

Web-Based Virtual Lab for Taxonomic Description

Alessandra Gomes, André Santanchè, and Fabiani de Souza
 Institute of Computing, University of Campinas
 Campinas, SP, Brazil
 ra075936@students.ic.unicamp.br, santanche@ic.unicamp.br,
 ra108171@students.ic.unicamp.br

ABSTRACT

The taxonomic description of a specimen is an essential task carried out by biologists aimed to identify and study living beings. The usual approach involves analysing and describing a given specimen in a physical laboratory. Nevertheless, several tasks are being virtualized. Images, sounds, and videos of living beings are being digitalized; records are stored in spreadsheets and databases and the description task itself are being supported by specialized software. This work investigates a step beyond, where the laboratory itself becomes virtual.

Categories and Subject Descriptors

D [Software]: [Miscellaneous]; J.3 [Life and Medical Sciences]: [Biology and genetics]

General Terms

Experimentation

Keywords

virtual lab, software component, taxonomic description, web

1. INTRODUCTION

Virtual laboratories simulate physical equipments and the infrastructure of a physical laboratory by using computational techniques. They can represent experiments by graphical interfaces and offer interactive simulations.

Virtual labs can be tooled to afford learning experiences comprising exercises, theoretical explanations, and interactive assistants that explain experiments step by step. In many cases, they can be used any time and from anywhere. This kind of laboratories are also known as simulated laboratories or e-laboratories [4].

By handling and combining visual software components, users can describe specimens in a virtual laboratory. This paper presents our work of such a tool involving the description of

living beings. In this work we investigate a specific kind of Biology virtual laboratory to support taxonomic description of specimens, in which the basic lab elements are virtualized as visual software components.

This is an ongoing work and, in order to validate our proposal, we have implemented a preliminary prototype with components to describe monitor lizards of the genus *Varanus*. The lab is based on a system called Varan-ID.

Varan-ID is an online determination system for monitor lizards. It is based on a morphological knowledge base of a group on carnivorous lizards, the genus *Varanus*. The system is based on the idea that not only experts are involved with monitor lizards. Students, curious, breeders, keepers are either interested in this subject, but may not have the necessary knowledge to work with the specimen. The process of identification in the Varan-ID system is based on descriptors.

In this paper we present a prototype of our virtual web laboratory to describe and to identify living beings. It is based on visual components handled by direct manipulation, which play roles of building blocks for descriptions and lab tools. Therefore, when a user inserts a component that represents a tail in the composition, he/she will add a related tail descriptor in the lizard description. The entire lab runs over the web on top of the *Componere* authoring environment[8], which explores the Rich Internet Application (RIA) approach to provide an interactive interface.

The remaining of the paper is organized as follows: Section 2 presents a taxonomic description model of the context of the developed tool. Section 3 presents implementation details of the tool. Section 4 presents future works and conclusions.

2. TAXONOMIC DESCRIPTION MODEL

The starting point for designing our lab was the software Xper2 (<http://lis-upmc.snv.jussieu.fr/lis/?q=en/resources/software/xper2/>). This tool supports the identification and description of specimens. It follows the character/character state (C,CS) [5] approach, organized in three phases:

- (i) to define descriptors and possible states;
- (ii) to relate descriptors/states to species;
- (iii) to identify a given specimen by recognizing values for each descriptor.

The Varan-ID system was developed over the Xper2. Its descriptors are organized in two distinct groups, easy-to-see descriptors and expert descriptors. Table 1 shows an example of some easy-to-see descriptors and their respective states.

Table 1: Easy-to-see Descriptors

Lizard Part	Descriptor	States
Tail	transversal section of the tail	roundish or laterally compressed
Head	position of nostrils between eyes and tip of snout	same distance from eyes than from tip of snout or nearer the tip of snout than the eyes or nearer the eyes than the tip of snout
Tongue	tongue coloration	red, light pink or whitish or blue, purple or black

Our tool is able to interact with Xper2 by accessing SDD files it can export. The Structure Descriptive Data format (SDD) is an open standard endorsed by the TDWG (Taxonomic Database Working Group) and DELTA (Descriptive Language for Taxonomy) for representing taxonomic descriptions in a XML format (<http://wiki.tdwg.org/SDD/>).

Figure 1 shows a diagram representing a fragment of a SDD file containing data to describe Varanus lizards. The hexagons represent elements, the rectangles represent texts and the ovals represent attributes. The CategoricalCharacter element defines a descriptor and possible states. In this example it defines the tongue coloration element and three possible states: red, light pink or whitish and blue, purple or black.

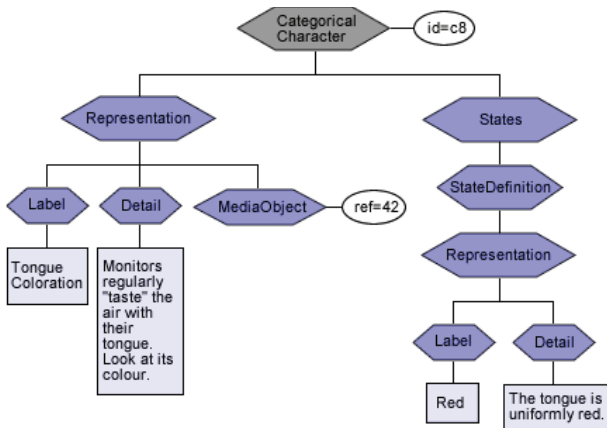


Figure 1: A Fragment of a SDD file of Varanus Lizards

The States element aggregates all possible states: StateDefinition elements. As can be seen in the diagram, the Representation element can be applied in many levels of the schema, containing textual and multimedia descriptions. This element is formed by a label, a detailed description (Detail) and references to multimedia resources (MediaObject)

In our tool we map these SDD description blocks in the following way:

- (i) each CategoricalCharacter becomes a description component;
- (ii) the set of states that the CategoricalCharacter can assume is transformed in a set of possible states that the component can assume;
- (iii) every time the description component assumes a state it provides a visual feedback.

These description components transform the task of describing specimens in selecting and customizing components, which are combined in compositions.

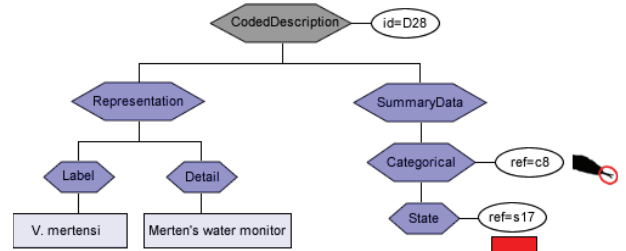


Figure 2: The Structure of Lizard SDD Base

Our resulting composition reflects another part of the SDD representation, illustrated in Figure 2. Besides the Representation element, described before, the diagram shows the summary data of the lizard. This element contains the element Categorical that represents the description of a given specimen (a lizard in the example), specifying Categorical Characters – referenced by the Categorical element – and setting values to it, through the State element. In this example, the categorical character "tongue coloration" assumes the state "red". Since each composition specifies the description of a given specimen, the resulting composition can be mapped to a SDD structure presented in Figure 2 and vice-versa.

In order to automatically derive SDD Categorical Characters to description components, a process was created to get all information on the SDD file and use it to fetch the respective description components, customizing them with the respective values. Figure 3 illustrates the creation process of lizard components and their use during a composition. A proxy reader component accesses a SDD file (step 1). This information is delivered to a component generator that creates all the description components (step 2 and 3) – lizard description components in the example. Therefore, the categorical element will be used to identify a specimen characterization – a lizard in the example – to generate its respective composition. The State element will generate the value that the parameter assume.

The author customizes and combines components, building a composition, whose parameters are recorded by the Settings component (step 4). When the author clicks in the Candidate button it triggers the Result component (step 5), which in turn retrieves the parameters from the Setting component (step 6), and uses the proxy component to fetch

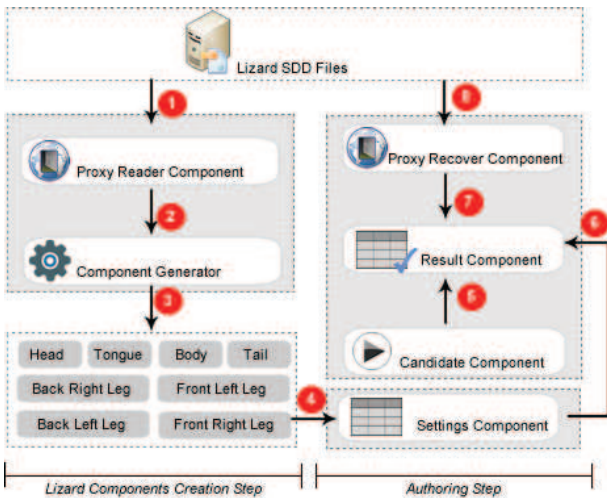


Figure 3: The Creation of Lizard Components and Composition Process

data from XML files in the base (step 7); which uses the HTTPXMLRequest API (step 8).

In our proposal, a description component is one type of the available components. Another type comprises tool components, a group that will be responsible to support the authoring process. The classification based on this two types of components is illustrated in Figure 4.

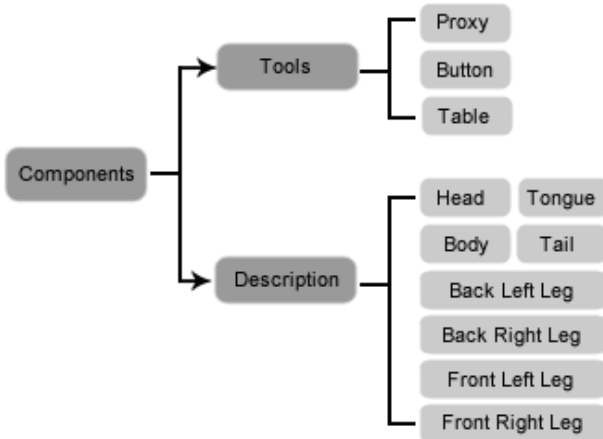


Figure 4: The Component Classification for the Lizard Lab

Our lab is designed to afford any kind of description component for living beings. However, in our prototype we have produced only description components representing each part of the Varanus lizard. They are visual components derived from Varan-ID easy-to-see descriptors. The tool components can be visual or not. For example, the button is a visual component that starts the execution process. The table is a visual component that contains data organized as rows and cols. The proxy is a non visual component that brings data stored in a database to the composition.

As mentioned before, our laboratory is built over the *Com-*

ponere environment. In the original *Componere* authoring environment all components play the role of building blocks. Our lab, on the other hand, introduces this new kind of component - the tool component - to assist the authoring task itself.

3. IMPLEMENTATION

The implementation of the proposed laboratory involved two steps: the construction of description components and the construction of the laboratory on *Componere*. As mentioned before, this first prototype is focused on a specific practical scenario involving the identification of Varanus lizards, based on the Varan-ID database.

3.1 Construction of Lizard Components

The Varan-ID base is composed by 7 knowledge bases: the Main base, *V. indicus*-group base, *V. prasinus*-group base, *V. timorensis*-group base, *V. gouldii*-group base, *V. salvator*-group base, and the Australian spiny-tailed base. Each base can be exported as an SDD file. To build the components, we analysed the available descriptors in these bases.

A component builder engine was developed to extract information from the exported database and to use them to build each one of the lizard components.

There are two ways to build *Componere* compositions. The first is by using a javascript code to instantiate and to connect components. The second is by embedding compositions in HTML pages through microformats based specifications [8]. During the authoring process the laboratory uses the first dynamic approach. Resulting compositions can be further materialized as HTML embedded compositions.

3.2 The Laboratory on Componere

As mentioned before, *Componere* is a framework based on javascript components that works over web browsers.

Thus, the Lizard Lab is an environment totally based on javascript, mainly directed to beginners in the monitor lizard identification process. An overview of the system is illustrated in Figure 5.

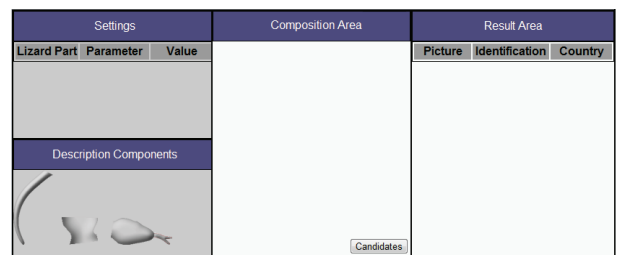


Figure 5: The Model of the Lizard Lab

The environment is organized in four areas: description components, settings, composition and result area. The description components area is where the components representing parts of the lizard are placed. Each one is independent and has its own set of parameters to be configured.

In order to produce a description the author drags description components to the Composition Area where they are

customized and connected. Each description component has two basic main actions: close and configure. The first will remove the component of the composition. The second will open a dialog box with the parameters to be configured. Whenever a component is changed, the Settings area (displayed in Figure 5) stores and shows a log of the values assigned to descriptors. Figure 3 illustrates this relation in the process.

During the description/identification process the author can access the base containing available descriptions of existing species – lizard in this case – whose descriptor/states match with those already assigned in the lab. For example, if the author assigns a specific tongue color and tail shape, the system will record these settings in the Settings area; whenever the author clicks in the “Candidates” button, available in the environment (see Figure 5), the system fetches and presents all lizards in the base which have the informed tongue color and tail shape. The steps to execute this process is illustrated in Figure 3. This technique to present progressive candidates is based in Xper2 approach to describe specimens.

The prototype is available at <http://fluidweb.sourceforge.net>. The page has many experiments using different kind of software components. To access the work proposed in this paper, click on the link “Lizard Prototype”.

4. RELATED WORK

Our work combines two approaches: virtual laboratories and tools to describe and identify specimens.

According to [4, 6] the laboratories can be classified in three categories: real, remote and virtual. Real labs are physical rooms, having concrete equipments and infrastructure. Remote labs enable access to physical resources of real labs by networks – as the Internet – through a simulation software, which replicates the remote environment. Virtual labs have the goal of offering a simulation environment to support virtual experiments. Our proposal can be considered a mixture of the three contexts, since it is a virtual lab that grabs data from the real world and is built over the web.

Laboratories usually offer specialized resources according to the context they are inserted. [1] proposes an educational environment for electronics and electrical engineering. [2] presents an environment for genetics learning. [3] presents a virtual lab of Chemical Vapor Deposition aimed to complement a physical laboratory in the undergraduate course curriculum. [7] proposes a virtual laboratory for medical digital analysis based on grids.

The tool to describe specimens – as Xper2, detailed before, and Lucid (<http://www.lucidcentral.org>) – are designed for specialists and does not follow a laboratory approach. As far as we know, there is no such a tool combining the characteristics of virtual laboratories and biology description/identification tools.

5. CONCLUSIONS

In this paper we present our virtual laboratory based on virtualized visual components. In other to validate it, we have implemented a prototype of the virtual lab to describe mon-

itor lizards of the genus *Varanus* based on a system called Varan-ID. The environment allows direct manipulation of visual components that work as basic elements of an authoring process, to support the identification and description of living beings. These basic elements derives from descriptions present on the Varan-ID database. The proposed tool runs over the web on top of the *Componere* framework.

The main contribution of this paper is our unified approach to produce a virtual lab for taxonomic description, combining the perspective of tools to describe specimens with the virtual laboratory model. It involved the design of a new component based description approach, in which components work as basic descriptive building blocks.

Future works include to expand the laboratory features, enabling it to better integrate with real world resources, i.e., fetching images and other kinds of media of real world specimens, including them in the description process. We are also working to generalize the process of building description components, enabling smooth expansion to new descriptors and other domains.

6. ACKNOWLEDGMENTS

This work was partially funded by CAPES, CNPq, FAPESP, CAPES-COFECUB (AMIB project) and INCT in Web Science (CNPq 557.128/2009-9).

References

- [1] M. Duarte and B. P. Butz. An intelligent universal virtual laboratory (uvl). *IEEE Transactions on Education*, 51(1):2 – 9, 2008.
- [2] K. M. B. et al. Genetics education. *Genetics*, 179:1151 – 1155, 2008.
- [3] M. K. et al. Enhancement of student learning in experimental design using a virtual laboratory. *IEEE Transactions on Education*, 51:76–85, 2008.
- [4] J. Ma and J. V. Nickerson. Hands-on, simulated, and remote laboratories: A comparative literature review. *ACM Comput. Survey*, 38(3):1–24, 2006.
- [5] P. Mabee, M. Ashburner, Q. Cronk, G. V. Gkoutos, M. Haendel, E. Segerdell, C. Mungall, and M. Westerfield. Phenotype ontologies: the bridge between genomics and evolution. *Trends in Ecology and Evolution*, 22(7):345 – 350, 2007.
- [6] Z. Nedic, J. Machotkd, and A. Najhlsk. Remote laboratories versus virtual and real laboratories. *Frontiers in Education, 2003. FIE 2003. 33rd Annual*, 1:1 – 6, 2003.
- [7] T. G. S. Olabarriaga and P. de Boer. A virtual laboratory for medical image analysis. *IEEE Transactions on Information Technology in Biomedicine (TITB)*, 14(4):979–985, 2012.
- [8] A. Santanche, M. Mota, D. Costa, N. Oliveira, and C. O. Dalforno. Componere autoria na web baseada em componentes. *WebMedia*, pages 91–98, 2009.