IKBS : an iterative knowledge base system for managing knowledge in remote systematics. Application to corals of the Mascarene Archipelago

Conf.

Auteurs

Conruyt N., Grosser D., Faure G., Pichon M., Geynet Y., Guillaume M., Bouchon C.

Abstract

IKBS is a generic software workshop for building knowledge bases in systematics on the internet. A knowledge base is an application which contains the know-how of experts for describing, classifying and identifying organisms in a given domain. IKBS manages complex knowledge (structured, variable, imprecise, noisy) with an object oriented representation formalism (written in java). It uses a knowledge management cycle to store the expertise in the knowledge base. The cycle is divided into three phases: acquisition, processing and refinement.

For the knowledge acquisition part, the expert defines his domain at a taxonomic level with observable characters (objects, attributes, values). They are structured in a description tree and the result is called a descriptive model. It is designed with the help of natural descriptive logics (decomposition, specialisation, contextual conditions, iteration, etc.). Then, an hypermedia questionnaire which matches the descriptive model is automatically built. It is used by different users as an observation guide for collecting descriptions of observed specimens in a case base, and the expert references each description by its species' name.

In the knowledge processing phase, a decision tree is automatically learnt by the system in discriminating from the pre-classified examples. Classification rules (diagnoses) can also be generated. Any user identify new observations by following tree nodes from the root to leaves and answering questions. IKBS uses a case-based reasoning approach to switch dynamically to alternative questions when an unknown answer is given: it then follows another decision sub-tree.

The knowledge refinement part uses an iterative process to evaluate the robustness of the previous descriptive model and descriptions. First, an experimentation is led with a part of the case base. Results of classification and identification as well as the use of the questionnaire with the test set give the expert the possibility to detect inconsistencies in the case base (errors in descriptions) and descriptive model (misunderstood characters, bad illustrations).

As an application, a knowledge base for the family *Pocilloporidæ* of corals of the Mascarene Archipelago is under construction. A first validation was made with 9 taxa (species and ecomorphs) of the genus *Pocillopora*. It gave on average 2/3 of good identifications with a sample of different non specialist users. The validation put stress on the difficulty for an expert to choose characters easy to observe and interpret. The

meticulous choice of vocabulary, drawings and images is then decisive for the robustness of the knowledge base.

This is why IKBS is designed to build knowledge bases cooperatively between experts. The aim is to draw up a thesaurus of terms and illustrations (i.e. the questionnaire) in order to reduce differences of interpretations that are the primary causes of misidentifications. As collections of specimens and experts are distributed around the world, we will bring them means to do remote systematics with satellite high speed broadband networks. Each expert is responsible for a family (*Pocilloporidæ, Fungiidæ, Poritidæ, Agariciidæ*). Using video-conference, they share their interpretations of observations synchronously with specimens under the microscope and a camera connected to it.

Key words

knowledge management, iterative approach, remote systematics, corals, Mascarene archipelago, descriptive model