**Proceedings of the ASME 2010 International Design Engineering Technical Conferences &
Computers and Information in Engineering Conference
IDETC/CIE 2010
August 15-18, 2010, Montreal, Quebec, Canada**

DETC2010-290

SUPPORTING BIOMIMETIC DESIGN BY
EMBEDDING METADATA IN NATURAL-LANGUAGE CORPORA

J. Ke   **I. Chiu**   **J.S. Wallace**   **L.H. Shu***

Dept. of Mechanical & Industrial Engineering
University of Toronto
Toronto, Ontario, Canada
*shu@mie.utoronto.ca

ABSTRACT

Biology is a good source of analogies for engineering design. One approach of retrieving biological analogies is to perform keyword searches on natural-language sources such as books, journals, etc. A challenge in retrieving information from natural-language sources is the potential requirement to process a large number of search results. This paper describes how inserting metadata such as part-of-speech, word sense and lexicographical data for each word in a natural-language source can help users identify relevant biological stimuli for biomimetic design. Although this research is still exploratory, initial qualitative observations demonstrate successful identification and separation of biological phenomena relevant to either desired functions or desired qualities. In addition, by incorporating the aforementioned metadata, we can automatically remove search results where search keywords act on abstract nouns or where keywords are used in irrelevant senses. The benefits of embedding metadata are demonstrated through a case study on the redesign of a fuel cell bipolar plate. In this case study, our method can be used to hide 64% of the search results that are unlikely to contain useful biological phenomena, reducing the effort to systematically identify relevant biological analogies.

1   INTRODUCTION

Analogies from conceptually different domains have been observed to result in more creative design solutions (Benami & Jin, 2002). Gordon (1961) noted that the specific domain of biology provides the richest source of direct analogies. Many successful biomimetic designs support the notion that biology is a good source of analogies. Despite the demonstrated usefulness of biological analogies in design, designers are likely limited by their personal knowledge of biology. Linsey et al. (2007) also support the idea that designers require tools and systematic methods to access cross-domain knowledge. Most biomimetic design methods, such as the systematic reverse

engineering of biological systems (Wilson & Rosen, 2007), can benefit from easy access of cross-domain knowledge.

Two main approaches exist for biological knowledge retrieval. One approach to support biomimetic design is to create a database of biological phenomena organized by engineering function (Vincent & Mann, 2002; Lindemann & Gramann, 2004). However, the creation of such a database may be both time consuming and resource intensive. The process may also be subject to the compilers' own knowledge and bias. Additionally, the rapid growth of biological knowledge provides further challenges for the updating of such a system (Rebholz-Schuhmann et al., 2005; Spasic et al., 2003).

Another approach to support biomimetic design involves searching for instances of functional keywords in natural-language biology knowledge sources, e.g., books, papers, etc. Matched text excerpts containing keywords are examined for relevant biological phenomena that can be applied to the engineering problem of interest. This method takes advantage of the extensive biological information already existing in natural-language format. A method was developed to use word collocation and frequency analyses to identify biologically meaningful keywords that bridge the different lexicons of the fields of biology and engineering (Chiu & Shu, 2007). Cheong et al. (2008) translated terms of the Functional Basis into biologically meaningful keywords, which were not obviously related to the functional keywords, to use as search keywords.

1.1 Natural Language Processing

Along with advantages, there are challenges involved with identifying biological analogies in natural-language knowledge sources. One challenge is the potentially high number of matches, some of which are irrelevant. Hacco and Shu (2002), Chiu and Shu (2004), and Shu (2006) described the following situations that may lead to irrelevant matches.