Biologically inspired design (BID) is a method for designing engineering systems by using analogies from the natural world. Given that knowledge of biological systems is normally provided in the form of textual documents, the question becomes how to extract design knowledge about biological processes from scientific articles? Among the difficulties that arise is relating the vocabulary of the engineer needing ideas with that of the biologist authoring the documents. Ontologies can play an important role in addressing this problem. In this talk, we present IBID, a semi-automated approach for extracting models of biological systems from documents describing them. IBID's use of ontologies and the reasoning approaches it uses, subsumption and analogy, are discussed as well as applications of the IBID approach to other domains.
Background

- Software Engineering: Dowser
- Artificial Intelligence: KA
• Automatic construction of UML class model diagrams from textual system descriptions

• Ontologies used: UML, WordNet

• Reasoning: rule-based

• [http://www.cc.gatech.edu/projects/dowser/](http://www.cc.gatech.edu/projects/dowser/)
UML: Ontology for System Modeling

• Unified Modeling Language

• De facto standard modeling notation in software engineering

• Structure and behavior of natural and artificial systems

• Constraint-based semantics

• http://www.omg.org/spec/UML/
WordNet:
Ontology of Verbal Concepts

• Based on a lexicon of common English nouns, verbs, adjectives and adverbs

• Semantics of a word is the set of synonyms (synset) to which it belongs

• Relationships: hypernymy/hyponymy, antonymy, and meronymy/holonymy

• Provides the ability to deal with concepts and not just words

• http://wordnet.princeton.edu
• Automatic construction of system descriptions from *The Way Things Work*

• Ontologies used: SBF

• Reasoning method: analogical reasoning

• [http://dilab.gatech.edu](http://dilab.gatech.edu)
SBF: Ontology for Mechanical Designs

- **Structure-Behavior-Function** notation for mechanical designs
  - (S) Elements and substances comprising a design solution and the connections among them
  - (B) Mechanism by which the solution is expressed
  - (F) Output of the system being modeled; purpose
  - Hierarchical nesting; systems made of subsystems
- Logical and operational semantics
- Extended to deal with biological systems

Biologically Inspired Design

• Use of biological processes as solution approaches to engineering problems
Problem

• How to support engineers searching for inspirational biological processes in text documents, specifically academic biology articles

• Issues:

  • Vocabulary differences between the engineer and the biologist author
  
  • How to locate similar rather than identical matches
  
  • Relevance checking of retrieved articles
Approach

• Domain-independent ontology of functions, structures and behaviors
  • Mapping of biological vocabulary onto them
• Document analysis using WordNet and VerbNet
• Representation via SBF models
• Analogical reasoning, in addition to subsumption
Function Ontology

- Domain-independent ontology of function concepts
- Constructed by merging existing ontologies
  - Functional Basis
  - Biomimicry Taxonomy
  - SBF
- Provides a hierarchical controlled vocabulary for retrieval by engineers
- Issue: Force the user into a way of thinking versus allowing them free rein (faceted search)
Functional Basis: Ontology for Mechanical Designs

• Mechanical engineering
• Itself a synthesis of functional ontologies
• Independently applied to biologically inspired design
• http://triumph.mie.uic.edu/ME250/handouts/wood_stone_function_basis_jmd.pdf
Biomimicry Taxonomy: Taxonomy for Biological Functions

- Biomimicry Institute
- Function related
- Fronts a database of biological strategies
Domain-Dependent Ontologies

- Collection of biology terms
- Required segregation along SBF lines
- Required alignment with the domain-independent SBF ontologies
- Issue: largely manual and time-consuming effort
Document Analysis

- Preprocessing of biology articles
- Extraction of biological processes
- Represented as SBF models
- Subsequent retrieval using faceted search
- Issue: Differentiate of biological processes from experimental ones
Analysis Details

- Sentence-by-sentence natural language processing
- For each sentence, determination of root verb
- Goal: map root verb to closest element of the domain-independent function ontology
- Use of WordNet relations (e.g. synonymy and hypernymy)
- Use of VerbNet thematic roles to provide semantics
- Issue: Word sense disambiguation
VerbNet: Ontology of Verb Contexts

- Lexicon of verbs and their *semantic roles*
- Word senses, semantic restrictions, synonyms, entailment, active/passive voice
- Enables roles to be assigned to constituents of a parsed sentence
- [http://verbs.colorado.edu/~mpalmer/projects/verbnet.html](http://verbs.colorado.edu/~mpalmer/projects/verbnet.html)
IBID

• Interactive tool for Intelligent Biologically Inspired Design

• Search through articles for suitable biological processes

• Annotation of semantic features

• Analogical matching
Architecture
Annotation

• Use of PDF toolkit to highlight areas of text that correspond to SBF model elements

• Aid readers in making relevance checks

In addition, I survey the ways that fruits and plant parts are modified to adhere to animals. I also put forth new hypotheses predicting how morphological characteristics of adhesive fruits influence both the grooming behavior of dispersal agents and the fruit retention times. Experimental data are provided to support these hypotheses. Suggestions for tests of hypotheses and future work on adhesive-fruit dispersal are provided wherever possible.

For convenience, all dispersal units are called *fruits*. [A fruit is the matured gynoecium with or without other floral organs or parts of organs. A *seed* is either the embryo embedded in endosperm, or the embryo and digested endosperm within the megasporangium, surrounded by one or more integuments (42).] Most adhesive fruits **adhere** to animals by barbs, hooks, or viscid outgrowths. In this review, fruits that adhere by barbs or hooks are called *burrs*. Fruits that adhere by viscid outgrowths are called *viscid fruits*. I
Analogical Reasoning

- Typical search engine: lexical retrieval (keywords)
- Basic IBID search: semantic based (domain-independent concepts)
- IBID analogical search: similarity based
- *Analogy*: mapping from one instance to another, conserving relations
- Uses Structure Mapping Engine from Northwestern
- Encode SBF models as a set or relations; does same for query; look for similar relationship patterns
IBID: Value Added

• Increased search recall because of synsets
• Increased search precision due to controlled vocabularies
• More efficient relevance checking because of annotations
• Richer search results due to analogical matching
• Issue: validation vs. gold standard
Extensions to Other Domains

• Finance: reports from analysts in support of question answering

• Cancer: articles on cancer pathways to find supportive or conflicting results
Future Work

• Targets: UML and SBF are examples of target notations
  • Issue: to what extent do the semantics of the target influence the analysis process?

• Natural language expression of problem vs. controlled vocabulary
  • Issue: tradeoff between precision of controlled vocabulary and convenience of nature language

• Question answering within documents
  • Issue: to what extent can the IBID approach be applied within an article to answer reader questions?