SIGGI-AACS, a Prototype for Archaeological Artifact Classification Using Computerized Agents

Robert Schlader, Skip E. Lohse, Corey Schou and Al Strickland A

Informatics Research Institute
Idaho State University, Pocatello, Idaho, USA
schlrobe@isu.edu
lohserne@isu.edu
schou@cob.isu.edu
stricka@isu.edu

Abstract. The SIGGI - AACS project is a prototype of an expert system for the identification of archaeological materials. To test the validity of the concept, projectile points were selected as the initial type of object to identify. To generate classifications SIGGI, the computerized agent, utilizes an artificial neural network that has been trained to identify projectile point types from three distinct regional typologies: The Columbia Plateau, the Great Basin, and a General Shape typology. As important as the actual identification of the archaeological materials is, the user interface developed for the interactions between the computer and the user. A set of interfaces has been designed for varying levels of users ranging from the public to dedicated research professionals.
Keywords: Neural Networks, Expert Systems, Auto classification

1. Introduction

The purpose of this update is twofold. First we hope to answer the questions posed to our design group by the archaeological community over the past year about the nature of the Archaeological Auto-Classification System (AACS) under development at the Informatics Research Institute at Idaho State University. Second we are presenting a revised version of the AACS application that is much more user friendly, as well as providing a basic set of instructions for how to use the new version of the software.

2. What are Computerized Agents?

This question is probably the most frequently asked question we have encountered. This term is used in a number of different contexts, and has been poorly defined in most due to a perceived understanding of the underlying concepts. There are in fact two concepts that make up this term, and we will define both.
Agent. The term agent describes an entity that does something on behalf of another.
Computerized Agent. A computerized agent is a piece of software that does something for the user.
Though these simple definitions may seem self explanatory, a couple of examples will help with the understanding of what a computerized agent really is and what they do.
Examples:
A web search engine:
These rather complicated pieces of software go out onto the internet and search through differing databases on behalf of the user to discover web sites that may be of interest to the user based on the set of key words, terms that the user typed into the search engine.
An online reservation system:
These extremely complicated software systems interface with the various computer systems that are used by the companies the user is trying to reserve resources from. They do the hard work of determining which computers to talk to, and what information needs to be moved where.

3. What is an Expert System?

The second most common question received concerns the nature of the software system we have developed. Expert systems capture and provide expert knowledge through software. They provide expert quality advice, diagnoses, and recommendations given real world problems. This is based on the knowledge of real-world experts. This is accomplished by capturing the expert's knowledge into a specialized database called a “knowledge base.”
Examples:
Discovery of tumors in MRI images:
One of the first expert systems ever developed this type of expert system uses the knowledge of hundreds of doctors to aid in the diagnosis of tumors in MRI images.
Facial recognition systems:
This type of expert system is used to not only identify individual users for access to secure facilities, but to isolate the faces of people in images or video in an attempt to discover the identity of people who may have broken the law.

3.1 SIGGI-AACS as an Expert System

AACS is intended to help archaeological investigators by providing an artificial expert in typological classification. This expert augments the knowledge of archaeologists by automating the process of identifying projectile points. This automation takes the form of examining digital images and returning the probability of a typological assignment.

The AACS system has three major components: SIGGI the artificial typological expert (AI), the AACS user interface, and the image manipulation software.

3.2 What is a Neural Network?

A neural network models a biological brain in a software system. A brain is composed of a set of neurons that interconnect into a mesh. Each neuron is composed of a set of structures (Figure 1).

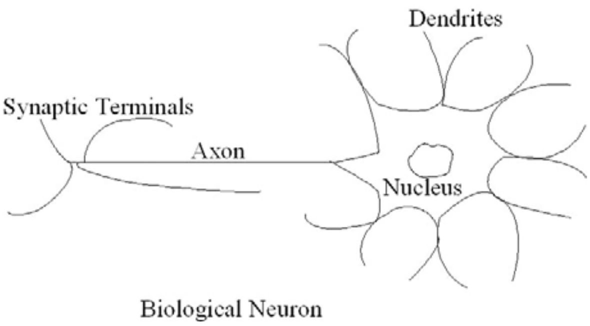

Fig. 1. Model of a neuron.

Within the nucleus of the neuron the cell combines the bioelectrical impulses coming into the cell through its dendrites.
The result of this combination is passed down the cell's axon to the other cells in the brain through the synaptic terminals. The artificial neuron of the neural network re-created all these structures in software (Figure 2).
The AI neuron has a nucleus, which is composed of an activation function. This function sums the values of the inputs to the neuron. The result is passed onto the other neurons in the neural network through the neuron's output. These artificial neurons are arranged into a series of layers that effectively filters the original input data down into more manageable pieces (Figure 3).
These pieces are each mapped to one of the output neurons, which provide an activation value (which we call an assignment index) for the output. Each output from the neural network is a virtual stand in for the individual types of the classification.

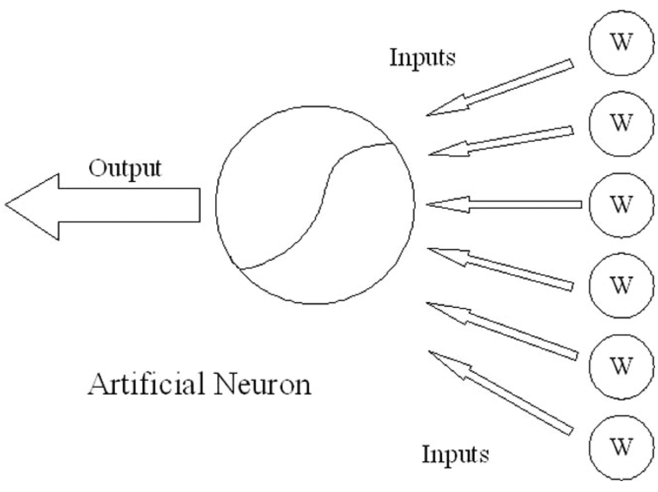

Fig. 2. Artificial neuron in a neural network.

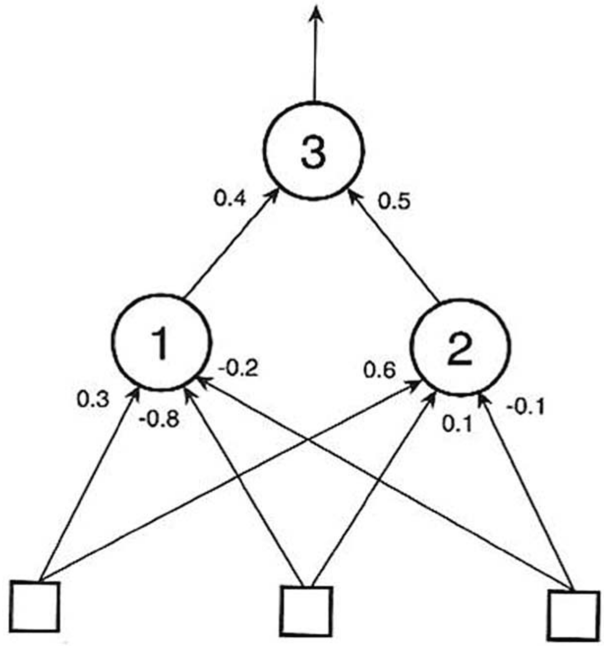

Fig. 3. Data filtering.

4. How Does SIGGI Think?

SIGGI uses a neural network to assign images of artifacts to types in a given classification. The image manipulation software of AACS breaks down the input image into the background and the artifact's edge. The images outline is then digitized into x and y coordinates, from which a random sample is taken. The sample coordinates are given to the neural network as inputs, and the neural network computes an assignment index for each type in the selected classification. The assignment indexes are then converted into a normalized set of probabilities for assignment to the different types.

4.1 What Typologies Does SIGGI Know?

SIGGI has been trained thus far on three Distinct Typological Systems: the Columbia Plateau Typology, the Great Basin Typology, and the General Shape Typology
Columbia Plateau Typology
The Columbia plateau typology is made up of twenty-two distinct types. They are:

- Windust A Windust B Windust C
- Lind Coulee
- Cascade A Cascade B Cascade C
- Cold Springs Side-Notched
- Plateau Side-Notched
- Columbia Corner-Notched A
- Columbia Corner-Notched B
- Quilomene Bar Corner-Notched
- Wallula Rectangular Stemmed
- Quilomene Bar Basal-Notched A
- Quilomene Bar Basal-Notched B
- Columbia Stemmed A
- Columbia Stemmed B
- Columbia Stemmed C

- Mahkin Shouldered
- Nespelem Bar
- Rabbit Island Stemmed A
- Rabbit Island Stemmed B

Great Basin Typology

The Great Basin typology is comprised of fourteen distinct types. They are:

- Gatecliff Stemmed or Shouldered
- Elko Corner Notched
- Rosegate Corner Notched
- Avonlea Side Notched
- Besant Side Notched
- Desert Side Notched
- Northern Side Notched
- Birch Creek Lanceoalate
- Clovis Lanceoalate
- Cottonwood Lanceoalate
- Folsom Lanceoalate
- Hasket Lanceoalate
- McKean Lanceoalate
- Whamuza Lanceoalate

General Shapes

The General Shapes typology is comprised of six distinct types, all defined by the basic shape of the object. They are:

- Simple Lanceoalate
- Complex Lanceoalate
- Corner Removed Triangular
- Corner Notched
- Basal Notched
- Side Notched

5. The New Version of the AACS Application

The new version of the AACS software has updated the interface between the users and SIGGI, the artificial expert. This new interface reduces the number of interactions the user needs to make with the software to five distinct processes: Start, Load / Edit, Threshold, Smooth, and Results.

5.1 AACS Application: Start

In this process, the user chooses which of the other processes the application should stop at and allow the user to interact with.

5.2 AACS Application: Load Edit

This process is where the user loads the image to be processed, and allows the user to open the selected image for editing in the computer's default image editing application.

5.3 AACS Application: Threshold

This process allows the user to set the grey scale value within the image that will be used to distinguish between the background and outline in the image. This value can be set either by moving a slider bar, or simply typing the value into the box provided.

5.4 AACS Application: Smoothing

This process is used to close the outline of the object sothat a proper sample can be collected. Smoothing can be done through a combination of expanding the object's area by one pixel, then contracting it by one pixel or contracting then expanding.

5.5 AACS Application: Results

This process is where the results from SIGGI are displayed. The most likely typological assignment is displayed in red, and the full set of probabilities is shown for reference.

6. How does SIGGI Know?

An explicit forward-chaining and backward-chaining expert system was developed from the knowledge base that was collected and stored into the database. The forward-chaining system is nothing more than the basic steps required to move an object from unknown to identified within a given typology. In essence, this is the methodology used to teach students to conduct typological assignment in the classroom.

6.1 Backward Chaining

The backward-chaining system is the set of all facts about an item that must be true in order for an item to belong to a given type. In order to develop a backward-chaining system the forward-chaining system must already be developed and be in a form where each step in the analysis is either true or false. This binary nature of the forward-chaining system allows the classification system to be walked through in a backward fashion to identify all the "facts" that need to be true for a given result to be true. This is why it is called a backward-chaining system.

SIGGI uses a backward-chaining system for classifying the items presented to it. The validity of the typological classifications comes from both the nature of the backward-chaining system and the assignment index generated by the neural network. Together these two methods reinforce and validate the decision made by the other.

Acknowledgements

We wish to acknowledge the support of the office of the Vice President through the Informatics Research Institute.

References

Bosque, M., 2002. *Understanding 99% of ANN's: Introduction and Tricks.* San Jose, New York, Lincoln, Shanghi, Writers Club Press.

Buckland, M., 2002. *AI Techniques for Game Programming*. Premier Press.

Fausett, L., 1994. *Fundamentals of Neural Networks: Architectures, Algorithms and Applications*. Englewood Cliffs NJ, Prentice Hall.

Freeman, J. and Skapura, D., 1991. *Neural Networks: Algorithms, Applications, and Programming Techniques*. Reading Mass, Addison-Wesley Publishing Co.

Gurney, K., 1997. *An Introduction to Neural Networks*. London & New York, Taylor and Francis Group.

Haykin, S., 1999. *Neural Networks: A Comprehensive Foundation*. 2nd Edition. Upper Saddle River NJ, Prentice Hall.

Jones, T. M., 2003. *AI Application Programming*. Hingham Mass, Charles River Media.

Khanna, T., 1990. *Foundations of Neural Networks*. Reading Mass, Addison-Wesley Publishing Co.

Masters, T., 1993. *Practical Neural Network Recipes in C++*. San Diego CA, Academic Press Inc.