APPLICATION OF A RULE-BASED KNOWLEDGE SYSTEM USING CLIPS FOR THE TAXONOMY OF SELECTED OPUNTIA SPECIES

Bart C. Heymans, Joel P. Onema, Joseph O. Kuti

Horticulture Research Laboratory, College of Agriculture
Texas A&I University, Kingsville, Texas 78363

Abstract. A rule-based knowledge system was developed in CLIPS (C-Language Integrated Production System) for identifying Opuntia species in the family Cactaceae, which contains approximately 1,500 different species. This botanist expert tool system is capable of identifying selected Opuntia plants from the family level down to the species level when given some basic characteristics of the plants. Many Opuntia species are cultivated as ornamental plants and some are significant as food crops. Opuntia plants are becoming of increasing importance because of their nutrition and human health potential especially in the treatment of diabetes mellitus. The expert tool system described in this paper can be extremely useful in an unequivocal identification of many useful Opuntia species.

INTRODUCTION

Rule-based expert tool systems have been implemented in a variety of applications (NASA 1989). One of the better suited applications of the tool systems is in botany or plant sciences, where they can be used to effect plant identification and taxonomic search. Plant taxonomy is particularly suited for the tool systems because it usually starts out from empirical characteristics, which are sometimes confusing and/or fuzzy in nature, until it narrows down the characteristics to a few but concrete ones for the eventual identification of particular plant species and their botanical synonyms, which are common to all scientific nomenclatures.

This particular tool systems may start out from plant kingdom and end in the eventual identification of the plant species. The overall taxonomic structure obtained constitutes a tree-like hierarchy [fig. 1] in the leaves represents individual species. This hierarchy will be covered in some details later. The CLIPS language, developed by NASA has been chosen for this system because of its portability, flexibility and the capability of its integration with other languages e.g. the C language (NASA 1990).

SYSTEM DESCRIPTION

A general overview of the system is shown in a directed graph [fig. 2]. The system is capable of identifying selected Opuntia and Nopalea species (Britton and Rose 1963) when adequate information are given on the basic characteristics that distinguish these species from each other (Buxbaum 1950). The system also enables the user to start species identification from any taxonomic level. For example, if the plant family is known, identification may start at this level, or if the plant series and forms are known, these will
become the starting points of the taxonomic search. If on the other hand, no plant speci-
can be identified, the user has the option to continue the search, to quit or to start all over
again.

To build the tool system we opted for a modular approach. [fig. 4] The system
consists of different knowledge segments modularly linked together as one logical know-
ledge base. Each knowledge segment could be considered as having a salience associated with
so that it can only be accessed in a predetermined order. As identification proceeds, an
item (e.g. in the plant Order, Class, Series, etc...) must be identified and loaded in a pare-
segment or a subsidiary segment provided the item in question does not constitute the em-
end of the search. Smaller subunits can be loaded from the subsidiary segment when they have
been duly selected. Thus, by means of this modular approach, a cascading effect will occur
and plant characteristics selected by the user may fire the necessary rules in the starting
segment. These rules will then propagate through intermediary segments and will eventual-
terminate in the identification of the plant species in one of the leaf segments.
Currently, the functioning modules in the system include a root or driver segment that
contains the rules for the plant taxa, several intermediary modules [fig. 5] and the leaf
segments which form core of the program.

The Opuntia and Nopalea species in the Cactaceae family [fig. 1], as described by
Britton and Rose (1963) and Pizzetti (1985), were chosen as model for the expert to-
system. Selected plants belong to the tribe Opuntiae, the the family Cactaceae and the order
Cactales [fig. 1]. In the tribe we identify plants in the genera Nopalea and Opuntia. Theh
include species in the series Dillenianae, Streptacantheae, Polyanthaeae, Ficus-indicae, and
Robustae, in the subgenus Platypuntia and in the genus Opuntia. The system has addition-
capabilities which include the possibility of keeping track of multitude and often confusing
botanical synonyms that are common with plant taxonomy. Our system is adaptable to
regular updating of the knowledge segment so that the system can display all the botanici-
synonyms, if so desired. Our system can also give the common name in any or all of five
selected foreign languages (if a common name is available in that language).

The system is menu driven in which the user is presented interactively with a choic-
between predefined plant characteristics, hierarchically structured [fig. 3] and demonstra-
tion option. Since we opted for a modular approach the system agenda is kept to the minimu-
so that fewer rules can be fired at a time. In consequence, the system closely mimics the
taxonomic categories or taxa used in plant and animal systematics. This allows for quick
implementations of new modules from the existing (i.e. not yet implemented) or undeфин-
(i.e. previously not defined in the taxonomy) items, easier merger of the two modules if the
two items can be unified taxonomically and deletion of the obsolete segments especially
when an item loses its independent status. When a new botanical name is used to identify
a plant, the name must only be changed inside its particular segment since its scope
limited to that segment. For example when the name Angiospermopsida is substituted f<
Angiospermae the name has to be changed only inside its segment (i.e. the regnum
segment).

The addition of new knowledge segments can potentially stretch the available
memory to its limits, since at a run time only the selected module (i.e. out of a potential
limitless number) will be active thus requiring only very limited run time capabilities. Our
modular approach addresses several potential problems one may encounter in plant
taxonomy e.g. constant changes in plant classification at species level and moving plant

506
genera from one family to the other and the use of different taxonomic methods (Cronsquist 1968). The rule based system described in this paper is capable of dealing with all these problems. When there is a change in taxonomic criteria agreed upon in the scientific community, such as in case of the Opuntia species, new rules can be added to reflect these changes. Old taxonomic criteria can be moved to a separate knowledge segment where other botanical synonyms or plant names are stored. When one is identifying a plant species, one can obtain the botanical synonyms or plant names that are currently set as the standard. This will ensure that people working with an older or different taxonomic criteria are still able to use the tool system and have opportunity to learn the current taxonomic criteria and terminology that are internationally acceptable for identifying a particular plant species.

CONCLUSION AND FUTURE PROSPECTS

This project has demonstrated the feasibility of using CLIPS to build an Opuntia expert tool system. Although, the knowledge base was implemented for a selected number of Opuntia species, their available common names and synonyms. The system can be expanded, at any time, to include more Opuntia species in the family Cactaceae in different foreign languages by adding or expanding the parent or subsidiary segments.

The nature of taxonomic definitions and search, in regards to modelling the real world definitions, contains inherently fuzzy concepts and definitions. Logically, fuzzy qualifiers are most appropriate to represent the taxonomic descriptions. In our case we only concern ourselves marginally with this problem while giving the user more options. In an improved version however, we will implement a different mechanism for dealing with fuzzy definitions. In this perspective, a probabilistic value will be attached to the different manifestations of a characteristic, therefore allowing the system to choose on the basis of the value of the characteristic adopted.

REFERENCES


System Overview

Figure 1.

Basic Characteristics

Species Identification

Knowledge-Base

Inference Engine

Expert System

Figure 2.
User \rightarrow Menu \rightarrow options

Hierarchical Structure

Characteristics

Demonstration

Decision

Facts

Knowledge Base

Figure 3.

LOGICAL KNOWLEDGE BASE

1 KNOWLEDGE SEGMENT

2 KNOWLEDGE SEGMENT

KNOWLEDGE SEGMENT

SUBSIDIARY SEGMENTS

SUBSIDIARY SEGMENTS

SUBSIDIARY SEGMENTS

SUBSIDIARY SEGMENTS

Figure 4. MODULAR CONFIGURATION
Figure 5. Functional Modules