A Software System for Describing a Grammar
of Machine Translation: GRADE

Jun-ichi Nakamura and Makoto Nagao

Kyoto University
Japan

ABSTRACT

A new software system for describing a grammar of a machine
translation system has been developed. This software system
is called GRADE (GRAmmar DEscriber). GRADE has the follow-
ing features:

1.   GRADE allows a grammar writer to divide a whole grammar
     into several parts. Each part of the grammar is called
     a subgrammar. A subgrammar describes a step of the
     translation process. A whole grammar is then described
     by a network of sub-grammars. This network is called a
     subgrammar network. A subgrammar network allows a gram-
     mar writer to control the process of the translation
     precisely. When a subgrammar network in the analysis
     phase consists of a subgrammar for a noun-phrase (SG1)
     and a subgrammar for a verb-phase (SG2) in this
     sequence, the subgrammar network first applies SG1 to an
     input sentence, then applies SG2 to the result of an
     application of SG1, thus getting a syntactic structure
     for the input sentence.

2.   A subgrammar consists of a set of rewriting rules.
     Rewriting rules in a subgrammar are applied for an input
     sentence in an appropriate order, which is specified in
     the description of the subgrammar. A rewriting rule
     transforms a tree structure into another tree structure.
     Rewriting rules use a powerful pattern matching algo-
     rithm to test their applicability to a tree structure.
     For example, a grammar writer can write a pattern that
     recognizes and parses an arbitrary numbers of sub-trees.
     Each node of a tree-structure has a list of pairs of a
     property name and a property value. A node can express
     a category name, a semantic marker, flags to control the
     translation process, and various other information.
     This tree-to-tree transformation operation by GRADE
     allows a grammar writer to describe all the processes of
     analysis, transfer and generation of a machine trans-
     lation system with this uniform description capability
     of GRADE.

3. A subgrammar network or a subgrammar can be written in
    an entry of the dictionaries for a machine translation
    system. A subgrammar network or a subgrammar written in
    a dictionary entry is called a dictionary rule, which is
    specific for a word. When an input sentence contains a
    word which has a dictionary rule, it is applied to an
    input sentence at an appropriate point of a translation
    process. It can express more precise processing appro-
    priate for that specific word that a general Subgrammar
    Network or Subgrammar. it also allows grammar writers
    to adjust a machine translation system to a specific
    domain easily.

4. GRADE is written in LISP. GRADE is implemented on FACOM
    M-382 and Symbolics 3600.

    GRADE is used in the machine translation system between
Japanese and English. The project was started by the
Japanese government in 1982. The effectiveness of GRADE has
been demonstrated in the project.