>Б</td>
<td>B</td>
<td>C</td>
<td>S</td>
</tr>
<tr>
<td>В</td>
<td>V</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>Г</td>
<td>G</td>
<td>У</td>
<td>U</td>
</tr>
<tr>
<td>Д</td>
<td>D</td>
<td>Φ</td>
<td>F</td>
</tr>
<tr>
<td>Е</td>
<td>E</td>
<td>Х</td>
<td>H</td>
</tr>
<tr>
<td>Е</td>
<td>ET</td>
<td>Ц</td>
<td>C</td>
</tr>
<tr>
<td>Ж</td>
<td>G</td>
<td>Ч</td>
<td>Y</td>
</tr>
<tr>
<td>З</td>
<td>Z</td>
<td>Ш</td>
<td>W</td>
</tr>
<tr>
<td>И</td>
<td>I</td>
<td>Щ</td>
<td>W2</td>
</tr>
<tr>
<td>Й</td>
<td>IT</td>
<td>Г</td>
<td>2</td>
</tr>
<tr>
<td>К</td>
<td>K</td>
<td>Ы</td>
<td>Ы</td>
</tr>
<tr>
<td>Л</td>
<td>L</td>
<td>Ы</td>
<td>Ы</td>
</tr>
<tr>
<td>М</td>
<td>M</td>
<td>Э</td>
<td>E2</td>
</tr>
<tr>
<td>Н</td>
<td>N</td>
<td>Ю</td>
<td>UT</td>
</tr>
<tr>
<td>О</td>
<td>O</td>
<td>Я</td>
<td>AT</td>
</tr>
<tr>
<td>П</td>
<td>P</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX B.

'EVOLUTIONARY LEARNING'.

WE ARE PRESENTING HERE A SHORT EXAMPLE, IN GERMAN, EXHIBITING AN ASPECT OF Z3IE'S CAPABILITIES. THE TRANSLATION RULE OF PATTERN P1 IS INCORRECT. AS MORE EXAMPLES ARE GIVEN, THOUGH, Z3IE BUILDS NEW PATTERNS P2 AND P3. THE SENTENCES WHICH WERE PREVIOUSLY TRANSLATED WHEN REACHING PATTERN P1 ARE NOW TRANSLATED WHEN REACHING PATTERNS P2 OR P3, SO THAT PATTERN P1 IS NOT REACHED ANY MORE. PATTERN P1 HAS EFFECTIVELY BEEN WASHED OUT.

THE OUTPUT IS REPRODUCED AS OBTAINED FROM THE PRINTER. INPUTS IN IPL-V FORM HAVE BEEN DELETED EXCEPT IN ONE ILLUSTRATIVE CASE.

THE IMPORTANT ATTRIBUTES AND VALUES OF A PATTERN PJ HAVE THE FOLLOWING MEANING:

TO(PJ) = TRANSLATION RULE (A LIST STRUCTURE).
PO(PJ) = A LIST OF PATTERNS WITH A P-LIST IDENTICAL TO PJ.
THE OTHER ATTRIBUTES ARE USED FOR BOOK-KEEPING AND MAY BE DISREGARDED.

COMMENTS ARE GIVEN AT THE RIGHT OF INPUTS, OR AT THE END OF
THE PROCESSING OF A SENTENCE.

LOOKING AT SENTENCE
(BE (WOMAN) HERE)
SIE IST HIER
(BE (WOMAN) THERE)
SIE IST DORT
PROCESS START
PUT INTO VOCABULARY
(BE (WOMAN))
(SIE IST)
P
PUT INTO VOCABULARY
HERE
HIER
A3
P
PUT INTO VOCABULARY
THERE
DORT
A3
P
NEW PATTERN

<9-1 = 25256>  <9-2 = 25482>  <9-3 = 25646>
02 9-1  04 0  04 0  04 0
00 A1  00 TO  02 9-3  00 Y1
Comment: the initialization phase is over. For the next sentence, we show the IPL-V input.

<table>
<thead>
<tr>
<th>IP Input Data</th>
<th>5</th>
<th>01</th>
</tr>
</thead>
<tbody>
<tr>
<td>( X_4 )</td>
<td>95</td>
<td>0</td>
</tr>
<tr>
<td>( X_3 )</td>
<td>95</td>
<td>0</td>
</tr>
<tr>
<td>( E_0 )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( E_2 )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( E_{10} )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( E_{13} )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( R_0 )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( R_4 )</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
LOOKING AT SENTENCE
(BE (MAN [NUMB 2]) THERE)
SIE SIND DORT
THE INPUTS ARE FIRST PRINTED
PROCESS START
TRY PATTERN
P
NOT MATCHED.
TRY LEARN FROM PATTERN
P
(Z DORT)
(BE (MAN [NUMB 2]))
(PATTERN P1 IS BEING BUILT
PUT INTO VOCABULARY
(SIE SIND)
PUT INTO VOCABULARY
(MAN [NUMB 2])
SIND
A2
P1
PUT INTO VOCABULARY
BE
SIE
A1
P1

NEW PATTERN

<table>
<thead>
<tr>
<th>&lt;P1&gt;</th>
<th>&lt;9-1 = 25256&gt;</th>
<th>&lt;9-2 = 25774&gt;</th>
<th>&lt;9-3 = 25420&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>02 9-1</td>
<td>04 0</td>
<td>04 0</td>
<td>04 0</td>
</tr>
<tr>
<td>00 A1</td>
<td>00 P0</td>
<td>00 P0</td>
<td>00 Y1</td>
</tr>
<tr>
<td>00 Y1</td>
<td>02 9-2</td>
<td></td>
<td>00 Y2</td>
</tr>
<tr>
<td>00 A2</td>
<td>00 Y0</td>
<td></td>
<td>00 Y3</td>
</tr>
<tr>
<td>00 Y2</td>
<td>00 J3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>00 A3</td>
<td>00 J4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>00 Y3</td>
<td>00 J3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>00 T0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>02 9-3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

COMMENT: PATTERN P1 HAS A P-LIST IDENTICAL TO THE P-LIST OF P0. HOWEVER, THE TRANSLATION RULES ARE DIFFERENT. ZBIE USED (BE(WOMAN)) → SIE IST, AND (BE(MAN[NUMB 2])) → SIE SIND, BOTH TRANSLATIONS IN THE CONTEXT OF THE PATTERN P0, FOR MATCH BACK THE COMMON PARTS IN FL AND NL, RESPECTIVELY 'BE' AND 'SIE', WERE PAIRED. THE COMPARISON OF THE NON-COMMON PARTS GAVE (MAN[NUMB 2]) → SIND (CONTEXT A2,P1). THE INFERENCE IS GRAMMATICALLY INCORRECT. NOTE THAT THE TRANSLATION RULE IS (Y1 Y2 Y3), IN THE SAME ORDER AS THE SETS. IT IS WORTH COMPARING THIS PATTERN P1 TO THE PATTERN P1 CREATED BY ZBIE WHEN LEARNING RUSSIAN (SEE CHAPTER
LOOKING AT SENTENCE
(BE (WOMAN) HERE)
SIE IST HIER
PROCESS START
TRY PATTERN
P1
PATTERN MATCHED. RESULT:
(SIE Z Z1)
TRY PATTERN
P
PATTERN MATCHED. RESULT:
(SIE IST HIER)
TRY LEARN MORE
TRY LEARN FROM PATTERN
P1
(SIE Z Z1)
GUESS
Z1
HIER
PUT INTO VOCABULARY
(WOMAN)
IST
A2
P1

THE FIRST SENTENCE AGAIN
(WOMAN) IS NOT KNOWN IN THE
CONTEXT (A2, P1), HENCE THE Z
ID. FOR HERE, (A3, P1) AND Z1

THE PATTERN P1 WAS THE FIRST TO
BE TOTALLY MATCHED, BUT WAS NOT
SUCCESSFULLY TRANSLATED
Z (I.E. (WOMAN)) IS NOT GUESSED
BUT Z1 IS, GIVING A CONSISTENT
TRANSLATION WITH THE INPUT
PUT INTO VOCABULARY
HERE NOW 'HERE' IS ALSO KNOWN
HIER IN THE CONTEXT (A3,P1)
A3
P1

COMMENT NOW (BE (WOMAN) HERE) WILL BE TRANSLATED BY PATTERN P1.

LOOKING AT SENTENCE
(BE (MAN) HERE)
ER IST HIER A NEW SENTENCE
PROCESS START
TRY PATTERN
P1
NOT MATCHED. (MAN) IS NOT IN A2
TRY PATTERN AS THE P-LISTS OF P0 AND P1 ARE
P IDENTICAL, IT IS NOT NECESSARY
NOT MATCHED. TO ACTUALLY MATCH P0
TRY LEARN FROM PATTERN
P1
(SIE Z HIER) NOT CONSISTENT WITH INPUT
TRY LEARN FROM PATTERN
P
(Z HIER) Z STANDS FOR THE UNKNOWN
PUT INTO VOCABULARY TRANSLATION OF (BE (MAN))
(BE (MAN))
ER IST
P
PUT INTO VOCABULARY CREATING PATTERN P2
(MAN )
er
A2
P2
PUT INTO VOCABULARY
BE
IST
A1
P2
NEW PATTERN

<table>
<thead>
<tr>
<th>&lt;P2&gt;</th>
<th>&lt;9-1 = 25814&gt;</th>
<th>&lt;9-2 = 26054&gt;</th>
<th>&lt;9-3 = 25854&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>02 9-1</td>
<td>04 0</td>
<td>04 0</td>
<td>04 0</td>
</tr>
<tr>
<td>00 A1</td>
<td>00 P0</td>
<td>00 P0</td>
<td>00 Y2</td>
</tr>
<tr>
<td>00 Y1</td>
<td>02 9-2</td>
<td>00 P1</td>
<td>00 Y1</td>
</tr>
<tr>
<td>00 A2</td>
<td>00 Y0</td>
<td>00 Y3</td>
<td></td>
</tr>
<tr>
<td>00 Y2</td>
<td>00 J3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>00 A3</td>
<td>00 J4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>00 Y3</td>
<td>00 J3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>00 T0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>02 9-3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

COMMENT: WE MATCHED BACK (BE(MAN)) → ER IST AND (BE(WOMAN)) → SIE IST, TO OBTAIN A NEW PATTERN P2. THE TRANSLATION RULE OF
P2, (Y2 Y1 Y3), is different from the translation rule of P1 and is grammatically correct. Note that the translation rule contains a transformation.

Looking at sentence
(BE (WOMAN) HERE)
SIE IST HIER

Again the first sentence
Process start
Try pattern
P2

Pattern matched. Result:
(Z IST Z1)

Try pattern
P1

Pattern matched. Result:
(SIE IST HIER)

Now translated by P1

Try learn more
Try learn from pattern
P2
(Z IST Z1)

Guess
Z1

HIER

Put into vocabulary
HERE

HIER
A3
P2
PUT INTO VOCABULARY
(WOMAN )
SIE
A2
NOW THE SENTENCE WILL BE
TRANSLATED BY P2
P2

LOOKING AT SENTENCE
(BE (SPEAKING MAN (NUMB 2 ))HERE )
WIR SIND HIER
A NEW SENTENCE
PROCESS START
TRY PATTERN
P3
(SPEAKING MAN (NUMB 2 )) IS
NOT MATCHED.
TRY PATTERN
P1
NOT MATCHED.
TRY PATTERN
P
NOT MATCHED.
TRY LEARN FROM PATTERN
P2
(Z IST HIER )
NOT CONSISTENT WITH INPUT
TRY LEARN FROM PATTERN
P1

TRY LEARN FROM PATTERN
P
\(Z\) HIER
PUT INTO VOCABULARY
\((BE \ (SPEAKING \ MAN \ (NUM\ 2 \ 1))\)
\(WIR\ SIND\)
P
PUT INTO VOCABULARY
CREATING PATTERN P3
\((SPEAKING \ MAN \ (NUM\ 2 \ 1))\)
\(WIR\)
A2
P3
PUT INTO VOCABULARY
BE
SIND
A1
P3
NEW PATTERN

\[
\begin{array}{ccccc}
\langle P3 \rangle & \langle 9-1 = 25584 \rangle & \langle 9-2 = 26328 \rangle & \langle 9-3 = 2413 \rangle \\
02 & 9-1 & 04 & 0 & 04 & 0 & 04 & 0 \\
00 & A1 & 00 & P0 & 00 & P0 & 00 & Y2 \\
00 & Y1 & 02 & 9-2 & 00 & P2 & 00 & Y1 \\
00 & A2 & 00 & Y0 & 00 & P1 & 00 & Y3 \\
00 & Y2 & 00 & J3 & & & & \\
00 & A3 & 00 & J4 & & & & \\
\end{array}
\]
COMMENT: THE TRANSLATION RULES OF P3 AND P2 ARE IDENTICAL, BUT 'BE' IS TRANSLATED DIFFERENTLY IN THE CONTEXTS (A1,P2) AND (A1,P3).

LOOKING AT SENTENCE
(BE (MAN (NUMB 2)) THERE)

SIE SIND DORT
THE THIRD SENTENCE AGAIN

PROCESS START
TRY PATTERN
P3
PATTERN MATCHED. RESULT:
(Z SIND Z1)
TRY PATTERN
P2
PATTERN MATCHED. RESULT:
(Z IST Z1)
TRY PATTERN
P1
PATTERN MATCHED. RESULT:
(SIE SIND Z)
TRY PATTERN
P
PATTERN MATCHED. RESULT:
Try learn more
Try learn from pattern
P3
(Z sind Z1)

Guess
Z1
Dort
Put into vocabulary
There
Dort
A3
P3
Put into vocabulary
(MAN (NUMB 2))
SIE
A2
Now the sentence will be
P3
Translated by P3

Comment: We shall now give sentences one and three again, and show how they are translated by P2 and P3 respectively. The pattern P1 is not reached any more.

Looking at sentence
(BE (WOMAN)HERE)
SIE IST HIER
Process start
TRY PATTERN
P3
PATTERN MATCHED. RESULT:
(Z SIND Z1)
TRY PATTERN
P2
PATTERN MATCHED. RESULT:
(SIE IST HIER)
TRY LEARN MORE
TRY LEARN FROM PATTERN
P3
(Z SIND Z1) NOT CONSISTENT

LOOKING AT SENTENCE
(BE (MAN (NUMB 2)) THERE)
SIE SIND DORT
PROCESS START
TRY PATTERN
P3
PATTERN MATCHED. RESULT:
(SIE SIND DORT)
END OF RUN
BIBLIOGRAPHY.


REICHENBACH, H., ELEMENTS OF SYMBOLIC LOGIC, MACMILLAN, NEW YORK, 1947.

RICHARDS, I. A., ET AL., RUSSIAN THROUGH PICTURES, BOOK I, WASHINGTON SQUARE PRESS, NEW YORK, 1961. THERE ARE SIMILAR BOOKS FOR: ENGLISH, FRENCH, HEBREW, ITALIAN, SPANISH AND GERMAN.

SIMON, H. A. AND KOTOVSKY, K., 'HUMAN ACQUISITION OF CONCEPTS FOR


SOLOMONOFF, R. J. 'A FORMAL THEORY OF INDUCTIVE INFFERENCE,' INFORMATION AND CONTROL, 1964, PP. 1-72, 224-254.


UHR, L. 'PATTERN STRING LEARNING PROGRAMS,' BEHAVIORAL SCIENCE, Vol. 9, JULY 1966, PP. 258-270.
Learning a natural language is taken as an improvement in a system’s ability to express situations in a natural language.

This dissertation describes a computer program, called Zbie, written in IPL-V, which accepts the description of situations in a uniform, structured functional language and tries to express these situations in a natural language. Examples are given for German and, mostly, Russian.

At run-time, Zbie builds simple memory structures. Patterns and sets are built on the functional language. The translation rules of the patterns and an in-context vocabulary provide the transition to the natural language. Zbie is a cautious learner, and avoids errors by several mechanisms. Zbie is capable of some evolutionary learning.