Representing Core Themes in a Biology E-Textbook

Author	**Nikhil Dinesh**
Supervisors	Vinay Chaudhri
Stage	Graduated in 2010, with a PhD in computer science (focus on logic and NLP). Working with ontologies since then.
Affiliation	SRI International
E-Mail	dinesh@ai.sri.com

Aims and Objectives of the Research

Inquire [1] is a new electronic textbook that integrates the Campbell biology textbook [2] with a reasoning system and a rich biology knowledge base (KB), allowing it to answer a variety of questions. It offers new ways for students to explore and interact with educational materials, and ultimately, improve their understanding of biology.

The KB and question answering system are crucial to many of Inquire's interactive features. Subject matter experts (SMEs), who are biology teachers, encode knowledge in the KB, using the process described in [3]. A focus of our research is to provide the SMEs with modelling guidelines or patterns, so that we may systematically acquire knowledge from the textbook. In the following section, we narrow our focus to a particularly challenging class of problems -- modelling the core themes in the textbook.

Justification for the Research Topic

The Campbell biology textbook is organized into eight core themes – relating structure to function, energy transfer, regulation, science as a process, continuity and change, evolution, and science, technology and society. We elaborate on two of these core themes below:

1. Relating Structure to Function – There is a correlation between the structure of biological entities and their functions. For example, the round shape of a red blood cell facilitates its function to transport oxygen.
2. Energy Transfer – In various biological processes, energy is transferred or transformed, and then used to do work. For example, photosynthesis transforms sunlight into chemical energy in the form of glucose and oxygen.

An explicit representation of these core themes in the KB and e-textbook will enable a student to explore it systematically. The core themes are not independent of each other. For example, the function of chlorophyll is to absorb light, which is also a transfer of energy from light to chlorophyll's electrons. However, in designing the ontology, question-answering and presentation methods, we have found a sufficient number of orthogonal issues in each core theme to warrant a separate exploration.

Research Questions

To represent a core theme, we need to answer the following questions:

1. What is the definition of the core theme?
2. What questions do we wish to answer about the core theme?
3. Given the definition and the questions, what should the KB contain?
4. How do we retrieve and present answers given the information in the KB?

Research Methodology

We adopted the following steps to handle the core theme relating structure to function:

1. Requirements Specification – SMEs define the core theme in biology terms. We then conduct a user study, where the definition of the core theme is discussed with teachers and students. During the study, we gather questions about the core theme.
2. Design – The questions are categorized into a set of templates. We then develop modelling vocabulary and guidelines, targeted at answering these questions.
3. Implementation -- The SMEs adopt the modelling guidelines during encoding. Algorithms to retrieve and present answers are developed. Based on the algorithms, we can classify the questions into four types – (a) descriptions (b) similarities and differences, (c) path-based relationships, and (d) slot-value queries.
4. Evaluation – Users (students and teachers) assess the quality of answers.

Research Results to Date

We have designed and implemented a representation for the core theme relating structure to function. An evaluation is planned in the near future. Here, we briefly describe the main relations that used for this core theme:

1. Structure is described using a set of meronymic relations (e.g., has-part, has-region), spatial relations (e.g., is-inside, is-outside), and properties (e.g., area, diameter). These relations are provided by the component library [4], our upper ontology.
2. Functions are described using the relation “has-function”. Our approach is similar in spirit to the proposal by Burek et al [5]. However, we do not distinguish between functions and functionings.
3. The relationship of structure to function is encoded in two ways. First, we have the relation “facilitates” which is used to state that a sub-structure is important to the function of a super-structure [6]. For example, the absorption of light by chlorophyll-A is facilitated by its poryphrin ring. Second, we have qualitative relations [7], which relate structural to functional properties. For example, the length of the loop of Henle is directly proportional to the rate of water re-absorption.

The SMEs are given guidelines to identify the relations based on sentences in the textbook. To answer questions, we need to account for various subtleties in terms of the transitivity of parts and in transferring functions from substructures to superstructures. Due to space restrictions, we omit a discussion of these issues.

References

1. Spaulding, A., Overholtzer, A., Pacheco, J., Tien, J., Chaudhri, V., Gunning, D., Clark, P.: Inquire for iPad: Bringing Question-Answering AI into the Classroom. In : Conference on AI in Education (2011)
2. Reece, J., Urry, L., Cain, M., Wasserman, S., Minorsky, P., Jackson, R.: Campbell Biology 9th edn. Pearson (2010)
3. Chaudhri, V., Dinesh, N., Pacheco, G., Ng, G., Clark, P., Goldenkranz, A., Seyed, P., Sharma, N.: Preliminary Steps towards a Knowledge Factory Process. In : Conference on Knowledge Capture (2011)
4. Barker, K., Porter, B., Clark, P.: A Library of Generic Concepts for Composing Knowledge Bases. In : Conference on Knowledge Capture (2001)
5. Burek, P., Hoehndorf, R., Loebe, F., Visagie, J.: A top-level ontology of functions and its application in the Open Biomedical Ontologies. Bioinformatics 22(14), e66-e73 (2006)
6. Umeda, Y., Tomiyama, T.: Functional Reasoning in Design. IEEE Expert 12(2), 42-48 (1997)
7. Forbus, K.: Qualitative Process Theory. Artificial Intelligence 24, 85-168 (1984)