N92-16614

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 ov again.

To build the tool system we opted for a modular approach. [fig. 4] The syste consists of different knowledge segments modularly linked together as one logical knowled 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, 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 e of the search. Smaller subunits can be loaded from the subsidiary segment when they ha been duly selected. Thus, by means of this modular approach, a cascading effect will occ and plant characteristics selected by the user may fire the necessary rules in the starti 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 th contains the rules for the plant taxa, several intermediary modules [fig. 5] and the le segments which form core of the program.

The *Opuntia* and *Nopalea* species in the Cactaceae family [fig. 1], as described Britton and Rose (1963) and Pizzeti (1985), were chosen as model for the expert to system. Selected plants belong to the tribe Opuntiae, the the family Cactaceae and the ord Cactales [fig. 1]. In the tribe we identify plants in the genera *Nopalea* and *Opuntia*. Th include species in the series *Dillenianae*, *Streptacanthae*, *Polyacanthae*, *Ficus-indicae*, ar *Robustae*, in the subgenus Platyopuntia and in the genus *Opuntia*. The system has addition capabilities which include the possibility of keeping track of multitude and often confusi 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 botanic synonyms, if so desired. Our system can also give the common name in any or all of fi 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 choi between predefined plant characteristics, hierarchically structured [fig. 3] and demonstratio 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 t 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 undefine (i.e. previously not defined in the taxonomy) items, easier merger of the two modules if t two items can be unified taxonomically and deletion of the obsolete segments especial when an item loses its independent status. When a new botanical name is used to identi 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 segement (i.e. the regnu segment).

The addition of new knowledge segments can potentially stretch the availab 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. O modular approach addresses several potential problems one may encounter in pla taxonomy e.g. constant changes in plant classification at species level and moving pla 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 feasability 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 symnonyms. The system can be expanded, at any time, to include more *Opuntia* species in the family Cactacaea 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 defintions, 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

Britton, N.L. and Rose, J.N. (1963) The Cactaceae. Vol. 1 & 2. Dover Press. New York.

Buxbaum, F. (1950). Morphology of cacti. Abbey Garden Press, Pasadena, California.

Cronquist, A. (1968). The evolution and classification of flowering plants. Houghton Mifflin Company, Boston.

Pizzeti, M. (1985). Guide to Cacti and Succulents Simon and Schuster. New York.

National Aeronautics and Space Administration. (1989). Clips Users Guide Version 4.3. Artificial Intelligence Section. Lyndon B. Johnson Space Center.

National Aeronautics and Space Administration. (1990). In CLIPS, first Clips Conference Proceedings, Houston.

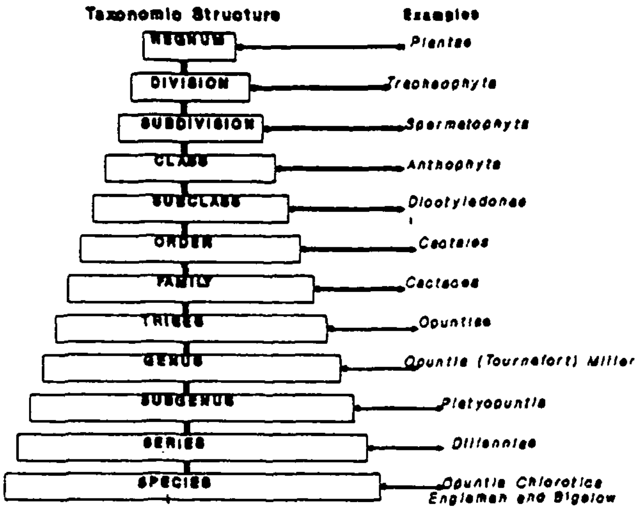

Figure 1.

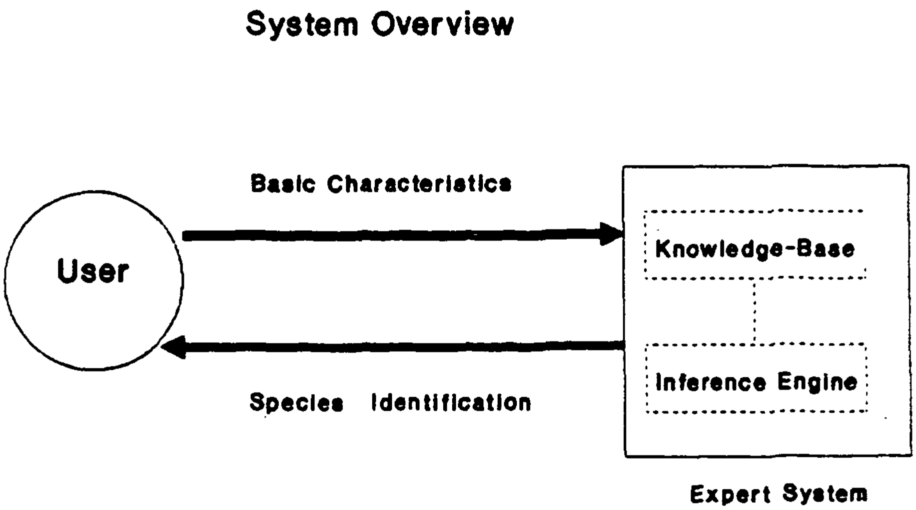

Figure 2.

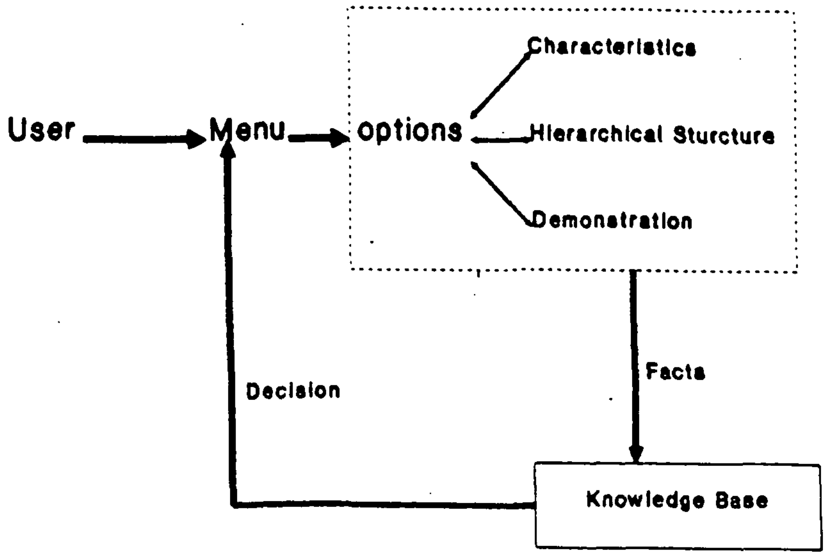

Figure 3.

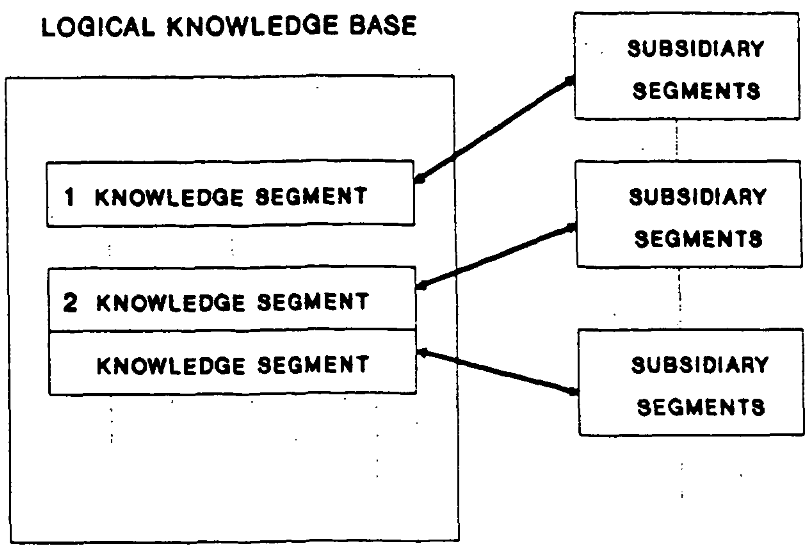

Figure 4. MODULAR CONFIGURATION

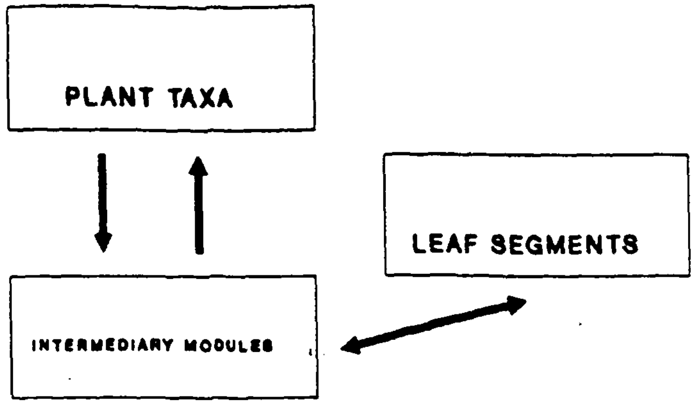

Figure 5. Functional Modules