Conceptual Models of Energy Transfer and Regulation

Vinay K. CHAUDHRI [a], Nikhil DINESH [a], and Stijn HEYMANS [a]
[a] *Artificial Intelligence Center, SRI International, Menlo Park, CA 94025*

Abstract. The rich descriptions that characterize biological processes can be expressed from multiple perspectives. One basic perspective involves specifying the process steps, how they are ordered, their participants, and how participants are involved in different steps. Two additional perspectives on processes are how energy is transferred and how the processes are regulated. Such perspectives are essential for understanding the basic mechanisms of biology. We consider a representation of energy transfer and regulation for biological processes, and consider several example questions that can be answered using this representation. Our work is driven by the description of biological processes in an introductory biology textbook and provides the conceptual design for curation of a knowledge base (KB) for an education application.

Keywords. process representation, multiple perspectives, energy transfer, regulation, educational ontology

Introduction

Creating conceptual models from a biology textbook has profound implications for both ontology research and student learning. For ontology research, it provides a circumscribed focus for making ontological decisions: the students studying from a textbook are expected to make certain distinctions and can be tested by asking them questions that have objective answers. Because the knowledge in a textbook is foundational and prepares a student for a variety of follow-up courses, laboratory work, and real-life situations, it provides an ideal test bed for developing reusable and multifunctional representations that have a high degree of consensus. Such conceptual models also have great utility in education because they can be used for answering questions in an intelligent textbook that has been shown to improve student learning [5].

A biology textbook such as *Campbell Biology* [16] contains rich descriptions of processes that capture how different mechanisms work. Representation of processes has been an active area of interest in knowledge representation, and reasoning (KRR), upper ontologies, and natural language processing (NLP). KRR researchers have developed a variety of action representation languages that can be used for modeling processes [2]. Most upper ontologies such as DOLCE [11], SUMO [15] and Cyc [13] support Event (or a comparable concepts such as Occurrent or Perdurant) as a distinction in the ontology and a variety of associated relationships. The NLP community has developed case roles and created event lexicons such as VerbNet [17] and FrameNet [1].

In our work, we have been using an upper ontology called Component Library (CLIB) [4]. CLIB defines a small number of basic distinctions, and has a vocabulary of actions and semantic relationships that domain experts have found easy to use for encoding knowledge from a biology textbook [12]. Although CLIB provides good vocabulary to encode the basic process structure (i.e., steps in a process and their relationships) and participants (i.e., entities that participate in different steps), it does not provide adequate guidelines and vocabulary for dealing with process regulation and energy transfer. After surveying many of the available ontologies, we found that none addressed those concepts adequately.

Accordingly, our focus here is on presenting an approach for conceptual modeling of energy transfer and regulation. We begin by providing background information on CLIB. We then define what is meant by energy transfer and regulation in the context of biological processes. To derive these definitions, we turn to three sources: the advanced placement standard defined by the College Board in the United States, the biology textbook, and biology teachers. We next define vocabulary to represent these concepts and provide illustrative sample representations. Our approach to defining the concepts of energy transfer takes into account how the textbook and biology teachers define these concepts and then state them from a knowledge engineering perspective. We also consider a few example questions that can be answered using these representations. We conclude with a discussion on unresolved problems and directions for future research.

1. Component Library

CLIB is a linguistically motivated ontology designed to support representation of knowledge for automated reasoning [4]. CLIB adopts four simple upper level distinctions: (1) *entities* (things that are), (2) *events* (things that happen), (3) *relations* (associations between things), and (4) *roles* (ways in which entities participate in events).

In CLIB [4], the class Action has 42 direct subclasses, with 147 subclasses in all. Examples of direct subclasses include Attach, Impair, and Move. Other subclasses include Move-Through (which is a subclass of Move), and Break (which is a subclass of Damage, which is a subclass of Impair). To ensure generality, these subclasses were developed by consulting lexical resources, such as WordNet, the Longman Dictionary of Contemporary English [18] and Roget's Thesaurus [14].

CLIB provides semantic relationships to define the participants of an action. These relations are based on a comprehensive study of case roles in linguistics [3] and include *agent*, *object*, *instrument*, *raw-material*, *result*, *source*, *destination*, and *site*. (The syntactic and semantic definitions we developed for these relations are available elsewhere [7].) As an example, we consider the definition of *raw-material*. The semantic definition of *raw-material* is that it is any entity that is consumed as an input to a process. The syntactic definition of *raw-material* is that it is either the grammatical object of verbs such as *to use* or *to consume*, or it is preceded by *using*.

2. Modeling Energy Transfer

We will first define energy transfer, and then introduce the conceptual vocabulary for representing energy concepts, followed by example competency questions.

2.1. Defining Energy Transfer

The College Board course description defines energy as the capacity to do work [9]. Energy transfer is considered as a core theme in biology: all living organisms are active (living) because of their abilities to link energy reactions to the biochemical reactions that take place within their cells. For example, the energy of sunlight, along with carbon dioxide and water, allows plant cells to make organic materials, synthesize chemical energy molecules, and ultimately release oxygen to the environment. Campbell Biology introduces this topic by saying *Life requires Energy Transfer and Transformation.* Here, *energy transformation* is a new term. While discussing these definitions with biology teachers, it was apparent that they indeed make a distinction between energy transfer (when energy changes location), and energy transformation (where energy changes form). In fact, the biology teachers' view was that most processes involve both energy transfer and transformation, but that in some cases one may be more important than the other. The biology teachers use the concept of energy flow to encompass both energy transfer and transformation. Thus, based on the initial analysis we needed to define: energy, energy transfer, energy transformation, and energy flow.

Energy: We define energy as ability or capacity to do work. Energy has a variety of forms, such as potential energy, kinetic energy, and light.

Energy transformation: A change in energy from one form to another (e.g., from potential energy to kinetic energy).

Energy transfer: A change in the location and/or possession of energy (e.g., transfer of light energy from the sun to the earth).

Energy flow: A combination of energy transfers and transformations (and other energy flows) (e.g., transfer of energy from the sun to the plants, which transform it into chemical energy, which is then transferred to other organisms).

2.2. Representing Energy Transfer

As mentioned above, we define energy as the ability or capacity to do work. In particular, we make Energy a subclass of Tangible-Entity. We based this choice on an analysis of sentences in Campbell Biology that treat energy as an independent object that is transferred between entities, and can serve as an input or output to chemical reactions just like other chemical entities. We use the relation *possesses* to relate other objects to energy (e.g., a Chemical-Entity *possesses* Chemical-Energy.

Because energy transfer is defined as a change in the location or possession of energy, we place it as a subclass of event Transfer in CLIB, by specializing the *object* relationship to be of a type of Energy.

Because energy transformation is defined as a change in the form of energy, we place it in the CLIB taxonomy as a subclass of Transform, which is a subclass of Change. We use the relations *raw-material* and *result* to specify the initial and final forms of energy, respectively.

Because energy flow is a combination of energy transfers, transformations, and other energy flows, we place it high up in the CLIB taxonomy as a subclass of Physical-Process. An energy flow process can have three kinds of subevents: energy transfer, energy transformation, or energy flow. Its required relationships are *raw-material* to denote the initial form of energy and *result* to denote the final form of energy. Other common relationships associated with an energy flow are —*donor* to denote the entity that possesses the energy prior to the energy flow, *recipient* to denote the entity that possesses the energy after the flow, *during* to state that one process happens during another (without being a step of the second process). In the next section, we illustrate how these concepts can be used to represent flow of energy during light absorption.

2.3. An Example Use of Energy Concepts

First, consider Figure 1, which defines the concept of Light-Absorption as a process with an Electron as its *instrument*. Additionally, the *base* and *agent* of light absorption is a Molecule; its *object* is Light.

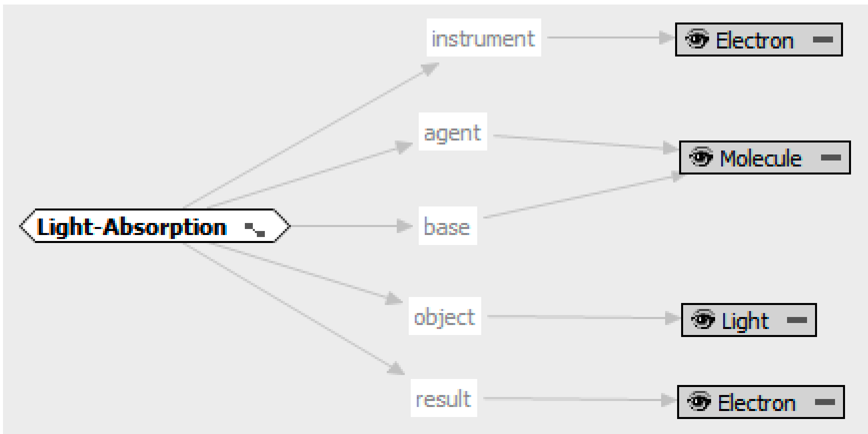

Figure 1. Light Absorption

In the graph in Figure 1, the white node (i.e., Light-Absorption), is universally quantified, and every other node is existentially quantified. We can formally state it in first-order logic as follows:

$$\begin{aligned}\forall x : \mathsf{Light\text{-}Absorption}(x) \Rightarrow\ & \mathit{instrument}(x, la_1(x)) \land \mathsf{Electron}(la_1(x)) \\ & \land\ \mathit{agent}(x, la_2(x)) \land \mathsf{Molecule}(la_2(x)) \\ & \land\ \mathit{base}(x, la_2(x)) \\ & \land\ \mathit{object}(x, la_3(x)) \land \mathsf{Light}(la_3(x)) \\ & \land\ \mathit{result}(x, la_4(x)) \land \mathsf{Electron}(la_4(x))\end{aligned} \tag{1}$$

In axiom (1), we use Skolem functions la_i, $1 \leq i \leq 4$, to represent the *existence* of 2 Electrons, 1 Molecule, and 1 Light.

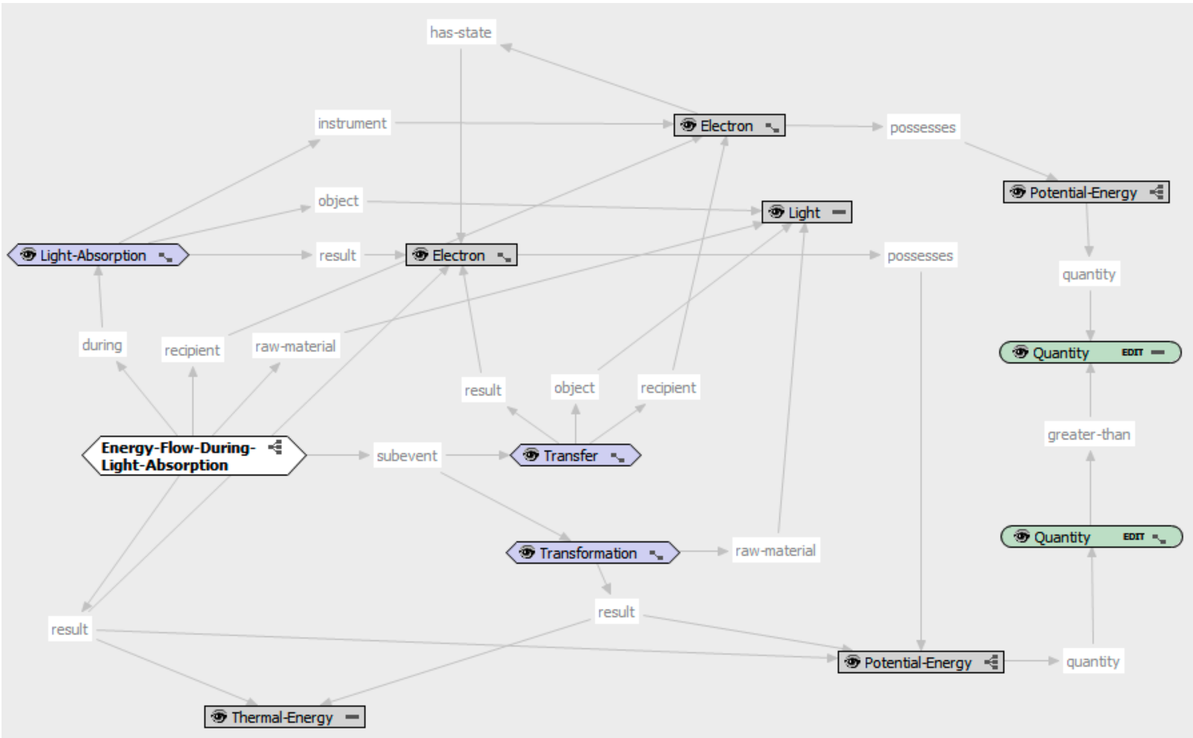

Figure 2. Energy Flow during Light Absorption

Next, in Figure 2, we show an example representation of energy flow during light absorption. When a molecule absorbs light, its electrons are excited (increase in potential energy). Energy (in the form of light) is transferred to the electrons. The overall concept of Energy-Flow-During-Light-Absorption has Light as its *raw-material*, and Electron as its *recipient*, and its *result*s are Potential-Energy and Thermal-Energy. Furthermore, it has two substeps: Transfer in which energy is transferred to electrons (from the sun or another light source), and Transformation during which energy is transformed from Light to Potential-Energy and Thermal-Energy. We also indicate that Energy-Flow-During-Light-Absorption takes place *during* the overall process of Light-Absorption, and further that, the Electron that is the *result* after the absorption has a higher Potential-Energy than the Electron that is its *raw-material*.

In the logical representation of the above graph shown in axiom (2) on the next page, we use Skolem functions named e_i, $1 \leq i \leq 11$. We indicate that some of the Skolem functions in this definition refer to the same individuals as defined by the axiom (1), by equality statements of the form $e_2(x) = la_1(e_1(x)$. Instead of using these equality statements, we could just use the nested Skolem functions directly in the formula, but doing so would make the formulas less compact and harder to read.

$$
\begin{aligned}
&\forall x: \mathsf{Energy\text{-}Flow\text{-}During\text{-}Light\text{-}Absorption}(x) \Rightarrow \\
&\quad \mathit{during}(x, e_1(x)) \wedge \mathsf{Light\text{-}Absorption}(e_1(x)) \\
&\quad\quad \wedge\ \mathit{instrument}(e_1(x), e_2(x)) \wedge \mathsf{Electron}(e_2(x)) \wedge e_2(x) = la_1(e_1(x)) \\
&\quad\quad\quad \wedge\ \mathit{has\text{-}state}(e_2(x), e_3(x)) \\
&\quad\quad\quad \wedge\ \mathit{possesses}(e_2(x), e_4(x)) \wedge \mathsf{Potential\text{-}Energy}(e_4(x)) \\
&\quad\quad\quad\quad \wedge\ \mathit{quantity}(e_4(x), e_5(x)) \wedge \mathsf{Quantity}(e_5(x)) \\
&\quad\quad \wedge\ \mathit{object}(e_1(x), e_6(x)) \wedge \mathsf{Light}(e_6(x)) \wedge e_6(x) = la_3(e_6(x)) \\
&\quad\quad \wedge\ \mathit{result}(e_1(x), e_3(x)) \wedge \mathsf{Electron}(e_3(x)) \wedge e_3(x) = la_4(e_1(x)) \\
&\quad\quad\quad \wedge\ \mathit{possesses}(e_3(x), e_7(x)) \wedge \mathsf{Potential\text{-}Energy}(e_7(x)) \\
&\quad\quad\quad\quad \wedge\ \mathit{quantity}(e_7(x), e_8(x)) \wedge \mathsf{Quantity}(e_8(x)) \wedge\ \mathit{greater\text{-}than}(e_8(x), e_5(x)) \\
&\quad \wedge\ \mathit{recipient}(x, e_2(x)) \\
&\quad \wedge\ \mathit{raw\text{-}material}(x, e_6(x)) \\
&\quad \wedge\ \mathit{result}(x, e_9(x)) \wedge \mathsf{Thermal\text{-}Energy}(e_9(x)) \wedge\ \mathit{result}(x, e_2(x)) \\
&\quad \wedge\ \mathit{subevent}(x, e_{10}(x)) \wedge \mathsf{Transformation}(e_{10}(x)) \\
&\quad\quad \wedge\ \mathit{result}(e_{10}(x), e_7(x)) \wedge\ \mathit{result}(e_{10}(x), e_9(x)) \wedge\ \mathit{raw\text{-}material}(e_{10}(x), e_6(x)) \\
&\quad \wedge\ \mathit{subevent}(x, e_{11}(x)) \wedge \mathsf{Transfer}(e_{11}(x)) \\
&\quad\quad \wedge\ \mathit{result}(e_{11}(x), e_3(x)) \wedge\ \mathit{object}(e_{11}(x), e_6(x)) \wedge\ \mathit{recipient}(e_{11}(x), e_2(x))
\end{aligned}
\tag{2}
$$

Axioms (1) and (2) could be viewed as multiple perspectives on the same process. Axiom (1) is a view that captures the process structure and its participants. Axiom (2) focuses on the energy transfer aspects of the process. The two perspectives are related through the use of the *during* relationship, and the sharing of Skolem functions. Use of such multiple perspectives with an ability to define relationships across them gives us an ability to factor the representation of a complex process into separate conceptual chunks that are easier to formalize and understand.

3. Modeling Process Regulation

We first define what is meant by regulation. We then introduce conceptual vocabulary for modeling it and then discuss competency questions.

3.1. Defining Process Regulation

The College Board curriculum introduces regulation by stating that *everything from cells to organisms to ecosystems is in a state of dynamic balance that must be controlled by positive or negative feedback mechanisms (e.g., body temperature is regulated by the brain via feedback mechanisms.)* The definition in Campbell Biology adds more detail by explaining feedback regulation. In feedback regulation, the output, or product, of a process regulates that very process. The most common form of regulation in living systems is negative feedback, in which accumulation of an end product of a process slows that process. For example, the cell's breakdown of sugar generates chemical energy in the form of a substance called adenosine triphosphate (ATP). When a cell makes more ATP than it can use, the excess ATP feeds back and inhibits an enzyme near the beginning of the pathway. Although less common than processes regulated by negative feedback, many biological processes are also regulated by positive feedback, in which an end product speeds up its own production. The clotting of blood in response to injury is one example. When a blood vessel is damaged, structures in the blood called platelets begin to aggregate at the site. Positive feedback occurs as chemicals released by the platelets attract more platelets. The platelet pileup then initiates a complex process that seals the wound with a clot. Biology teachers view process regulation as a dynamic phenomenon that maintains equilibrium of properties in all things from the cell to the ecosystem. Changes in the quantities in these systems drive the need for regulation. Signal and messenger molecules like hormones and regulators provide feedback to maintain the balance of properties. Rate changes in metabolic cycles and reactions occur through molecular interaction.

As the preceding discussion indicates, we can see that feedback mechanisms are fundamental to process regulation. The definitions used by biology teachers tend to be highly specific in regard to the different ways such feedback mechanisms occur in biology. Our goal, however, was to define regulation to allow design of a generic modeling pattern that fulfills the representation needs for biology but is also generic. Therefore, we adopted the following definition for process regulation: Biological systems need their properties (e.g., body temperature, quantity of available ATP molecules) to stay within certain limits. External factors (e.g., exposure to cold, consumption of ATP) may push these properties outside their desired limits. A regulatory mechanism adjusts the properties so that they return to their desired levels.

Given these considerations, we define the representation pattern for regulation to have the following elements: (1) regulated property (e.g., temperature, blood pH level); (2) limits of the property (i.e., the regulated property needs to be kept within certain limits for the organism to survive); (3) stimuli — events that can cause the property to move outside the desired limits (e.g., exercise can cause an increase in temperature); (4) mechanisms — events triggered by the stimuli (e.g., an increase in temperature triggers sweating); (5) opposing effects — the response mechanisms that oppose the effect of the stimuli on the property (e.g., sweating causes evaporation).

It is helpful to understand the relationship between process regulation and causality [10]. Causality is a statement about one event that directly influences another event. For example, "the attachment of RNA Polymerase causes transcription to begin" is a statement that introduces the causal relationship between two events. Regulation addresses more detail than causality: specifically, it tells us about the relationships to properties and how an event or outcome might be modified. For example, the statement "the abundance of RNA Polymerase has a positive impact on the rate of transcription" is a statement about regulation.

3.2. Representing Process Regulation

We define regulation as a process and place it as a subclass of Biological-Process. Its subclasses correspond either to regulation of a property of an entity or to regulation of the property of a process. For example, Thermoregulation is a regulation process that regulates body temperature. Regulation-Of-Glycolysis is a regulation process that regulates the rate of Glycolysis. Some regulatory processes such as Thermoregulation have biological names; others, however, such as

Regulation-Of-Glycolysis do not have biological names and thus are named using suitable naming conventions. In many situations, even the regulatory processes that do have biological names such as Thermoregulation may need synthesized names such as Thermoregulation-In-Human.

We introduce a new relation *object-property* with domain Regulation and range Property-Value. It is required to specify the *object-property* for all subclasses of the class Regulation. For example, for the regulation process Thermoregulation-In-Human, the *object-property* is the temperature of the human.

We next introduce the relations to specify the limits of the regulated property. We use the relation *from-value* and *to-value* to encode the limits (or ideal range) of the regulated property. For example, for the Thermoregulation-In-Human, body temperature must remain between 36° and 37° Celsius. Because the temperature may indeed go outside these limits, the limits cannot be specified as integrity constraints. The intent of these properties is to represent the ideal range for the regulated property. In many situations, the textbook may be silent on the ideal range of a regulated property value; for example, it does not give the ideal limits for Regulation-Of-Glycolysis.

A stimulus is an event that is external to the regulated process that brings the regulated property out of the desired range. We model a stimulus for regulation as an increase or decrease in a property, which we represent with CLIB events Increase and Decrease, respectively. There are two cases to consider. In the regulation of properties of entities, the stimulus is always an increase or decrease in the regulated property. In the regulation of rates of processes, the stimulus can be an increase or decrease in the *raw-material* or a *result* of the process or a related process.

A mechanism of regulation is a process that is directly responsible for increasing or decreasing the regulated property to restore its value to desired limits. For example, for Thermoregulation-In-Human, the mechanisms are secretion of sweat and vasodilation, both of which decrease temperature. These mechanisms are in turn triggered by hormones secreted by the hypothalamus. However, the secretion of hormones is not considered a mechanism because the effect on the regulated property is indirect.

When a stimulus occurs, it may trigger a series of events that lead to an effect that opposes the stimulus. This pattern is described using a causal chain — a series of events linked by the *causes* relation. We divide the discussion into two parts (1) causal chain from stimuli to mechanisms, and (2) causal chain from mechanisms to effects. If some parts of the causal chain are not described in Campbell Biology (as discussed above), they may be omitted, as long as one coherent causal chain remains. For example, we may have only causal the chain from Stimulus to Mechanism, or only one from Mechanism to Opposing Effect.

3.3. Example Uses of Regulation Concepts

We will consider two example uses of the regulation concepts we have introduced. To illustrate regulation of a property of an entity, we consider Thermoregulation-In-Human, and to illustrate the regulation of a process, we consider Regulation-Of-Glycolysis.

3.3.1. Regulation of the Property of an Entity

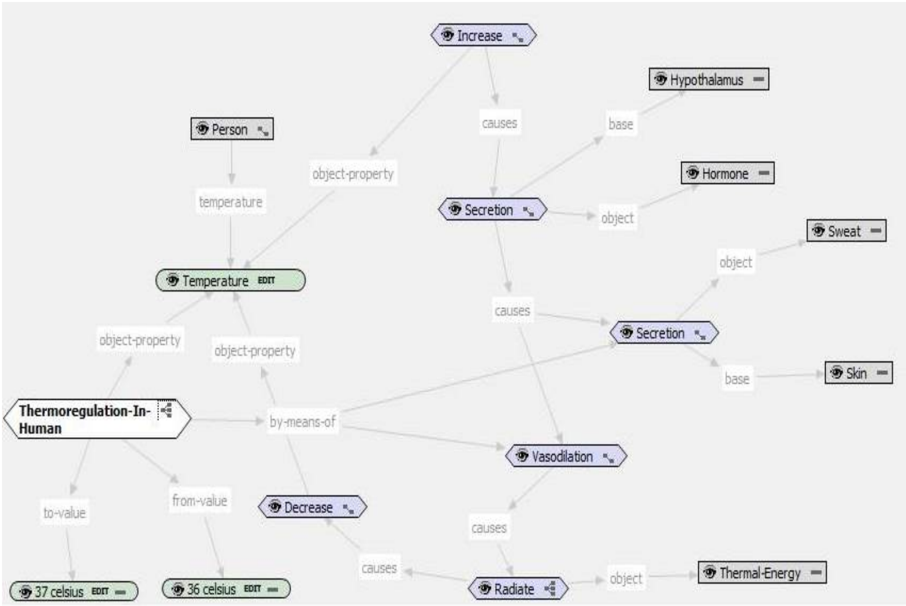

Figure 3. Thermoregulation in Humans

Figure 3 presents a representation of how the body temperature of a human is regulated. Because the logical meaning of these concept graphs can be interpreted in a manner similar to the earlier concepts, we omit the detailed logical axioms for them, limiting our explanation to conceptual representation. The regulated property (i.e., the temperature of a human body) is related to the regulation process using the *object-property* relation. We use the relations *from-value* and *to-value* to encode the limits (or ideal range) of the regulated property. Figure 3 shows that the temperature in humans is regulated between between 36° and 37° Celsius. The Stimulus could be either increase or decrease in temperature, but for illustration we show increase in temperature only. The mechanisms are related to the regulation process using relations *by-means-of*. We model two mechanisms of thermoregulation in humans: (1) secretion of sweat, and (2) vasodilation–both of which decrease temperature. The increase in body temperature triggers a Hormone secreted by the Hypothalamus that in turn triggers the Secretion of Sweat, thus, creating a causal chain from Stimulus to Response. The Secretion of Hormone is not considered a mechanism because its effect on the regulated property is indirect. As an illustration of a causal chain from a mechanism to the regulated property, we show that Vasodilation *causes* Radiation, which in turn *causes* a Decrease in body temperature. Thus, the Stimulus and Response have opposing effects on the regulated property.

3.3.2. Regulation of the Property of a Process

Figure 4 shows the regulation of Glycolysis. Here, the regulated property is the rate of Glycolysis. Because the textbook does not provide information on the

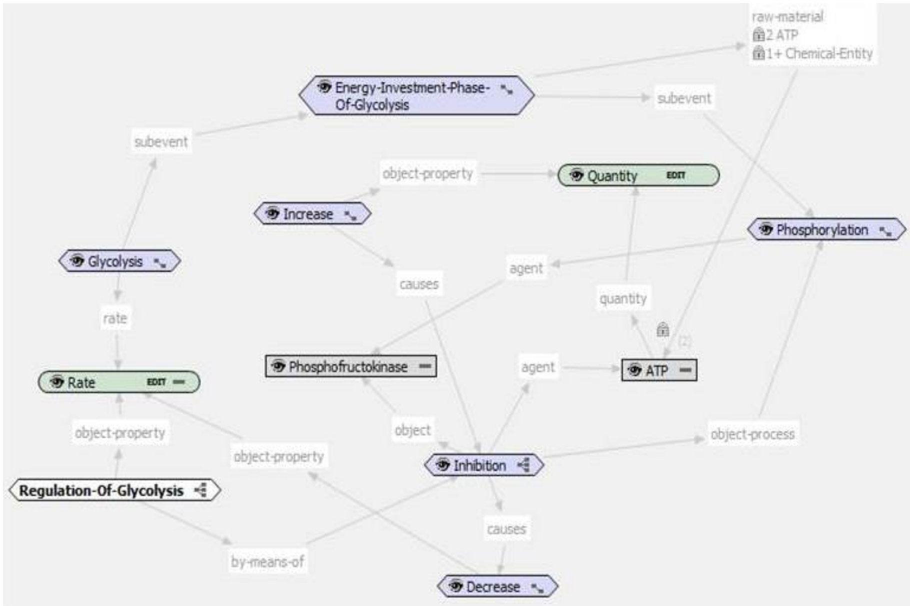

Figure 4. Regulation of Glycolysis

limits at which the rate should be maintained, the limits are omitted from the representation. The external stimuli on the rate of Glycolysis can be an increase in the concentration of ATP, adenosine monophosphate (AMP), or Citrate. For our example, we consider only the increase in the concentration of ATP. A mechanism for regulating the rate of glycolysis is the Inhibition of Phosphofructokinase, which is the *agent* of the Phosphorylation step of the Energy-Investment-Phase of Glycolysis. The Response mechanism of Inhibition of Phosphofructokinase directly causes a Decrease in the rate of Glycolysis. The representation constructs used are the same as in Thermoregulation-of-Human, except the representation of the regulation of Glycolysis considered here can be considered another perspective on Glycolysis. Furthermore, the representation of regulation of Glycolysis makes references to subevents and the participants of Glycolysis.

4. Using Representations for Answering Questions

We describe our question development process, indicate different categories of questions, and illustrate sample answers.

4.1. Question Development

Because our goal is to use these conceptual models in an intelligent textbook, we needed to identify a set of educationally useful questions. To determine the questions that would be useful and interesting to answer, we first convened a focus group of teachers and students who generated a list of questions that they considered both typical and educationally useful. Next, we analyzed those questions so that they could be mapped to well known computational approaches for answering them. Several example questions we gathered during the process, the reasoning approach, and sample answers follow.

4.2. Questions for Energy Transfer

For the example questions pertaining to energy transfer, we first show the question in its raw form as suggested by the focus group, followed by its reformulation to a form that the system can compute, an abstract characterization of that form, and a sample answer.

Q1: How are anabolic pathways energetically related to catabolic pathways?
Reformulation: What is the energetic relationship between an anabolic pathway and a catabolic pathway?
Abstract form: What is the ⟨modifier⟩ relationship between ⟨X⟩ and ⟨Y⟩?
Answer: An endergonic process, which is a subclass of catabolic pathway, uses free energy resulting from an exergonic process, which is a subclass of an anabolic pathway.

Q2: The mitochondria in your muscle cells produce ATP energy through the citric acid cycle. What provides the energy for this process?
Reformulation: What processes provide raw materials for Citric Acid Cycle?
Abstract form: During ⟨X⟩, what processes provide the raw materials for ⟨Y⟩?
Answer: Oxidation: an electron is oxidized from a malate to an NAD+ resulting in an oxaloacetate. This process requires an activation energy. Additionally, a Redox reaction consumes a malate and an NAD+ and produces an NADH and an oxaloacetate.

Q3: In terms of energy, it has been said that bioluminescence is the opposite of photosynthesis. Why is this the case?
Reformulation: What is the energetic difference between bioluminescence and photosynthesis?
Abstract form: What is the ⟨modifier⟩ difference between ⟨X⟩ and ⟨Y⟩?
Answer: In photosynthesis, light energy is consumed and a chemical is produced. In bioluminescence, chemical energy is consumed and light energy is produced.

Q4: In terms of potential energy, what does the Calvin cycle accomplish?
Reformulation: Compare the potential energies of the raw materials of the Calvin Cycle and the results of the Calvin Cycle.
Abstract form: Compare the ⟨modifier⟩ energies of ⟨X⟩ and ⟨Y⟩.
Answer: The Calvin cycle raises the relatively low potential energy of carbon dioxide to the high potential energy of sugar molecules.

Q5: In glycolysis, a series of reactions during fermentation converts glucose to pyruvate, lowering the free potential energy of the molecules. What are those reactions, and in what order do they take place?
Reformulation: During Fermentation, what sequence of steps of Glycolysis converts Glucose to Pyruvate?
Abstract form: During ⟨X⟩, what sequence of steps of ⟨Y⟩ converts ⟨Z⟩ to ⟨U⟩?
Answer: The energy investment phase of glycolysis is followed by the energy payoff phase of glycolysis.

We reformulate a raw question form to a form used for computation to keep the focus on representing and reasoning with energy-related concepts. Attempting to answer the question stated in its raw natural language form would require dealing with the problems of arbitrary natural language understanding; that would diffuse the focus on the conceptual modeling of energy-related concepts that is the primary focus of our work. Question Q1 above queries for specific relationships between two individuals. The computation involved in answering this question has been previously explained in [6]. When we are specifically interested in energetic relationships, the search is restricted to relationships such as *raw-material*, *result*, *possesses*, etc. The ⟨modifier⟩ can take other values such as *structural* or *regulatory*. Question Q2 is a straightforward query about the *raw-material* of a process. Question Q3 asks for energetic differences between two concepts. The computation for the generic differences question has been previously explained in [8]. While computing energetic differences, the computation is restricted to energy-related relationships such as *raw-material*, *result*, etc. Q4 also asks for a difference, but here the entities of interest are involved in the same reaction. Q5 asks for processes that *convert* Glucose into Pyruvate. Here, the KB must contain a definition of convert; that is, a series of steps convert A to B, if A is an input to the first step, and B is a result of the last step.

4.3. Questions for Process Regulation

We consider here a few representative example questions pertaining to process regulation. The format for presentation is analogous to the one we used for energy transfer in the previous section.

Q6: What role is played by Thermoregulation in the life of Penguins?
Reformulation: What property is maintained by Thermoregulation?
Abstract form: What property is maintained by ⟨X⟩?
Answer: The temperature of an animal.

Q7: What is conventional set point for human body temperature?
Reformulation: Above what value does Thermoregulation in Human maintain temperature?
Abstract form: $\langle\{\text{Above}, \text{Below}\}\rangle$ what value does ⟨X⟩ maintain ⟨property⟩?
Answer: 36° Celsius.

Q8: Describe the regulation of the rate of glycolysis. (How is the rate of glycolysis controlled?)
Reformulation: What are the mechanisms of the Regulation of Glycolysis?
Abstract form: What are the mechanisms of ⟨X⟩?
Answer: A citrate inhibits a phosphofructokinase, an AMP activates a phosphofructokinase, and an ATP inhibits a phosphofructokinase.

Q9: List the effector organs for Thermoregulation and how they would respond to an increase in core body temperature.
Reformulation: How are the mechanisms of thermoregulation in human triggered?
Abstract form: How are the mechanisms of ⟨X⟩ triggered?
Answer: An increase in temperature of a person causes a hormone to be secreted at the hypothalamus which causes Vasodilation and secretion of sweat at the skin.

Q10: Phosphofructokinase (PFK) is allosterically regulated by ATP. Considering the result of glycolysis, is the allosteric regulation of PFK likely to increase or decrease the rate of activity for this enzyme.
Reformulation: How do the mechanisms of Regulation of Glycolysis affect the regulated property?
Abstract form: How do the mechanisms of ⟨X⟩ affect the regulated property?
Answer: The rate of glycolysis is decreased by a citrate inhibiting a phosphofructokinase, increased by an AMP activating a phosphofructokinase, and decreased by an ATP inhibiting a phosphofructokinase.

Just like the questions for energy transfer, the reformulation of questions from their raw form is aimed at factoring away the complexities of natural language. Here, Q6 asks for the regulated property, Q7 asks for the limits of the regulated property and Q8 asks for a description of how a process is regulated. In Q9 the query is for the stimuli of regulation, whereas in Q10 the query is for response mechanisms. We have not included examples of relationship or comparison questions, but such questions are equally applicable here.

5. Future Work and Conclusions

We have used the concepts introduced in this paper to model a range of energy transfer and regulation concepts across the textbook. For energy transfer, in addition to Light-Absorption, we have modeled Phosphorylation, Light-Reaction, and Redox-Reaction. For regulation, we have modeled concepts such as the regulation of Cell-Cycle and Mitosis. However, these concepts are not a part of the intelligent textbook yet which is open for future work.

A specific challenging application for these concepts is regulation of Cell-Cycle (Campbell chapter 12) because of multiple levels of abstraction. Most of the text is at a general level (e.g., start/stop signals), but some goes into more biological detail (regulation of the M-Phase checkpoint by fluctuating Cyclin concentration). The "checkpoints" do not directly affect the rate of the cell cycle. The regulatory process is not primarily in response to external stimuli. It is a cyclical process that continues all the time on its own accord. Thus, strictly speaking, it may not be a case of regulation according to the definition we have introduced here. However, some cases of external stimuli mentioned in Campbell do affect this process (e.g., release of growth hormone). Resolving these modeling choices is open for future work.

In summary, we considered the detailed representation of energy transfer and regulation of biological processes. We introduced the conceptual primitives necessary to model such information. We argued that with the goal of factoring the representation of complex processes into simpler modules, it is advantageous to view such information as additional perspectives on the basic view of a process which captures steps of the process and its participants. We showed how we could relate those multiple perspectives to each other by using semantic relationships as well as by the sharing of Skolem functions. We also considered several example questions that can be answered using such representations. Creating representations from a biology textbook has provided an extremely useful and powerful framework that has allowed us to pursue the ontology research described here, and we encourage others to follow similar methodologies, in which conceptual decisions can be rooted in objective criteria.

References

[1] Collin F Baker, Charles J Fillmore, and John B Lowe. The Berkeley FrameNet project. In *Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics-Volume 1*, pages 86–90, 1998.

[2] Chitta Baral. *Knowledge representation, reasoning and declarative problem solving*. Cambridge University Press, 2003.

[3] K. Barker, T. Copeck, S. Delisle, and S. Szpakowicz. Systematic construction of a versatile case systems. *Journal of Natural Language Engineering*, 3(4):279–315, 1997.

[4] K. Barker, B. Porter, and P. Clark. A library of generic concepts for composing knowledge bases. In *First International Conference on Knowledge Capture*, 2001.

[5] Vinay K Chaudhri, Britte Cheng, Adam Overholtzer, Jeremy Roschelle, Aaron Spaulding, Peter Clark, Mark Greaves, and Dave Gunning. *Inquire Biology*: A textbook that answers questions. *AI Magazine*, 34(3):55–72, 2013.

[6] Vinay K. Chaudhri, Nikhil Dinesh, and Craig Heller. Conceptual models of structure and function. In *Second Annual Conference on Advances in Cognitive Systems*, 2013.

[7] Vinay K. Chaudhri, Nikhil Dinesh, and Daniela Inclezan. Three lessons in creating a knowledge base to enable explanation, reasoning and dialog. In *Second Annual Conference on Advances in Cognitive Systems*, 2013.

[8] Vinay K. Chaudhri, Stijn Heymans, Adam Overholtzer, Aaron Spaulding, and Michael Wessel. Large scale analogical reasoning. Technical report, SRI International, 2014.

[9] College Board. Biology: Course description. `http://apcentral.collegeboard.com/apc/public/repository/ap-biology-course-description.pdf`, 2010.

[10] Antony Galton. States, processes and events, and the ontology of causal relations. In *Proceedings of the Seventh International Conference on Formal Ontology in Information Systems (FOIS)*, pages 279–292, 2012.

[11] Aldo Gangemi, Nicola Guarino, Claudio Masolo, and Alessandro Oltramari. Sweetening WordNet with DOLCE. *AI magazine*, 24(3):13, 2003.

[12] David Gunning, Vinay K. Chaudhri, Peter Clark, Ken Barker, Shaw-Yi Chaw, Mark Greaves, Benjamin Grosof, Alice Leung, David McDonald, Sunil Mishra, John Pacheco, Bruce Porter, Aaron Spaulding, Dan Tecuci, and Jing Tien. Project Halo update: Progress toward digital Aristotle. *AI Magazine*, 31(3):33–58, 2010.

[13] Douglas B Lenat. CYC: A Large Scale Investment in Knowledge Infrastructure. *Communications of the ACM*, 38(11):33–38, 1995.

[14] S. M. Lloyd, editor. *Roget's thesaurus*. Longman, 1982.

[15] Ian Niles and Adam Pease. Towards a standard upper ontology. In *Proceedings of the International Conference on Formal Ontology in Information Systems-Volume 2001*, pages 2–9. ACM, 2001.

[16] Jane B. Reece, Lisa A. Urry, Michael L. Cain, Steven A. Wasserman, Peter V. Minorsky, and Robert B. Jackson. *Campbell biology*. Benjamin Cummings imprint of Pearson, Boston, 2011.

[17] Karin Kipper Schuler. *VerbNet: A broad-coverage, comprehensive verb lexicons*. PhD thesis, University of Pennsylvania, 2005.

[18] D. Summers, editor. *Longman dictionary of contemporary English*. Longman, 1987.