NORTHWESTERN UNIVERSITY

Capturing Qualitative Science Knowledge with Multimodal Instructional Analogies

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Computer Science

By

Maria de los Angeles Chang

EVANSTON, ILLINOIS

September 2016

ABSTRACT

This thesis explores a communication method that is relevant to learning by reading systems and educational software systems: instructional analogy. It is widely recognized that analogical reasoning plays a vital role in our ability to detect similarities and differences and to transfer knowledge between topics. Building software that can understand these analogies can advance the state of the art in knowledge capture and is an important step toward human-level learning by reading. However, this goal involves several challenges. Instructional analogies are expressed in natural language and are often accompanied by rich spatial representations. Successful interpretation of these analogies requires substantial background knowledge about everyday experiences and the physical world. Lastly, the target knowledge for instructional analogies for introductory science is often conceptual and qualitative. The claims of this thesis are that a combination of simplified natural language understanding, sketch understanding, and structure-mapping can be used to build qualitative, conceptual knowledge from multimodal instructional analogies, and that such knowledge can be used to answer questions about the new domain. A model of instructional analogy interpretation was developed and tested on a set of 11 multimodal analogies. The model achieves 83% accuracy on queries of target knowledge. Ablation experiments indicate that the use of base domain knowledge, implicit background knowledge, and sketch knowledge make varying contributions toward accuracy. The knowledge acquired from these analogies was used to answer 11 out of 14 related questions on middle school science exams. Continued research in this area can improve reading systems and opens the door to new kinds of intelligent tutoring systems.

ACKNOWLEDGEMENTS

I thank the National Science Foundation and the Spatial Intelligence and Learning Center for their generous financial support of my research.

I thank my thesis committee, Dedre Gentner, Chris Riesbeck, and Matthew Easterday for their constructive feedback. I especially thank Dedre Gentner who supported me as if I was one of her own students.

I thank Ken Forbus, for being an outstanding advisor and for reminding me that everything is hard until it's easy. That battle is always worth it.

I thank Nadine Gaab, Hector Muñoz-Avila, and Yolanda Gil for their mentorship and constructive feedback.

I thank my fellow students, researchers, and staff at the Qualitative Reasoning Group and the Cognitive Systems Division, for their friendship, intellectual inspiration, and support: Subu Kandaswamy, Jon Wetzel, David Barbella, Jason Taylor, Andrew Lovett, Maddy Usher, Tom Hinrichs, Jenn Stedillie, and Carrie Ost. I thank the next generation of QRG students, Joe Blass, Chen Liang, CJ McFate, Irina Rabkina, and Max Crouse, for keeping it fresh. I especially thank Scott Friedman, Matt McLure, and Mark Cartwright for being great friends to me and, in varying ways, knowing just what to do to make my days better.

I thank my family and friends for their genuine interest in my work and well-being. My parents, Peppina Liano and Jorge Arturo Chang, have shown me unconditional support. To my siblings, Blanca, Todd, Aurora, Andres, George, and Jose, thank you for reminding me how crazy our family is and how that's a good thing. To my second family, Rick and Sue Porteus, Karissa, Mike, Jacob, and Ashley Reardon, I'm so honored to have all of you in my corner. Marco and Elliot Hidalgo-Greenberger, you know you two are what I'll miss most about Chicago. Thank you, Matt O'Connor, for all the political nonsense that you know I love. I thank my friends back home, Leah Rubin, Jeremiah Romm, and Diane Scheiner, for providing the friendship that I need even when we're far away. I thank my dogs, Julep and Marmalade, for always listening to my practice talks.

Most of all, I thank my wife, Shauna Porteus, for being the most incredibly loving and supportive partner I could have hoped for.

For my parents,

Peppina Liano & Jorge Arturo Chang

who truly embody strength, courage, and sacrifice.

For my wife,

Shauna Porteus

who loves and supports me through everything, over and over again.

Contents

List of Tables

List of Figures

Chapter 1: Introduction

This thesis explores a communication method that is relevant to learning by reading systems and educational software systems: instructional analogy. It is widely recognized that analogical reasoning plays a vital role in our ability to detect similarities and differences and to transfer knowledge between topics. Instructional analogies are very frequently used in textbooks and by educators to communicate new ideas to people, especially in science, technology, engineering, and mathematics (STEM) domains. What is the nature of these analogies? What are the reasoning requirements for interpreting them? Do they provide knowledge that is useful to an intelligent system? These are the main questions that this thesis addresses.

This thesis shows that (at least) the following capabilities are needed to interpret instructional analogies:

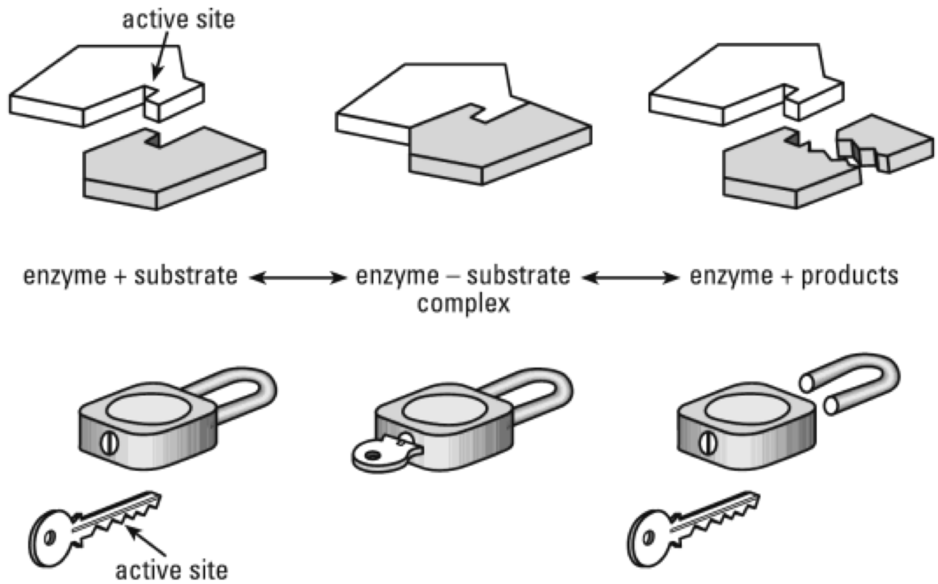
1) Multimodal reasoning, including natural language understanding, spatial reasoning, and the ability to combine information from both modalities
2) Use of background knowledge, including partial knowledge about everyday concepts or experiences
3) Analogical reasoning, including transferring knowledge from one topic to another

Without limits, these requirements can expand to AI-complete problems. However, the task constraints imposed by instructional analogies can be used to contextualize these requirements, and to develop heuristics that make interpretation feasible. This thesis demonstrates how these requirements, constrained by the reasoning goals of instructional analogies, can be fulfilled with a combination of simplified natural language understanding, sketch understanding, and analogical reasoning governed by the structure-mapping theory of analogy (Gentner, 2010).

1.1 Motivation

Analogical reasoning is a hallmark of interpersonal communication and learning. Analogies are powerful tools for communicating abstract ideas, highlighting important similarities and differences, and focusing inference. *Instructional analogies* are comparisons that are often used to teach introductory science topics (Glynn, 2008; Harrison & Coll, 2007; Zeitoun, 1984). The purpose of instructional analogies is to describe new, unfamiliar concepts in terms of familiar ones. They tend to use basic, everyday experiences or objects as a source of knowledge (often referred to as a *base domain*) that can be projected onto a new topic (often referred to as a *target domain*). For example, in biology, enzyme activation is described as a process that is similar to unlocking a lock. Instructional analogies are often presented explicitly, with individual similarities identified, e.g. *An enzyme molecule is like a key / A substrate molecule is like a lock,* Figure 1. Analogies may also be implicit, where the comparison is deeply embedded in the language used to describe the topic, e.g. *The electrical current **flows** through the wire*. In some cases, it can be very difficult to describe things without implicitly invoking an analogy or conceptual metaphor (Lakoff & Johnson, 2003).

Instructional analogies may also take the form of comparisons between highly similar things. Figure 2 shows a classic *bridging analogy* (Clement, 1993). The spatial arrangement of the two images is intended to convey an analogy between a hand pressing down on a spring and a book at rest on a table. It is also intended to suggest that, just as the spring pushes up on the hand, the table pushes up on the book. This analogy can be used to teach students about forces on objects in static equilibrium. Unlike analogies for enzyme action and electric current, this analogy involves two scenarios of the same domain. For this reason, it is referred to as a *within-domain* analogy, whereas analogies between two completely different topics are referred to as *cross-domain* analogies.



Lock and key	Enzymes and substrates
Key	Enzyme Molecule
Lock (padlock)	Substrate Molecule
Notched part of key (has unique shape)	Active site (has unique chemical makeup)
Keys unlock only specific locks	Enzymes react only with specific substrates
Key breaks apart (unlocks) a padlock	Enzyme action breaks apart substrate molecules.
Key comes out of the lock unchanged and can be reused.	The enzyme comes out of the reaction unchanged and can be reused.

Figure 1: An analogy used to explain how enzymes and substrates work from Harrison & Coll (2007), p. 96.

Spatial representations provide additional benefits to instructional analogies. As illustrated by Figure 2, the spatial arrangement of elements in a drawing or diagram can invite comparison between individual items that are laid out in a similar way. For STEM domains in particular, students often have to learn about concepts that they cannot observe directly. In geoscience, for example, students must understand processes that cannot be fully observed visually because they are mostly invisible (e.g. the carbon cycle), they occur over thousands of years (e.g. subduction), or they exist on very large scales (e.g. layers of the earth). Similar issues are encountered in biology (e.g. microbiology) and physics (e.g. electricity). Spatial representations provide a way for students to visually observe things that they would not be able to see otherwise.

Instructional analogies also tend to focus on *qualitative knowledge*, which is important for solving problems with partial information and for representing continuous phenomena with discrete labels. Qualitative knowledge allows us to understand scenarios with little or no detailed quantitative information. The connection between analogical reasoning and the development of abstract knowledge is important because abstract, qualitative knowledge is a hallmark of expert knowledge (Chi et al., 1981). This has impacted teaching strategies and assessments. For example, the *force concept inventory* is a test of qualitative physics knowledge that was developed after discovering that quantitative assessments were poor measures of physics expertise (Hestenes & Halloun, 1995; Hestenes et al., 1992). Analogies are useful for building qualitative knowledge because they enable people to use a familiar vocabulary to partially describe new topics and, as described next, they help people focus on abstract similarities and differences.

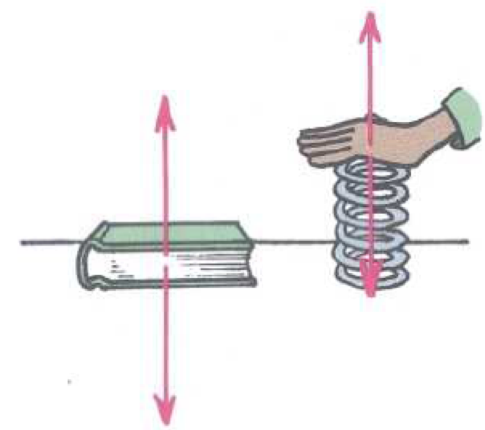


Figure 2: An analogy between a book at rest on a table and a hand pushing down on a spring, *Conceptual Physics*, Hewitt (p. 28).

1.1.1. Structure Mapping Theory

Structure-mapping theory (SMT) (Gentner, 1983) provides an account of how analogical reasoning works and why it is useful. According to SMT, there is an important distinction between surface similarity and structural similarity. *Surface similarity* refers to shared attributes between two things, e.g. color, overall shape. *Structural similarity* (also called relational similarity) refers to shared relational structure, e.g. that two objects cause something else to happen. The distinction between relational similarity and surface similarity is important because SMT claims that relational similarity is the type of similarity that has the greatest inferential power. The presence of shared systems of relations is referred to as *systematicity*, and the importance of shared systems of relations for detecting structural alignment is referred to as the *systematicity principle*. One can see how the distinction between surface and structural similarity maps to instructional analogies described earlier: within-domain analogies have high surface similarity (and possibly, but not necessarily, high structural similarity), and cross-domain analogies (if they are effective instructionally) have high structural similarity (and possibly, but not necessarily, some surface similarity).

Psychological evidence supports the claim that structural alignment drives what analogical inferences people make. In a seminal study on problem solving, Gick and Holyoak (1980) demonstrated that cross-domain analogies between stories could be used for problem solving. They demonstrated not only that people can use stories from distant domains to solve new problems, but that the level of structural similarity impacts the likelihood of arriving at an analogous solution. In a similar study, Blanchette and Dunbar (2002) showed that people have a selective bias for inferences that have high systematicity, i.e. are supported by high structural similarity. The structural similarity that supports analogical inference can be spatial as well (Gentner et al., 2015). In a study at the Children's Museum in Chicago, children were shown pairs of towers that varied in terms of their surface and structural similarity to each other. In this case, structural similarity didn't involve semantic relations in a story, but spatial relations between parts of the towers. Children were told to compare the two towers to each other and were asked, "which one is stronger?" (i.e. more stable?). In each pair, only one tower used a diagonal brace in its construction, which made it more stable than the tower that lacked a diagonal brace. The pairs were either highly alignable, meaning that their parts shared many spatial relationships, or less alignable, meaning that their parts shared fewer spatial relationships. Children between the ages 6 and 8 who were shown the highly alignable pairs were more likely to apply the diagonal brace principle to a tower repair task than children who were in the low alignment condition. This indicates that the important difference between highly alignable towers was made more salient by the presence of many shared spatial relations. Other studies on SMT illustrate the importance of systematicity (Clement & Gentner, 1991), the impact of relational language on category learning (Gentner & Namy, 1999), and the *career of similarity* whereby the ability to attend to abstract relational similarities increases with age and expertise (Gentner & Ratterman, 1991).

Structure-mapping theory also suggests that the alignment of relational structures enables important similarities to emerge, while abstracting away irrelevant details (Kotovsky & Gentner, 1996; Namy & Gentner, 2002). Analogical reasoning commonly refers to transferring knowledge from a known topic to an unknown topic, but researchers have shown that even when presented with two scenarios that are both only partially understood, mutual alignment can promote deep understanding (Kurtz et al., 2001; Mason, 2004). This leads to the formation of abstract knowledge, or schemas, which can readily be transferred to novel situations. Gick and Holyoak (1983) found that abstract schemas for solving particular story problems can be induced by comparing two analogous stories, and that the act of comparing analogous stories is more effective than studying and summarizing one analogous story before being reminded of it at transfer time. Lowenstein et al. (1999) conducted a transfer learning experiment with business school students and found that comparing negotiation scenarios was more effective for transfer than reading identical scenarios independently. Gadgil, Nokes-Malach and Chi (2012) found that comparison of different mental models of the circulatory system was more effective than self-explanation at undoing misconceptions. In another investigation, analogical comparison and self-explanation were used to promote far transfer in physics problem solving (Nokes-Malach et al., 2013), and both were independently more effective than simply reading from worked examples. Additionally, the quality of the schemas produced by participants predicts how often they are able to transfer that schema to new examples (Gentner et al., 2003; Gick & Holyoak, 1983; Loewenstein et al., 1999). These studies indicate that analogical comparisons facilitate the acquisition of abstract knowledge. Given the importance of qualitative knowledge in early science education, it makes sense that analogies are appealing instructional tools.

Despite evidence of their efficacy, there are circumstances that can cause analogies to be misleading. Spiro et al. (1989) developed a taxonomy of eight types of analogy pitfalls, where there is a tendency to oversimplify complex ideas based on individual analogies. These pitfalls include various ways of over-projecting information from the base domain (i.e. the known topic) to the target domain (i.e. the new topic) or missing important parts of the target domain because there are no corresponding parts in the base domain. These pitfalls can be thought of as problems with the alignment process or as inadequate evaluation of candidate inferences. There are also two pitfalls that relate to the specific terminology that is used in analogies. Some terms have common-knowledge connotations that conflict with their technical meanings. Spiro et al. provide an example from biology, where the word “compliance” refers to the flexibility of blood vessels, and students develop misconceptions about blood vessels “giving way to” or “surrendering to” blood. This type of pitfall is due to the recruitment of erroneous or irrelevant background knowledge that makes its way into the analogical mapping. While the taxonomy identified by Spiro and colleagues describes the various misconceptions that can arise from analogies, structure-mapping can provide additional insight into why these misconceptions take place. The systematicity principle indicates that common relational structures are critical to making structural alignments. It is therefore critical to make common relations known and block potential mismatches. Our tendency for noticing surface similarity also plays a role in these pitfalls, since we process local surface matches before (or faster than) relational ones (Goldstone & Medin, 1994). If two items in a mapping are intended to correspond, but a competing match shares greater surface similarity, it is more difficult to arrive at the intended mapping. Such analogies are called *cross-mapped* comparisons because common surface attributes invite an unintended correspondence (Gentner & Toupin, 1986). Cross-mapped comparisons are more difficult for novices (Gentner & Ratterman, 1991), indicating that, especially when teaching a new topic, surface similarity should be as consistent as possible with relational matches. Other factors that impede the ability to detect relational matches include excessive cognitive load (Markman & Gentner, 1993) and unfamiliarity with the base domain.

These pitfalls and recommended solutions are consistent with the teaching strategies identified by the education literature, which are described next.

There are three prominent guides for teaching with analogies: the General Model of Analogy Teaching (GMAT) (Zeitoun, 1984), the Teaching with Analogies (TWA) guide (Glynn, 2008), and the Focus Action Reflection (FAR) guide (Harrison & Coll, 2007). The general principles for all three guides are very similar and consistent. They all stress the importance of assessing learners' knowledge prior to instruction as well as explicitly identifying similarities and differences between the topics in the analogy. To use Glynn's terms, *simple analogies* only give the student the base and the target, e.g. "An animal cell is like a city," whereas *elaborate analogies* systematically map properties of one description to properties of the other. Elaborate analogies are preferred. It is important to identify differences as well because learners need to be made aware of how the analogy breaks down. Without knowing the differences, the analogy may be taken too far, leading to incorrect inferences or misconceptions. The use of multiple analogies is also recommended to refine knowledge and fill in gaps that other analogies leave open. These are important parts of the teaching process because they address some of the pitfalls identified by Spiro et al. (1989). Holyoak and Richland (2014) take these practices even further, recommending (1) multiple bases/sources for analogies presented in order of difficulty to facilitate progressive alignment (Gentner & Medina, 1998), (2) techniques for reducing working memory load (e.g. using visual representations), and (3) the use visual representations and gestures that highlight relational commonalities, making them highly alignable. These recommendations are consistent with the literature supporting the use of progressive alignment, relational language, and alignability to promote learning (Gentner et al., 2015; Gentner & Ratterman, 1991; Kotovsky & Gentner, 1996). In general, these separate recommendations are complementary. The FAR Guide is a particularly useful resource because it contains a collection of instructional analogies for many different science domains. For this reason, it is used as a source of material for this thesis.

The education and cognitive science literature indicates that multimodal instructional analogies play important roles for positive learning outcomes in people (Alfieri et al., 2013). Interactive AI systems should therefore know how to interpret these analogies so that they can communicate more naturally with people and so that they can capture knowledge in the same ways that people do. However, current AI systems lack the ability to understand or make use of multimodal instructional analogies.

1.1.2. Impact in AI

Using analogy to transfer knowledge between domains, as people do, is a longstanding problem in AI with many applications. Researchers believed early on that the ability to evaluate a new scenario with respect to some prior experience could be very useful in planning (Veloso & Carbonell, 1993), learning (Burstein, 1985), and problem solving (Falkenhainer, 1990; Klenk & Forbus, 2013). In addition to being able to use analogies to transfer knowledge, intelligent systems need to know how to interpret analogies for human-level learning by reading (Barbella & Forbus, 2011). As we move toward the goal of building interactive intelligent systems, analogical reasoning becomes an important requirement for learning from human collaborators, who may use novel analogies to explain complex phenomena. Similarly, interactive intelligent systems need the ability to communicate via analogy to engage in rich interactions with their human collaborators or to explain new concepts to them. Given that analogies are such powerful tools for learning, building analogy-based tutoring systems has been proposed (Lulis et al., 2004; Murray et al., 1990), but of the three that have been implemented, all are intended to operate in a single domain and on a restricted topic (i.e. support forces (Murray et al., 1990), how the circulatory system works (Lulis et al., 2004), and topics in data structures (Harsley et al., 2016)).

As with many forms of knowledge capture, reasoning about analogies within one modality is not enough. Multimodal instructional analogies are pervasive in textbooks and classrooms. Analogies in textbooks are often accompanied by pictures or diagrams, and the use of spatial representations is part of best practices for teaching with analogies, as mentioned above (Holyoak & Richland, 2014). The importance of spatial representations is not surprising because they lighten working memory load (Larkin & Simon, 1987) and support students' spatial reasoning, which has been shown to be a strong predictor for success in many STEM domains (Wai et al., 2009). For analogies in particular, spatial representations can facilitate alignment, making analogies like the one shown in Figure 2 very easy for people to align. It is therefore important for an intelligent system to be able to take advantage of the spatial information that is depicted with instructional analogies.

Intelligent systems that can combine analogical reasoning, qualitative reasoning, and spatial reasoning, can advance the state of the art in knowledge capture. Understanding the requirements for reasoning about multimodal instructional analogies is an important step toward this goal. The claims of this thesis are that a combination of simplified natural language understanding, spatial reasoning, and structure-mapping can be used to build qualitative, conceptual knowledge from multimodal instructional analogies, and that such knowledge can be used to answer questions about the new domain. We examine those claims in detail next.

1.2 Claims and Contributions

Claim 1: Structure-mapping can be used to build qualitative knowledge from multimodal instructional analogies.

This claim is supported by experiments in Chapter 4. It can be decomposed into three sub claims:

1) Multimodal instructional analogies use visual representations to facilitate interpretation.

2) Instructional analogies use background knowledge to facilitate interpretation.
3) Multimodal integration and analogy interpretation can be achieved using structure-mapping.

A qualitative analysis of common analogies used to teach introductory science provides evidence that visual representations are pervasive and that accurate interpretation relies on implicit background knowledge. Based on this analysis, I implemented an analogy interpretation system that takes as input simplified natural language passages and digital sketches. This system uses spatial and conceptual information depicted in each sketch and expressed in natural language to build and interpret instructional analogies. The resulting knowledge is qualitative because it includes causal relationships between continuous quantities and type-level rules that may not be universally true, but are nonetheless useful for reasoning. Structure-mapping serves two main purposes in this system. It is the process by which visual and text-based representations are aligned and integrated. It is also the process by which the instructional analogy is interpreted. The interpretation allows the system to construct facts about the new topic that use ontologies from a large scale knowledge base. Variations in the reasoning components of the system demonstrate the relative contribution of:

1) Spatial Information / Multimodal integration
2) Background knowledge
3) Explicit base domain knowledge

Contribution 1: Interpretation requirements and strategies for multimodal instructional analogies.

As a result of the experiments used to support claim 1, a set of interpretation strategies are identified for multimodal instructional analogy interpretation. These strategies are:

1) The use of visual representations as a natural language disambiguation heuristic

2) The use of visual-conceptual elaboration for interpreting sketches
3) The use of structure-mapping for combining visual and sketch-based representations
4) The use of dialogue acts and event interpretations for determining cross-domain match constraints
5) Type-level summarization of facts prior to analogical mapping

Claim 2: Qualitative knowledge captured via multimodal instructional analogies can be used to answer questions.

This claim is supported by experiment in Chapter 5, which demonstrates that the knowledge captured from a set of multimodal instructional analogies, although qualitative and incomplete, can be used to answer questions from middle school science exams. It is not the case that the knowledge captured from multimodal instructional analogies is sufficient for answering all (or most) of the exam questions. The experiments do suggest, however, that the qualitative knowledge plus some simple question answering heuristics are useful for answering questions about basic functions and properties of concepts in the tested domain.

Contribution 2: Question answering strategies using knowledge captured via analogy.

Question answering is carried out with the following basic heuristics:

1) Detecting object properties via part-whole reasoning and base domain reasoning
2) Reasoning about example situations with qualitative background knowledge
3) Using concept and relation hierarchies to estimate semantic similarity of statements

1.3 Preview

Chapter 2 covers the theoretical background for this work and explains the systems used in this thesis. Chapter 3 provides a qualitative analysis of the analogies in the FAR guide. Chapter 4 describes my approach for interpreting multimodal instructional analogies and Chapter 5 describes my approach for answering middle school science questions using the knowledge captured in Chatper 4. Chapter 6 covers related work in analogical reasoning, multimodal reasoning, and question answering. I close with a general discussion and suggestions for future work.

Chapter 2: Background

2.1 Structure-Mapping Engine

The structure mapping engine is a computational model of analogy and similarity based on the structure-mapping theory of analogy. The algorithm has been described in detail elsewhere (Falkenhainer et al., 1989; Forbus et al., 1994), but here I describe the general inputs and outputs as well as how constraints and analogy control predicates are used in this task.

The structure-mapping engine (SME) takes as input two structured descriptions: a *base* and a *target*. The structured descriptions consist of expressions that describe the domain. Given a base and a target, SME produces one or more mappings between them. Each mapping has a set of correspondences, which say how things in the base correspond to things in the target, a structural evaluation score, which is an estimate of match quality (i.e. systematicity), and (optionally) candidate inferences, which are things that are true in the base and hypothesized to be true in the target. There can also be reverse candidate inferences, which are things that are true in the target and hypothesized to be true in the base. Essentially, candidate inferences (regular or reverse) represent differences between the base and the target. They have the potential to introduce new terms via *analogy skolems*. Analogy skolems are terms that exist in one description, but not the other.

SME provides optional match constraints that operate as advice to encode task demands. *Partition constraints* require that items of the same type correspond to each other (e.g. apples to apples and oranges to oranges). These are useful for literal similarity matches where types or surface attributes matter. *Required correspondences* are constraints on individual items (e.g. apple1 must correspond to apple2). These are useful in situations where correspondences are given explicitly, as is often the case in instructional analogies. Partition constraints and required correspondences are used the interpretation model described in Chapter 4.

SME also uses a vocabulary of *analogy control predicates*, which allow for certain predicates to be ignored or for certain predicates to be declared as *non-atomic functions.* A non-atomic function is a predicate that results in a *non-atomic term,* which is statement that should be considered a single term. For example, if there is a predicate, `LiquidFn`, that takes one argument and denotes the liquid form of that argument, we might want to treat facts like, `(LiquidFn Nitrogen)`, as being unique terms. Declaring `LiquidFn` a non-atomic function makes that happen.

2.2 Cyc

ResearchCyc (Cyc) is a large scale knowledge base (KB) with a representational language called CycL.[1] The Cyc KB has an ontology that describes concept and relational hierarchies that can be useful for reasoning. In Cyc, concepts are represented using *Collections,* which are like sets. Entities are instances or members of collections. For example, if I have a piece of fruit in my hand, that piece of fruit is an

[1] www.cyc.com

entity that belongs to the collection `Fruit`. *Microtheories* are logical environments that can be used to control reasoning. For example, to query the knowledge about a fictional world in Cyc, one must restrict the query to only examine facts in the microtheory that represents that fictional world. Predicates are used to form statements and they can be instance-level or type-level predicates. Instance-level predicates operate on instances, while type-level predicates can operate on collections and predicates.

Rule macro predicate examples	Meaning
`(relationAllExists cityMayor City Person)`	All cities have a mayor.
`(relationAllExistsRange citizens City Person Many-Quant)`	All cities have many citizens.
`(relationAllExistsCount physicalParts Person Eye 2)`	All people have 2 eyes.
`(relationAllInstance citizens Person PlanetEarth)`	All people are citizens of planet Earth.
`(relationExistsAll doneBy (ControllingFn City) CityGovernment)`	All city governments have a city that they control.
`(relationExistsRangeAll anatomicalParts Mammal Foot-AnimalBodyPart AFew-Quant)`	All mammals have a few feet.
`(relationExistsCountAll birthChild BirthEvent Person 1)`	Every person is born once.
`(relationExistsInstance temporallyCoexist Martian PlanetMars)`	Loosely: There are Martians on Mars. Precisely: There exists a Martian that temporally coexists with the planet Mars.
`(relationInstanceAll temporallyCoexist PlanetMars Martian)`	All Martians are on planet Mars.
`(relationInstanceExists bordersOn NileRiver Country)`	There are countries that share a border with the Nile river.
`(relationInstanceExistsRange citizens Tokyo-PrefectureJapan Person Millions-Quant)`	There are millions of Tokyo citizens.
`(relationInstanceExistsCount presidentOfCountry UnitedStatesOfAmerica Person 1)`	There is one president of the US.
`(relationAll repeatedEvent Sunrise)`	Sunrises repeat.
`(relationExists inProgressEvent Writing)`	There is writing happening right now.

Table 1: Examples of statements that use rule macro predicates in CycL and their meanings in English.

CycL has a class of type-level predicates that expand into rules. These predicates are called *rule macro predicates* (Table 1). They are useful for expressing general rules that apply to an entire collection. Since introductory science knowledge often contains information about classes of things or processes, they are an important part of the representational system used for this thesis.

Introductory science knowledge also includes information about functions and behaviors. These are typically represented in Cyc as events or situations (i.e. instances of the collection `Event` and/or instances of the collection `Situation`). When an action or event is described in Cyc, it is described using a *neo-Davidsonian* representation, which means that the event itself is an entity and properties of that event are related to the event separately. For example, the sentence "Jane kicked the ball" could be described using a three-argument predicate for the verb "kick."

```
(kicked Jane the-ball)
```

However, the neo-Davidsonian representation would look like this:

```
(isa kick123 KickingAnObject)

(doneBy kick123 Jane)

(objectActedOn kick123 the-ball)
```

This way of representing the event is more flexible. If more information were added, all that is needed is one more statement, rather than a new kick predicate with another argument slot. For example, "Jane kicked the ball to Bob" could be represented with the previous three statements and the following:

```
(toAgent kick123 Bob)
```

Neo-Davidsonian representations are used in the model for interpreting instructional analogies.

Functional predicates are predicates that denote other things dynamically. *Collection denoting functions* are functional predicates that denote collections do not already exist in the knowledge base or that need to be defined compositionally. For example, the functional predicate `SubcollectionOfWithRelationToTypeFn` can be used to denote subcollections with particular relation to other collections. To denote the collection of events where a ball is kicked, we can use this:

```
(SubcollectionOfWithRelationToTypeFn KickingAnObject objectActedOn Ball)
```

This term denotes the collection of events that are instances of `KickingAnObject` and are related to an instance of `Ball` via `objectActedOn`. To say that a particular event, `kick123`, belongs to this collection, we could use this statement:

```
(isa kick123
       (SubcollectionOfWithRelationToTypeFn KickingAnObject objectActedOn Ball))
```

While neo-Davidsonian representations are useful for flexibly representing individual events, collection denoting functions are useful for precisely describing collections of events. Paired with rule macro predicates, collection denoting functions can be useful for building rules. For example, to say that all children kick balls, we could use this statement:

```
(relationExistsAll performedBy
       (SubcollectionOfWithRelationToTypeFn KickingAnObject objectActedOn Ball)
       HumanChild)
```

Collection denoting functions and rule macro predicates are both used in this system to represent type level knowledge.

2.3 Qualitative Reasoning

Qualitative Process Theory (QPT) (Forbus, 1984) is a representational system for carving up continuous quantities and processes into qualitative representations. Continuous quantities include things like age, height, weight, account balance, calories, etc. Each of these quantities can have a numerical value associated with them. However, we often reason about quantities without knowing the exact numerical values by using terms like "short" or "low-calorie." In Cyc, such quantities are instances of the collection `ScalarInterval`, which is the collection of things that can be ranked on some scale. This includes quantities like temperature, which can have numerical values, and quantities like happiness, which don't have numerical values.

In QPT, processes cause changes in quantities. Processes can be represented using *model fragments* (Friedman & Forbus, 2011; Klenk & Forbus, 2009), which represent partial information about a process type. Each model fragment has participants (i.e. the entities involved), constraints on what can be involved, conditions for the process to actually happen, and consequences of the process. Model fragments can also represent conceptual or physical views that provide a perspective on a situation. Such model fragments are not processes, so while they may involve causal relationships, they do not directly cause quantities to increase or decrease. The consequences of model fragments typically include causal relationships, ordinal relationships, or correspondences between quantities.

Causality is embedded in relationships called *influences*. *Direct influences* describe changes to extensive parameters, e.g. "A growth processes directly influences weight." Using QPT syntax, this statement would be `(I+ (WeightFn dog) (RateFn dog-growing))`, where the token `dog-growing` would refer to the process of the dog growing. The rate of the growth process is a causal antecedent of the weight of the dog. A looser form of causality can be represented with *indirect influences*, which can describe changes to extensive or intensive parameters that are not necessarily additive. Indirect influences are represented by predicates for positive and negative *qualitative proportionalities,* denoted by the predicates qprop and qprop-. For example, in the winter, the higher the temperature, the happier I am:

```
(qprop (HappinessFn me) (TemperatureFn Chicago))
```

The fact that this sentence is qualified with "in the winter" indicates that this qualitative proportionality could be a consequence to a model fragment describing winter, meaning that when that model fragment is active, the relationship holds. In the summer, the relationship might change to:

```
(qprop- (HappinessFn me) (TemperatureFn Chicago))
```

This negative qualitative proportionality indicates that I am happier when it is cooler. As is the case with direct influences, the second argument of an indirect influence is the cause and the first argument is the effect. *Correspondences* represent points of equality between quantities. Consider the following expression:

```
(qpCorrespondence (PopulationFn GrantPark-Chicago) Zero
                  (TemperatureFn Chicago) Zero))
```

This correspondence means that when the temperature in Chicago is zero, the number of people at Grant Park (which is outside) is zero. Causal influences enable a qualitative analysis of a situation in terms of what quantities are changing. Given a scenario with at least one active process, the scenario can be characterized in terms of what quantities are increasing, decreasing, and staying the same. This characterization is called influence resolution. Qualitative simulation also enables limit analysis, which involves the prediction of future states, although this aspect of QPT is not used for this thesis. QPT is useful for supporting inference with partial information, which is exactly the type of information that exists when a topic is first being learned.

2.4 EANLU

The Explanation Agent Natural Language Understanding system (EANLU) (Tomai & Forbus, 2009) produces semantic representations for simplified natural language text. EANLU takes a pragmatic approach to natural language understanding. Semantic interpretation of text is made tractable by using sentences with simplified syntax. In other words, the goal is not to have complete coverage of natural language inputs, but rather to have very broad coverage of the knowledge that can be expressed to the system using simple sentences. For each sentence EANLU processes, it generates *choice sets* to represent the possible interpretation choices that can be made. Several heuristics for making semantic interpretation choices are built in to EANLU. For this thesis, semantic choices are selected using two simple heuristics: information gain and favored context. These heuristics and the way EANLU is used to interpret multimodal instructional analogies are described in section 4.2.3.

2.5 CogSketch

CogSketch is a domain-independent sketch understanding system (Forbus et al., 2011). *Sketch understanding* refers to the process of using spatial reasoning to interpret sketches, independent of sketch recognition. Sketch understanding is intended to cover the types of sketches that are qualitative, abstract, and may not necessarily reflect realistic images. CogSketch achieves this by using qualitative spatial reasoning over ink that is manually segmented and labeled by the user. In CogSketch, the basic building blocks of a sketch are called *glyphs*. To draw a glyph, the user draws ink using a mouse or stylus and tells the software when the glyph is done by clicking a *Finish Glyph* button. This means that the user manually segments ink into meaningful objects. Ink editing tools allow the user to merge and re-segment ink as they draw. Grouping tools also allow the user to group conceptual items together (e.g. two wheels and a frame can be grouped together to represent a bicycle). The user can also provide conceptual labels for glyphs using collections and relations from the Cyc knowledge base so that the software has a model of the user's intent that is tied to the Cyc ontology. By labeling an individual glyph with a collection from Cyc, the user is indicating that they have drawn an instance of that collection. For example, in Figure 3, the innermost circle represents a cell's nucleus. To convey this to CogSketch, the user would draw the nucleus, click *Finish Glyph* to indicate that the glyph is done, and label it with the collection `CellNucleus` from the Cyc knowledge base. Additionally, the parts of the cell in Figure 3 can be grouped together by the user. By labeling the grouped glyphs with the Cyc collection `Cell`, the user indicates that, when grouped together, the glyphs represent an individual cell. Although this approach requires extra drawing and labeling effort, it has two advantages over recognition of raw ink. The first is that it avoids segmentation and recognition errors because the user explicitly tells the software how to group ink and what they want the ink to represent. The second is that this draw-and-label interface is amenable to educational settings because students are required to label sketches and explicitly provide their (possibly incorrect) interpretation of what they have drawn. Also, ink recognition without labeling would not work across multiple domains, since the mapping from shapes to concepts is many to many. It is especially problematic when a new domain is being introduced, since training recognizers requires many examples.

CogSketch automatically generates qualitative spatial representations for what is drawn in a sketch. Topological relations (e.g. intersection, containment) and positional relations (e.g. above, right of) are automatically computed between adjacent glyphs. Spatial relations between nonadjacent glyphs can be computed on-demand. The conceptual labels provided by the user are also used by CogSketch so that spatial and conceptual information exist in the same reasoning environment (i.e. in the same microtheory). CogSketch sketches can (optionally) consist of multiple subsketches, which are useful for representing individual states in a sketch. Each subsketch has its own microtheory so that things about the same object can be true in one subsketch/state and not true in another.

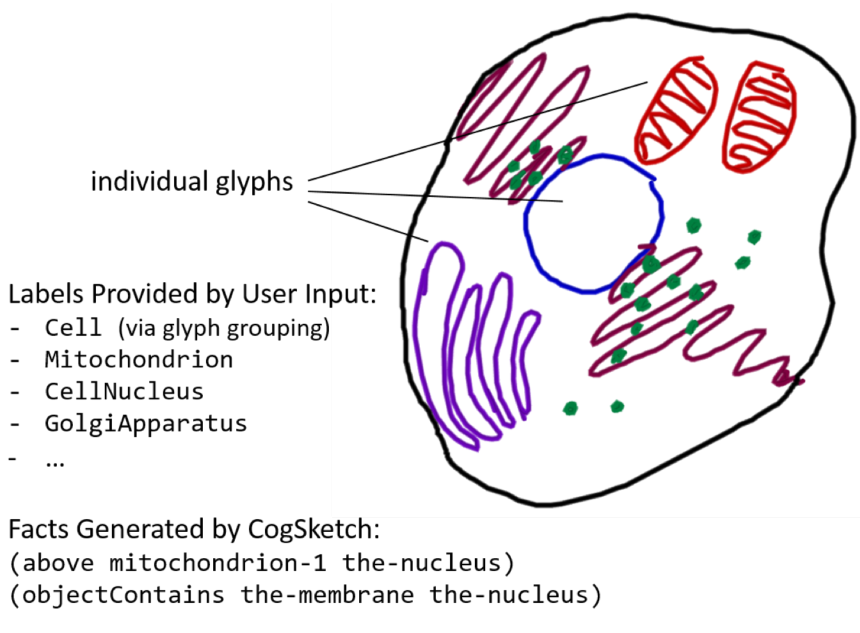


Figure 3: An example sketch with partial representations from CogSketch. Colors indicate separate glyphs drawn by the user. User-provided labels from the Cyc knowledge base give explicit conceptual information about the glyphs. CogSketch does not perform sketch recognition of the ink. Instead, it relies on the labels provided by the user. In this example, the different colored glyphs are all grouped together and labeled again, allowing CogSketch to represent each individual part with its own label (e.g. CellNucleus), as well as the group of glyphs with another label (i.e. Cell).

For interpreting the visual portion of instructional analogies, I created sketches using CogSketch. Each sketch was manually organized into individual glyphs, and labeled with concepts from the Cyc knowledge base. When objects were depicted as individual parts of another object, glyph grouping was used to represent part-whole relationships. In other words, CogSketch did not perform any sketch recognition. Rather, it generated qualitative spatial representations for conceptually labeled glyphs.

2.6 Companions

The Companion Cognitive Architecture (Forbus et al., 2009) is based on the idea that intelligent systems are social organisms that collaborate with others and learn over extended periods of time (e.g. from experience). In a Companion, analogy is a central reasoning mechanism. Each Companion is capable of using multiple modes of interaction. It has a natural language interface that is built upon EANLU and a sketching interface that uses CogSketch. I use the Companion cognitive architecture to model the interpretation of multimodal instructional analogies and to answer multiple choice questions. Instructional analogy interpretation is coordinated by HTN plans. Question answering strategies are represented by a cost-based AND/OR tree, where AND nodes represent subgoals for solving a question (i.e. they must all succeed) and OR nodes represent different strategies for solving a single question (i.e. only one must succeed). There are costs associated with OR nodes, so that some strategies can be prioritized over others.

Chapter 3: Characterization of Instructional Analogies

In general, the goal of instructional analogies is to build knowledge about a new topic. However, this is achieved in many different ways. Some analogies highlight functional similarities to help novices understand complex phenomena. Other analogies simply provide a physical model for something that is not directly observable. In this chapter, I provide a qualitative analysis of the reasoning requirements of instructional analogies found in a guide for in-service teachers.

3.1 The FAR Guide

The FAR guide (Harrison & Coll, 2007) contains analogies used for teaching middle school and secondary school science, collected from the educational literature, with recommendations for how to use them effectively in the classroom. The analogies cover several topics over four domains: Biology (12), Chemistry (13), Earth and Space Science (10), and Physics (14).

Each analogy uses a common sense domain, topic, or physical model to describe a science topic. The basic goal is to get the learner to transfer their knowledge about everyday situations to their knowledge about science. However, the analogies vary in terms of what similarities should be attended to and what types of inferences should be made. They also vary in what kind of spatial representations are associated with them. Although similarities and differences are often explicitly stated in each analogy, there is often background knowledge that the learner is expected to use. The following sections provide a qualitative analysis of these properties, which provides insight into what the reasoning requirements are for interpretation.

3.2 Focused Alignment and Inference

As others have noted (Spiro et al., 1989), analogies can sometimes lead to inaccurate conclusions. The authors of the FAR guide (and other guides in the education literature) agree with this caveat, and therefore recommend that teachers explicitly identify individual similarities and differences. This is especially important for cross-domain analogies, where there are many irrelevant similarities and differences that can distract students or lead them to incorrect conclusions. As a result, the reasoning goals and target knowledge of an analogy determines what similarities and differences are described.

Table 2 shows the number of analogies in the FAR that belong to different categories of reasoning goals. Out of the 49 analogies in the FAR guide, each generally fit into one, or more of these categories:

1) Analogies about functions, e.g. providing energy, building something

These analogies describe functions or behaviors in terms of known phenomena. For example, in the analogy *a cell is like a city,* parts of the city are compared to parts of the cell because they have similar functions. Power stations are like mitochondria because they both provide energy. Construction companies are like ribosomes because they both build things: homes and proteins, respectively.

2) Analogies about physical properties, e.g. containment, relative size

These analogies describe relative size, shape, or physical configurations. For example, in the analogy *a cell is like planet Earth*, parts of the planet are compared to parts of the cell to illustrate relative size and proper part relations. A cell is already small, but relative to chromosomes, the cell is very large. Similarly, the planet is much larger than an individual town or city.

3) Physical model analogies

These types of analogies are more common in domains that aim to explain processes or things at a level of detail that cannot be directly observed. For example, analogies about electricity use physical models to express the functions of the parts of a simple series circuit and the causal relationships between quantities (e.g. voltage and current flow). It is not the case that electricity cannot be observed at all, since one could build a circuit and observe its behavior. However, the explanation for *how* that artifact works is initially dependent on the physical properties of a different process where individual behaviors are more directly observable. Students do not *see* electricity *flowing*, but they have seen water flow.

Topic	Functional Similarities	Physical Similarities	Physical Model	Total
Biology	8	3	2	12
Chemistry	7	8	4	13
Earth/Space Science	8	8	6	10
Physics	14	2	9	14
Total	37	21	21	49

Table 2: Number of analogies in the FAR guide that involve reasoning about functional similarities, physical similarities, and physical models. Note that the three categories are not mutually exclusive.

Most of the instructional analogies in the FAR guide involve similarities in functions or behaviors. This is not surprising because cross-domain analogies usually lack surface similarity. In biology, for example, protein synthesis is described as following a master plan to build a house. The focus of this analogy is not on surface characteristics. Instead, the analogy focuses on the roles that DNA, the cell nucleus, and amino acids play in building proteins. Students familiar with the notion of blueprints and master plans can use that knowledge to interpret DNA as *instructions* for building proteins. Another example is in physics, where conservation of electric current in a simple series circuit is explained using an analogy to a passenger train. In a simple series circuit, with a battery and a light bulb, the conversion of energy by the battery and light bulb are represented by passengers entering and exiting the train. The train continues to move so long as it is in a closed track, which mimics current in a closed circuit. In this analogy, it's not the physical configuration that matters, but rather the requirements for movement (i.e. flow) in a circuit.

Other analogies focus on similarities in relative size, structure, and/or shape. These analogies are used to deal with the difficulties in observing quantities at very large or very small scales. For example, to help students understand Avogadro's number, the FAR guide describes an analogy to grains of rice that fill the United States or oranges that fill the planet. This is intended to help students understand how large Avogadro's number is, and to be able to conceptualize a mole as a single unit of something. Similarly, to understand the length of the human genome, instructors may also use an analogy to a very long road trip. In this analogy, regions along the long trip (e.g. states) represent chromosomes. Monotonous parts of the trip (e.g. long stretches of empty land) represent repetitive DNA that does not code for proteins. Interesting parts, like busy cities, represent genes that play an important role in protein synthesis. This analogy is intended to help students understand that there are varying parts of the genome, and that most of it consists of DNA that does not code for proteins. Switching between two very different scales is intended to help students understand the relative magnitude of things they cannot directly observe.

Another characteristic of some analogies is that they use real physical models or demonstrations to illustrate a new topic. Sometimes these analogies are enriched versions of analogies that convey similarities in size, shape, or structure. Although other times, there is little physical similarity and the focus is on functions and behaviors. For example, an analogy can be used to teach students about the interdependence of ecosystems and food webs. By having students physically connect to each other with string, they can understand how disturbances in one area can affect many others. The connections between students bear no physical resemblance to real ecosystems, but the goal is to have them transfer their observations about movement and stability to the stability of ecosystems. On the other hand, a more physically focused analogy with a physical demonstration is used to teach students about the geologic record of the earth. This analogy works by having students create a timeline in real space to illustrate the large distances (i.e. times) between, for example, the extinction of dinosaurs and the present era. Like the genome analogy, the geologic record analogy illustrates the relative magnitude of different spans of time, but it achieves this with a real physical model.

Note that these categories are not mutually exclusive. Some analogies involve similar physical features as well as similar functions. For example, in earth science, layers of sponges are used to model salinity in soil (Figure 4). Using water, salt, and a heat source, water evaporation and salt crystal formation can be observed. This allows students to understand dryland salinity in terms of an observed physical model. The physical model is aimed at illustrating that the heat causes water evaporation, which leaves behind salt crystals. However, these causal relationships are also tied to spatial information. The water rises to dry layers that are above wet ones. The water table is at the lowest level of sponges. The physical configuration of the sponges also reflects the relative physical structure of the layers of soil. For these reasons, this analogy involves both functional and physical similarities. This analogy also happens to achieve this through the use of a real physical demonstration.

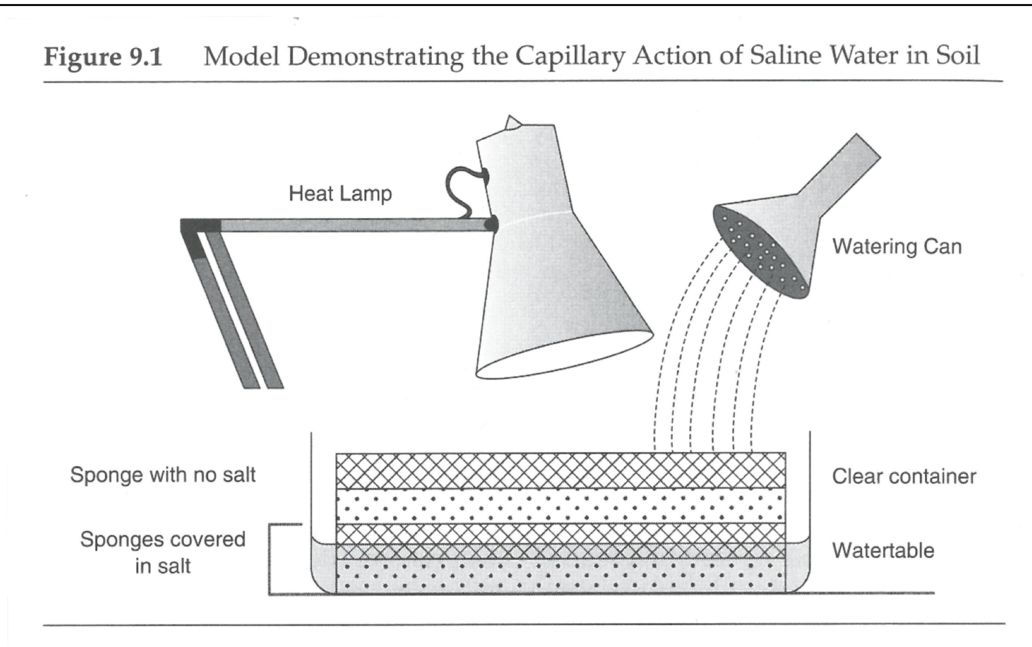


Figure 4: An example of an analogy from the FAR guide (Harrison & Coll, 2007, p. 234). This analogy expresses similarities in functions/behaviors and spatial relations through a physical model.

The wide range of analogies found in the FAR guide illustrate that they can be characterized by multiple overlapping categories. The reasoning goals of these analogies indicate that a model of interpretation needs to represent functional and physical properties at a level of abstraction that transcends domain. In other words, even though the cross-domain functional and physical properties are not literally the same, they need to be represented as if they are. This is similar to the way novices represent this information before they learn the domain's technical vocabulary. For example, mitochondria are the sites where phosphate molecules bond with adenosine diphosphate (ADP) to create adenosine triphosphate (ATP), which is the most common source of cellular energy. Rather than representing this process at this level of detail, it can simply be represented as mitochondria *providing* energy. Similarly, one could go into greater detail about how power stations generate electricity, but instead, saying that power stations provide energy is more amenable to a match with our description of mitochondria. Language plays a critical role in determining what level of abstraction is used and a successful interpretation must make this determination correctly.

3.3 Spatial Representations

Most of the analogies in the FAR guide are accompanied by pictures or spatial representations of some kind (Table 3). For the purposes of this investigation, I define spatial representations to be any non-linguistic input that provides spatial information. This includes photos, sketches, diagrams, and physical demonstrations where the learner visually observes something. Of the 49 analogies in the FAR guide, 45 had a spatial representation of some kind associated with it. In most cases, the spatial representation depicted the base domain (44). In fewer cases (13), there was a spatial representation for the target. In all but one of the analogies, anytime there was a spatial representation for the target there was also one for the base.

Topic	Base Spatial Representation	Target Spatial Representation	No Spatial Representation
Biology	10	6	2
Chemistry	11	4	1
Earth/Space Science	10	0	0
Physics	13	3	1
Total	44	13	4

Table 3: Number of analogies in the FAR guide that use spatial representations.

Although there were some analogies that were accompanied with pictures, the vast majority (39 out of 45) were presented with either diagrams or a physical demonstration of some kind. The presence of diagrams indicates the need to derive conceptual information from abstract symbols like arrows and topological relations like containment. For example, the diagram for the analogy between ATP and a battery uses arrows to indicate the transformation of ATP to ADP and vice versa. The presence of physical demonstrations indicates the need to understand physical processes or situations in a way that supports the acquisition of qualitative knowledge. For example, in the analogy for describing interdependence in the ecosystem, students need to understand that connections between students indicate interdependence and that just as in the physical model, a change in the ecosystem ripples through to impact connected species. The presence and use of spatial representations to convey conceptual information indicates that spatial reasoning is an important aid in understanding these analogies. The experiments in Chapter 4 investigate how often the spatial representations are actually *required* for interpretation.

3.4 Background Knowledge

Instructional analogies vary in how they are presented and in how much implicit background knowledge is used by the learner. Implicit background knowledge is knowledge about the base domain that isn't explicitly stated or shown in the presentation of the analogy. All analogies in the FAR guide are presented in table format, so the many of the important correspondences are usually explicit. However, even this type of explicit representation can leave out important information. Take for example the analogy between ATP and batteries in Table 4 and the analogy between the cell and Earth in Table 5.

Rechargeable Battery	**ATP and ADP**
A charged battery is energy sufficient.	ATP is energy sufficient.
Batteries move energy to where it's needed to power an electronic device.	ATP is moved to where energy is needed in the cell.
A charged battery converts into a flat battery when energy is used.	ATP converts to ADP when energy is used.
A rechargeable battery can be used over and over again.	ATP can be used over and over again: ADP + P → ATP
A battery recharger is the site where energy is reintroduced to the flat batteries	Mitochondria are the sites where energy is used to change ADP to ATP.
Recharged batteries can be used for a variety of machines for a variety of tasks.	ATP can be used at many sites in the cell (ribosomes or cell membrane) for tasks such as protein production or active transport.

Table 4: Analogy between ATP/ADP and a rechargeable battery, Harrison & Coll (2007), p. 94

The Earth to a street address	**The cell to a codon**
The Earth	A cell
A continent	The cell nucleus
A state	A chromosome
A city	Chromosomal DNA fragment
Street Address	Codon

Table 5: Analogy used to explain the relative size of cells and some of their parts, Harrison & Coll (2007), p. 118

The analogy used to explain ATP is much more detailed and contains more technical language. The primary goal is to understand how ATP provides energy for various cell activities and that it can be used repeatedly. Understanding this process in terms of batteries can be a useful foundation, but the same information is also given explicitly in the text of the analogy. This is not to say that the base domain of batteries serves no purpose, but it does demonstrate that sometimes explanations of the base domain are included explicitly with the analogy. These statements serve as indicators for how the mapping between the two domains should work. In contrast, the cell/earth analogy in Table 5 only lists correspondences between entities. It is much less detailed. The purpose of this analogy is to understand the relative size of the cell and its parts, even though size and part-whole relationships are not explicit in the language. This indicates that the learner is expected to fill in the gaps with background knowledge (and visual information if there is any). As with spatial information, experiments in Chapter 4 provide evidence that background knowledge does impact interpretation.

3.5 Target Knowledge

The target knowledge for instructional analogies at this age level (i.e. middle school) is qualitative. Specific technical, quantitative information is rarely addressed in these analogies. Analogies that focus on functional or behavioral similarities rely on qualitative representations and abstract (sometimes metaphorical) language. Analogies that focus on physical similarities have to do with relative, not absolute, size and structure. Using the mitochondria and power station example from earlier, the production of energy is described as an act of making something available to the cell, rather than a chemical reaction that creates ATP molecules. In another analogy, homeostasis is compared to walking up a down escalator. The speed of the escalator going down and the speed of the person going up must be equal in order to maintain homeostasis. The quantitative values of those speeds and the resulting body temperature are not communicated in the analogy. Instead, it is communicated that the speed of the escalator has a negative influence on the body's temperature and the speed of the person has a positive influence on the body's temperature. These qualitative relationships convey the meaning of this analogy. The only quantitative value that is given is the normal body temperature of the human body, because it represents boundary condition between qualitatively different states (i.e. hyperthermia and hypothermia). In the cell/earth analogy (Table 5), the cell and its parts are compared to the Earth and its parts to convey the very large size differences between small parts and big parts. However, quantities are only used to indicate that there are *many* sub parts in both domains. The actual cardinality values are not used, and for good reason, since using those values might bring a student to conclude that every cell has six or seven nuclei. In all cases, things that are explicitly mentioned in the analogy provide clues as to what is important. However, because implicit background knowledge is sometimes required, just because something isn't explicitly mentioned, does not mean it is unimportant.

The target knowledge for instructional analogies is intended to be general so that it can be applied to new situations or domain instances. This presents a challenge for AI systems attempting to learn from these analogies because the analogies do not explicitly identify which information is general and which information is incidental. This is a known problem in natural language understanding, referred to as the problem of generics (Leslie, 2008), where statements about particular things may be intended to apply to all (or most) other things of the same kind. For example, given a diagram that depicts an enzyme, the analogy may express, in language, that *the enzyme* has an active site. Despite the use of the definite article, *the*, the fact that the enzyme has an active site is intended to be understood as a property of all enzymes. Successfully interpreting these analogies depends on knowing how to generalize certain parts of the analogy, and knowing which parts to generalize.

Chapter 4: Interpretation of Multimodal Instructional Analogies

Building software that can understand instructional analogies has two potential benefits: knowledge capture and education. In this chapter I describe my approach for building a system that can understand multimodal instructional analogies, using examples from the FAR guide.

4.1 Problem

There are several challenges in building an algorithm for interpreting multimodal instructional analogies. These challenges overlap with the reasoning requirements identified in Chapter 3: focused alignment, spatial reasoning, background knowledge, and qualitative/type-level representations. As shown in Chapter 3, instructional analogies often consist of natural language and visual representations. Interpreting these two sources of information independently (including natural language processing and spatial reasoning) is a challenge, as is combining them in a meaningful way. Information about the base and target domains is often interspersed and it is up to the reader to figure out which information describes the base domain and which information describes the target. People also bring background information into the foreground so that it can be used in the analogy. This allows them to go beyond what is explicitly stated. Using qualitative and/or type-level knowledge enables individuals to generalize the knowledge being described in an individual analogy to an entire concept or topic. Building this ability into an instructional analogy learning system would enable it to make inferences about types of things rather than individual examples. There is also the challenge of building the analogy itself. In this thesis, I use the structure-mapping engine to build mappings with constraints that are guided by the types of cross-domain analogies found in the FAR guide. Lastly, once a mapping is created via structure-mapping, inferences need to be evaluated to determine which ones can be accepted as correct.

There is no single way to interpret an instructional analogy. Different approaches can vary with respect to modality, level of human-computer interaction, and level of incremental reasoning. I chose to build a system that analyzed simplified English text and a CogSketch sketch file to build qualitative knowledge without human intervention. I chose this approach for two reasons. The first reason is an assumption that learning independently represents the most basic way to illustrate the effectiveness of a learning system. That is, if a system can learn a topic independently, that provides the strongest evidence that its learning approach is sufficient. The second reason is that excluding human intervention during the learning process could reveal the obstacles to building independent learning by reading systems that can make use of instructional analogies found in textbooks and online sources. Points of failure provide insight into how the learning approach could be improved, either through improved sketch and/or natural language understanding or through human intervention.

4.2 Approach

The general approach for interpretation is summarized in Figure 5. As a testbed for this problem, I gathered analogies from the FAR guide used to describe topics in biology (9) and electricity (2). Each analogy consists of natural language text describing similarities and differences. Most analogies also consist of some kind of spatial representation (either a picture or diagram). Each analogy was manually adapted from its original form to simplified English and a CogSketch sketch. The interpretation model takes as input a text file and a sketch file, and produces symbolic knowledge to represent the analogy described in those files and information about the newly learned target domain. The interpretation process occurs in five steps: sketch interpretation, text interpretation, multimodal integration, case extraction, and mapping. Each of these steps is outlined below, after a description of the background knowledge that is given to the system independent of the individual analogies.

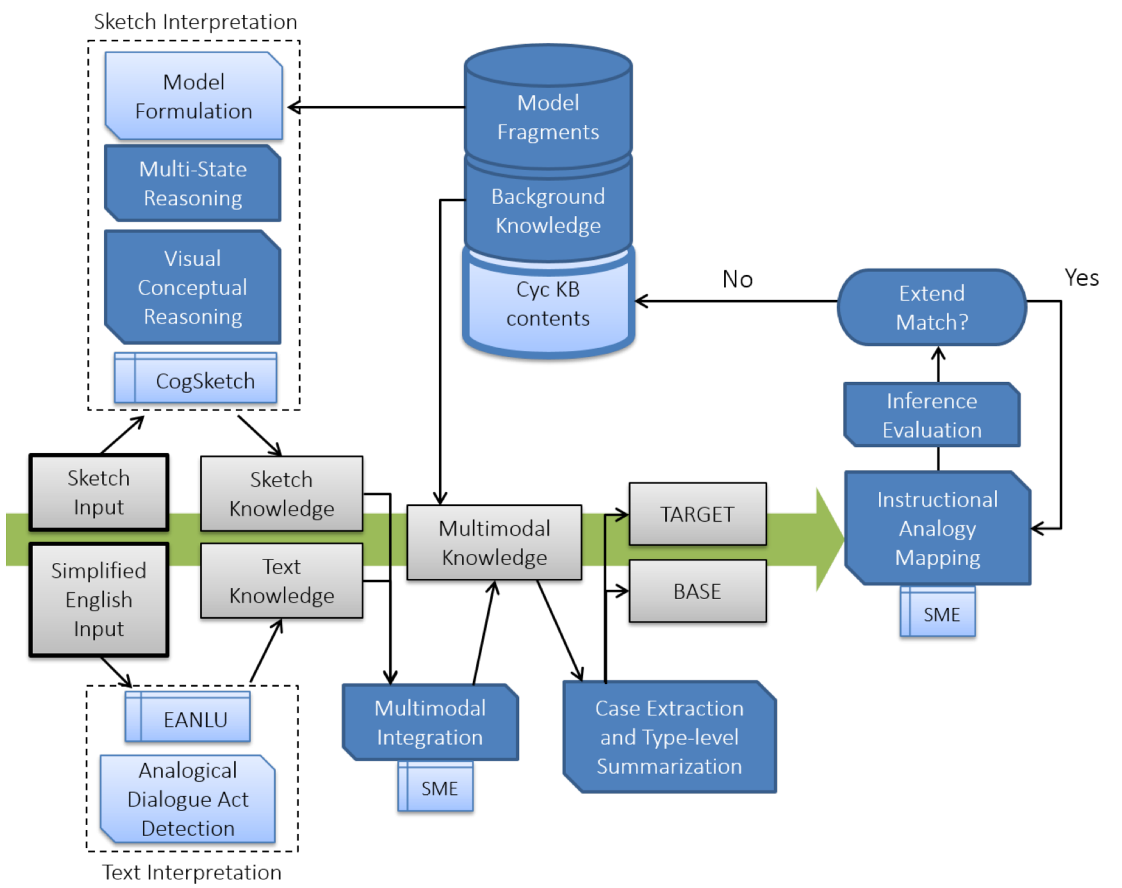


Figure 5: Overview of the system. The input to the model is a sketch (i.e. CogSketch sketch file) and a text file with a simplified English description of the analogy. Light blue items represent systems or knowledge sources that pre-date this thesis. Dark blue items indicate new techniques or knowledge sources developed for this thesis.

4.2.1. Background Knowledge

As illustrated in Chapter 3, instructional analogies depend on background knowledge about everyday concepts or experiences. This general knowledge can be represented using type-level rule macro predicates in Cyc and model fragments. Cyc has many rule macro predicates built in (examples shown in Table 1). However, out of the box, they do not provide sufficient coverage for the types of basic knowledge that were relied upon for the instructional analogies in the FAR guide. To supplement existing knowledge in Cyc, rule macro predicates like the ones shown in Table 1 were manually written and added to the knowledge base. All of the rule macro predicate statements captured knowledge about the base domain only. For example, in an analogy that compares a cell to a city, background knowledge about cities (e.g. that all cities have mayors) is imported into the system's interpretation of the analogy, even though such knowledge may not be explicitly stated. The background knowledge statements do not represent rules that are universally true. That is, it is possible that there are cities without mayors, but for the purposes of interpreting science analogies, these statements are accepted as part of the basic knowledge that is assumed to be known by the learner.

The second source of background knowledge was represented with model fragments, which are useful for representing commonsense knowledge about processes and scenarios. As was the case for rule macro predicates, model fragments were only used to represent knowledge in the base domain. Model formulation has been used previously with sketches (Chang et al., 2014; Klenk et al., 2011), and I extended the 2-dimensional mechanics domain theories in Chang et al. (2014) with those described in Table 6. The analogies in the FAR guide did not always involve qualitative processes and quantities, but when they did, the model fragment definitions in Table 6 were needed to capture the phenomena being described. For example, the analogy depicted in Figure 7 required a very specific type of situation: walking up an escalator that is moving down. This specific scenario is intended to teach students about homeostasis. Although this scenario is very specific, the model fragment used to describe it is made up of two more general purpose model fragments for movement. The escalator moving down is represented with an instance of a motion process, while the person moving up is represented with another instance of a motion process. When these two processes are active under the right conditions (i.e. the person is walking up, the escalator is moving down, and the person is *on* the escalator), then the more specific model fragment is also active. To interpret the one of the analogies about electricity, the system used a model fragment that described liquid flowing out of a container. To interpret an analogy about ecosystems, the system used a model fragment that described a physical role play scenario, where students form a web made of string.

Background knowledge, whether from rule macro predicates or model fragments, provide implicit knowledge about the base domain. When interpreting analogies, this implicit background knowledge is useful because it allows base domain knowledge to be projected onto the target domain even if that base domain knowledge is not explicitly stated.

Analogy	Model Fragment Type	Quantities
Escalator / Homeostasis	Walking up an escalator that is moving down (compositional; more general motion processes are participants to this fragment)	Height/Elevation, speed of walker, speed of elevator
Voltage / Pressure	Water flowing out of a container	Pressure, depth, flow rate
Ecosystem/Web	A student web role playing scenario; movement within the web increases instability.	Stability, speed of movement

Table 6: Model fragments used for interpreting analogies.

4.2.2. Sketch Interpretation

Sketch interpretation involved analyzing the CogSketch sketch file associated with an analogy and performing spatial and conceptual reasoning to extract relevant facts from the sketch. The facts included a combination of information automatically computed by CogSketch (i.e. topological and positional relations) along with inferences that were generated from visual conceptual reasoning, qualitative modeling, and multi-state reasoning. The extensions that were used to enable these three types of reasoning are detailed next.

4.2.2.1 Extensions to visual conceptual relations

The first extension required writing rules to bridge spatial relations that are computed by CogSketch with ontologies in Cyc. Because CogSketch uses a conceptual labeling interface, a significant amount of conceptual knowledge comes directly from the user. The spatial relations that are computed by CogSketch (automatically and on-demand) are combined with conceptual information to make new inferences, which are referred to as *visual-conceptual relations* because their meaning depends on spatial and conceptual information (Forbus et al., 2005). There are two categories of visual-conceptual relations that I implemented: part relations, and relations resulting from drawn arrows.

Relation	Spatial Antecedents	Conceptual Antecedents
`visualSubsetOf`	Hollow containment	Sets or Collections, Containers, Events
`visualSubsetOf`	Glyph grouping	None
`possessiveRelation`	Visual subset	None
`physicalParts`	Visual subset	Partially tangible entities
`subsetOf`	Visual subset	Sets or Collections
`subEvents`	Visual subset	Events

Table 7: New visual conceptual part relations.

Part relations provide a general way for representing things and their parts. Parts can be physical or conceptual.

Table 7 shows the new visual conceptual relations that were developed to support part-whole reasoning. The inference chain for part relations stems from topological relations, which are automatically computed by CogSketch, and by glyph grouping, which is initiated by the user. Proper part relations in the rcc8 vocabulary (Cohn et al., 1997) indicate when one object contains another. If the containing glyph represents a set, container, or event, then the contained glyph is a `visualSubsetOf` the container glyph. For example, if a glyph of a bottle contains a glyph of water, then the water glyph is a `visualSubsetOf` the bottle glyph. Glyph grouping can also indicate visual subsets, although unlike topological relations, glyph grouping is an explicit user action, so it can be used to infer `visualSubsetOf` no matter what the conceptual labels are. For example, the cell drawn in Figure 3 uses glyph grouping to indicate that when those glyphs are grouped together, they represent a cell. In that case, each of the grouped glyphs is a `visualSubsetOf` the glyph that represents the cell. The visual subset relation is a spatial antecedent for the rest of the part relations, which can be inferred depending on the conceptual labels of the glyphs. If the glyphs are instances of the `PartiallyTangible` collection in Cyc, then the `physicalParts` relationship holds. If the glyphs are sets or collections (i.e. instances of `SetOrCollection` in Cyc), then the `subsetOf` relation holds. If the glyphs are events (i.e. instances of `Event` in Cyc), then the `subEvents` relation holds. Possession (`possessiveRelation`) is the most general of the part relations, holding for sets, collections, containers, and anything that is involved in a glyph grouping action. For example, because the cell nucleus in Figure 3 is a `visualSubsetOf` the cell, the relation `(possessiveRelation the-cell the-nucleus)` holds. The use of `possessiveRelation`, rather than or in addition to more specific relations like `physicalParts`, matches the way such relationships are often expressed in natural language, e.g. "The cell *has* a nucleus." This statement, although imprecise, is also at a level of abstraction that transcends domain, which is useful for interpreting instructional analogies.

Arrows are used in a variety of ways in sketches. Sometimes they are used to indicate motion, e.g. a person walking *up* an escalator. Sometimes they are used to indicate the transfer of energy or a change in state, e.g. a battery changing from charged to dead. Sometimes they are used to indicate causality, e.g. one state causes or leads to another. In CogSketch, each arrow has a conceptual label, but usually, additional information can be inferred based on what the arrow represents and how it is oriented. CogSketch already interprets arrows that are labeled as direction of movement annotations and arrows that are labeled as explicit relations, but prior to this thesis it did not reason about arrows as events. Within the Cyc ontology, neo-Davidsonian representations are used to indicate thematic role fillers, and we can thus make hypotheses about the roles filled in these events based on how the arrow is drawn (e.g. Figure 6). My extension to these capabilities was to write rules that capture the meaning of the role relations in Table 8.

The set of visual conceptual relations described here are by no means complete, but they serve as a vocabulary for arrows in diagrammatic representations that are sufficient for achieving the results described below.

Event Type	Inferred On Demand

Creation event	`(doneBy arrow-object arrow-src)`
	`(outputsCreated arrow-object arrow-dest)`
	`(products arrow-obj arrow-dest)`

Destruction event	`(doneBy arrow-obj arrow-src)` `(inputsDestroyed arrow-obj arrow-dest)`

Making Something Available	`(doneBy arrow-object arrow-src)`
	`(objectActedOn arrow-object arrow-dest)`
	`(transferredObject arrow-object arrow-dest)`

Using an object	`(doneBy arrow-object arrow-src)` `(instrument-generic arrow-object arrow-dest)`
Intrinsic state change event	`(objectOfStateChange arrow-obj arrow-src)` `(toState arrow-obj arrow-src)`

Table 8: New event role relations based on arrow label and arrow direction. For each row, if there exists an object named `arrow-object` that belongs to the specified event type and points from `arrow-src` and to `arrow-dest`, then the specified facts can be inferred on-demand.

Given:

```
(isa top-arrow
     IntrinsicStateChangeEvent)
(isa bottom-arrow
     IntrinsicStateChangeEvent)
```

Available on demand:

```
(objectOfStateChange top-arrow
     charged-battery)
(toState top-arrow
     drained-battery)
(objectOfStateChange bottom-arrow
     drained-battery)
(toState bottom-arrow charged-
     battery)
```

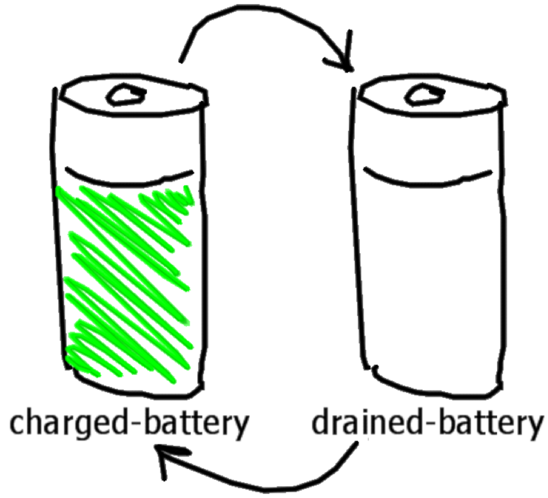
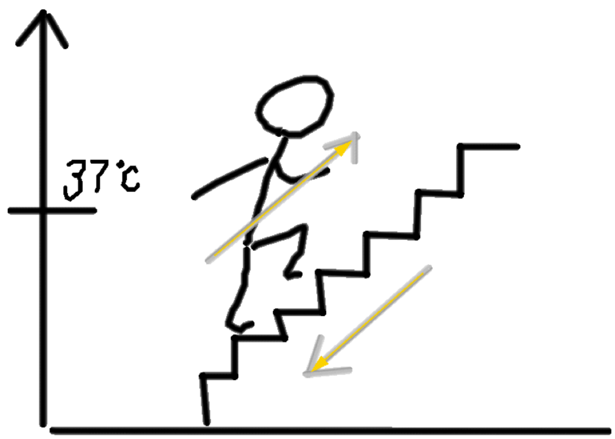


Figure 6: An illustration showing how new visual conceptual relations allow events to be drawn as arrows. The top arrow represents using a battery. The bottom arrow represents charging a battery.

4.2.2.2 Qualitative Modeling

Using the domain theories for 2D mechanics (Chang et al., 2014) with extensions shown in Table 6, model formulation (Friedman & Forbus, 2011) was used to model analogies that involved qualitative causal relationships. Three new model fragment types were defined for this system (Table 6) and they are part of the general background knowledge about the base domain. To determine if qualitative modeling was necessary for a particular analogy, the system checked for the presence of quantities in the analogy. If there were any quantities, the system searched for model fragments that might apply to the sketched scenario. The relevance of a particular model fragment type was determined by its constraints and conditions. For any active model fragments, the system asserts consequences of the model fragment, which typically indicate causal relationships between quantities. For example, in the homeostasis/escalator analogy, one of the inferences is that there is a positive qualitative proportionality between the height (i.e. level) of the person walking and their speed, as well as a negative qualitative proportionality between the height (i.e. level) of the person walking and the speed of the escalator (Figure 7). There is also a qualitative correspondence that indicates that when the speeds of the person and the escalator are equal, the rate of change of the person's height is zero. These relationships are part of the target knowledge for this analogy. Qualitative modeling was used to interpret three out of eleven analogies in the current dataset.




```
(qprop ((QPQuantityFn Height) ?person) (RateFn ?person-motion)) (qprop-
((QPQuantityFn Height) ?person) (RateFn ?esc-motion)))
(qpCorrespondence
       (RateFn ?person-motion) (RateFn ?esc-motion)
       (QPDerivativeFn ((QPQuantityFn Height) ?person)) Zero))
```

Figure 7: The sketch associated with an analogy used to explain homeostasis. The person is walking up an escalator that is moving down. The model fragment that describes this scenario has the three consequences shown below the sketch. The first statement means that the faster the person walks, the greater their height. The second statement means that the faster the escalator moves, the lesser the person's height. The third statement represents a quantity correspondence: when the person and escalator move at the same speed, the height of the person is not changing.

4.2.2.3 Interpreting Multi-state Processes

Many of the analogies described processes. In one of the analogies examined, the diagram depicted the process using multiple states (Figure 1, Figure 8). The enzyme/key analogy depicts three states in the target domain: (1) the enzyme and substrate are independent, (2) the enzyme binds with the substrate, (3) the enzyme breaks the substrate apart and unbinds.

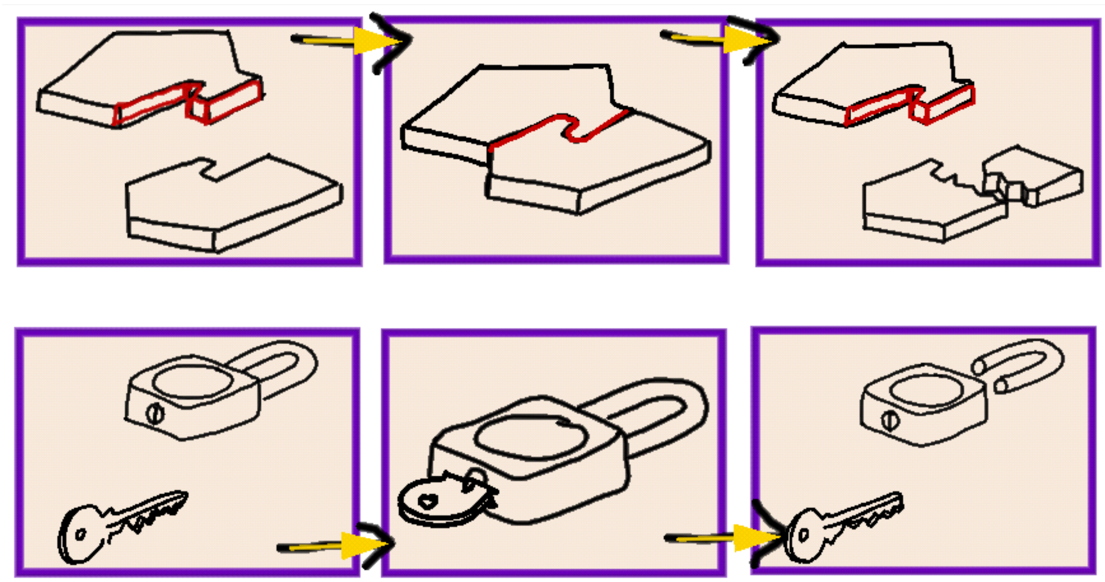


Figure 8: The CogSketch sketch for the enzyme-key analogy. Each purple box is an individual subsketch, which is used to represent an individual state. The arrows between states are enablement relations to indicate the order of the sequence.

The first challenge of representing multi-state processes has to do with contextualizing knowledge correctly. CogSketch already supports proper contextualization through subsketches. In Figure 8, each purple box is an individual subsketch with its own logical reasoning environment. This allows the sketch to represent the same object under different conditions.

The second challenge arises when there are relationships between subsketches that need to be generalized to an entire concept. Because instructional analogies at this age level tend to focus on qualitative, type-level knowledge, it is important that relationships can be generalized to an entire concept, rather than just an individual instance. In this case, the arrows represent enablement relationships to convey that one state leads to the other from left to right. The purpose of this analogy, however, is not just to learn on this single situation, but on other situations like it, i.e. enzyme activation events. Conceptual labels provide some additional information: the final state is labeled as a breaking event (because the substrate molecule breaks apart). The first two states are simply sub-events of an enzyme activation event. If the enablement relationships were generalized using only this information, we would get underspecified type-level relationships, e.g. "sub events of enzyme activations enable sub events of enzyme activation." A better approach would be to analyze the sub events to understand what makes them different from each other.

The system does this via structure-mapping. Each subsketch is compared to the others to look for properties that are unique to each state. This allows the model to describe each state more precisely. More specifically, for each state, the model finds pairwise candidate inferences by comparing it to each of the other states using the structure-mapping engine. Each candidate inference is a fact that is true in the base state, but not the target state. Each reverse candidate inference is a fact that is true in the target state, but not the base state. These facts can be generalized into existential rule macro predicates (like the ones shown in Table 1) that qualify the collections of each state. For example, when comparing the first state to the second, where the first is the base and the second is the target, there is a reverse candidate inference that the enzyme intersects with the substrate molecule. This is another way of saying that an important difference between the first state and the second state is that the enzyme intersects with the substrate in the second state. The model translates this into a property of the first state: the enzyme and the substrate do *not* intersect. This fact can then be generalized into a type-level rule macro predicate, which states that there does not exist an enzyme that intersects with a substrate molecule. This rule macro predicate becomes a general property of the first state. Now, instead of knowing only that the first state is a sub-event of enzyme activation, the model now knows that the first state is a sub-event of enzyme activation where the enzyme is not in contact with the substrate molecule. By qualifying the collections of all three states, the enablement relations become more meaningful. Where they used to express that one event enabled another, it can now express that a state where the molecules are *not* in contact enables a state where they *are* in contact (which in turn enables the final state where the substrate has broken apart).

4.2.3. Text Interpretation

Text interpretation was completed using the existing EANLU pipeline, with one new approach for disambiguation. EANLU has several optional disambiguation heuristics that guide the selection of interpretations choices. The two heuristics that are used to make disambiguation choices for instructional analogies are *favored context* and *information gain* heuristics. The favored context heuristic takes as input a microtheory name, which can be used as a frame of reference for interpreting the current reading. For instructional analogies, the only input to the favored context heuristic is the name of the microtheory that contains the information stored in the sketch. Since glyphs in CogSketch have conceptual labels that are taken directly from the Cyc knowledge base, they provide direct evidence for particular interpretation choices. Given an interpretation choice, the favored context heuristic counts the number of collections and predicates in the choice that are present in the sketch. The number of common collections and predicates is the preference weight assigned to that choice. This means that choices that involve more things that are also mentioned in the sketch are given preference. All of these heuristics predate this thesis, except for the use of predicates (in addition to collections) in the favored context heuristic. Lastly, I used Barbella's (Barbella & Forbus, 2011) analogical dialogue act interpretation to detect when the analogy is being introduced, when correspondences are being introduced, and to detect which key elements are parts of the base and target domain respectively.

4.2.4. Multimodal Integration

Once the initial sketch and text interpretation is complete, the two representations need to be merged to take advantage of information from both sources. This problem can be characterized as an alignment problem (Lockwood & Forbus, 2009). I extended the approach used in Lockwood's work by using language that has been automatically disambiguated and by using event interpretations to provide support for accurate alignments.

Structure-mapping is used to align the knowledge gathered from the analogy's text and sketch sources. Aligning these two representations can be characterized as a very near, within-domain analogy. The two descriptions are literally describing the same things. Using partition constraints for all collections is a way to ensure that terms can only map to each other if they belong to the same collection(s), for example, keys match with keys, locks with locks and so on. This works very well when the conceptual labels in the sketch are identical to the interpretations generated by EANLU. However, even when there is considerable overlap in conceptual information, each modality typically includes unique information. For example, the enzyme/key analogy shown in Figure 1 (with CogSketch sketch shown in Figure 8) has substantial overlap in conceptual information because each glyph in the sketch has a label (`EnzymeMolecule`, `SubstrateMolecule`, `EnzymaticBindingSite`, etc.). These labels guarantee that the enzyme molecule mentioned in the text modality will align with the enzyme shown in the sketch because partition constraints are used. However, there are several things that are mentioned in the text that are not present in the sketch, e.g. the binding site's unique chemical makeup, the key's unique shape, the fact that enzymes can be reused, etc. Similarly, there is information in the sketch that is not present in the text, e.g. the spatial intersection of the enzyme and substrate, the sequential nature of enzyme activation, etc. The differences between the two modalities mean that there are opportunities for mismatches. This is especially true for analogies that involve concepts that are not explicitly drawn, e.g. energy. The system reduces the chances for mismatches by reasoning about events in the text interpretation and projecting that information to the sketch modality.

Generic event interpretation takes the neo-Davidsonian event interpretations generated by EANLU and projects them to the sketch modality (Figure 9). It is based on the assumption that when instructional analogies describe events, they describe possible generic roles for the primary participants of the events. For example, in the cell/city analogy, the sentence “Power stations provide electricity” is interpreted as a statement about the functions of power stations. Using the neo-Davidsonian representation, such event would be represented like this:

```
(isa provide34501 MakingSomethingAvailable)
(performedBy provide34501 power-station88374)
(objectActedOn provide34501 electricity18186)
```

To project this information into the sketch modality, the system looks at the primary actor of the event: `power-station88374`. This entity is an instance of the collection `PowerGenerationComplex`. The system searches for instances of this collection in the sketch, and if it finds one, searches for instances of `MakingSomethingAvailable` and of Electricity. In the case of the cell/city analogy, there is an instance of `Windmill`, which is a subcollection of `PowerGenerationComplex`, but there are no instances of `MakingSomethingAvailable` or `Electricity`. The system then generates two speculative entities: an event that is an instance of `MakingSomethingAvailable` and an entity that is instance of `Electricity`. The system asserts role relations between the entities, resulting in the following representation in the sketch modality:

```
(isa (SpeculativeThingFn event52582) MakingSomethingAvailable)
```

```
(performedBy (SpeculativeThingFn event52582) windmills)

(objectActedOn (SpeculativeThingFn event52582) (SpeculativeThingFn electricity52843))
```

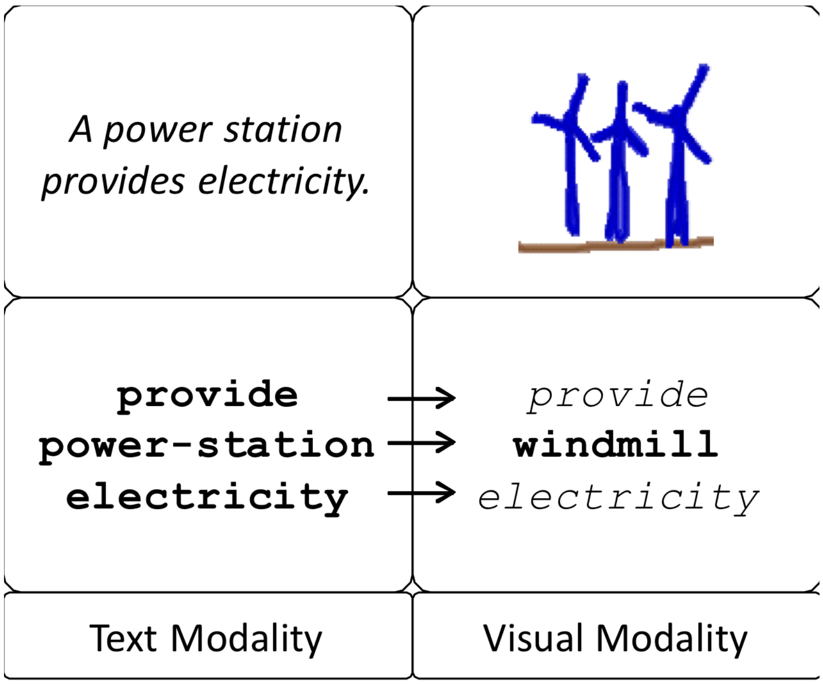


Figure 9: An illustration of how generic event interpretation works. Event information from the text modality is assumed to apply to elements in the visual modality. Providing electricity is projected onto windmills because they are specializations of power stations..

The system also generates statements that describe the events with more specific collections, similar to the way multi-state processes (described above) are interpreted. Because the goal of these analogies is often to build type-level knowledge, denoting collections with specific terms can be useful. For example, the cell/city analogy conveys functions and behaviors of the parts of the cell. The initial language interpretations use verbs to detect what kinds of events are being described. Using power stations as an example, the verb "providing" indicates that the event being described is an instance of `MakingSomethingAvailable`. The role relations provide further information, and the generic event interpretation procedure uses collection denoting functions (described in section 2.2) to indicate that the electricity-providing event is an instance of a more specific collection:

```
(isa provide34501 (MakingAbstractAvailableFn Electricity))
```

Similarly, this information is projected into the visual domain as well:

```
(isa (SpeculativeThingFn event52582) (MakingAbstractAvailableFn Electricity))
```

These collection membership statements provide greater structural support for putting windmills and power stations into correspondence. This is particularly useful in this case because the conceptual attributes for these entities do not match exactly, so it is not guaranteed that they will correspond based on partition constraints alone. Lastly, the use of collection denoting functions is also useful for representing type-level information. Rather than making general assertions about events where something is made available, the system can now make general assertions about events where electricity is made available.

The analogical mapping between the text information and sketch information determines how the information is merged. Corresponding entities and facts are merged together and all facts that were not in the mapping are included in the final representation as is. The advantage to using alignment and merging (rather than a simple union of facts) is that it allows aligned entities to refer to the same conceptual thing. For example, when combining facts about the cell/city analogy, we do not want a multimodal representation where there are two nuclei. The nucleus that is mentioned in the text is assumed to be the same nucleus that is drawn in the sketch because the two nuclei correspond to each other in the analogical mapping. Consequently, the final representation contains knowledge about the nucleus controlling the cell (a fact which comes from the text) and about the nucleus being a proper part of the cell (a fact which comes from the sketch). Including facts that were not in the mapping is also important, because it allows the system to include facts that are present in one modality but not the other.

4.2.5. Case Extraction

Case extraction is the process by which the base and target descriptions are built from the integrated multimodal description. Case extraction uses three sources of information to split facts in the multimodal description into a base and target: analogical dialogue acts, event role relations, and part-whole relationships. These information sources are used to build sets of entities that are hypothesized to be core elements of the base and target domain respectively. These sets are then used in an iterative search over the multimodal description. Analogical dialogue acts determine the initial members of each entity set. If an analogical dialogue act introduces a correspondence between items B and T, then B is an element of the base domain and T is an element of the target domain. For each element identified in analogical dialogue acts, the system also includes any sub parts or sub objects of that element. Sub objects are determined by possessive relations (described in section 4.2.2.1), which come from natural language interpretation (e.g. “ATP *has* energy”) and from visual conceptual relations. For each of these entities (and their sub objects), the system searches for events that are related to them. Each event along with any other participants of that event, are included in the entity set as well. Using these entity sets, the model proceeds with a very simple voting heuristic that classifies any remaining facts as belonging to the base description, the target description or as being ambiguous, based on number of known base or target items mentioned in the expression:

```
classify-fact(fact, base-items, target-items)
    base-count ← length(intersection(base-items, args(fact)))
    target-count ← length(intersection(target-items, args(fact)))
    IF base-count > target-count THEN RETURN :BASE
    ELSE IF target-count > base-count THEN RETURN :TARGET
    ELSE RETURN :AMBIGUOUS
```

If a fact involves more base entities than target entities, it is classified as being a fact about the base domain. Ambiguous statements, which involve an equal number of base and target entities, are discarded.

4.2.6. Type-level Summarization

As noted in Chapter 3, the analogies in the FAR guide are aimed at teaching type-level conceptual knowledge. Although detecting generics and representing them accurately is a known problem in natural language understanding, this model takes an aggressive approach to generalization, and assumes that certain predicates generalize as follows. Part predicates (e.g. physical part predicates, subset of) generalize with a universal quantifier over the whole and an existential quantifier over the part. For example, a sketch that illustrates a cell nucleus as part of a cell will generalize to a rule macro predicate that states that for all cells, there exists a nucleus that is a part of it. Primary actor role relations generalize with a universal quantifier over the primary actor and an existential quantifier over the event. This is especially useful for analogies that focus on behaviors and functions. When an analogy states, for example, that mitochondria provide energy, this generalizes to a fact that states that all mitochondria provide energy. Temporally coexisting predicates (i.e. binary relations between nonevents that temporally coexist) are generalized with a universal quantifier over the first argument and an existential over the second. For example, a sketch that illustrates that a cell is bigger than a nucleus, will generalize to a rule macro predicate that states that for all cells, there exists a nucleus that it is bigger than. Facts about instances are discarded unless they represent quantities or constants in the KB, e.g. `Zero` or `PlanetEarth`. By summarizing the information in the base and target with type-level predicates, the base and target descriptions become suitable for a type-level analogy, where correspondences and inferences about collections (rather than strictly entities) can be made.

4.2.7. Mapping and Inference Evaluation

The instructional analogy is created using SME with constraints that have been identified by analogical dialogue acts and event interpretations. As with case extraction, analogical dialogue acts provide the starting point for determining what constraints should be put on the mapping process. Correspondences that have been introduced in the text are used as hard constraints on the match. Events that those entities participate in and the other participants of those events are used as required correspondences as well. However, since the analogy is operating at the type-level, those match constraints are translated so that they are between concepts (i.e. collections) rather than instances (i.e. entities). For a given constraint between two entities, A and B, the collections that those entities belong to are used as constraints. Using the cell/city analogy as an example, the sentence “A cell is like a city” represents a dialogue act that introduces a correspondence between an instance of `City` and an instance of `Cell`. This correspondence is translated to a required correspondence between the collections `City` and `Cell`. This translation is necessary for capturing the type-level semantics of the analogy and for enabling inferences about categories or classes of things, rather than single examples of them.

Once an initial match is constructed, its inferences are evaluated and the model attempts to resolve skolems. Because all the statements in the match are type-level statements, skolems can be either: collections, predicates, or entities that passed type-level filtering. Such entities may be known terms in the Cyc knowledge base, e.g. `PlanetEarth`, or qualitative quantities, e.g. `Many-Quant`. Abstract entities, like quantities, and predicates are accepted as is. Collections require additional inspection to determine if they are known to be very abstract or if they involve functional predicates that were treated as entities and whose arguments have known matches. Constants and abstract collections were manually identified based on initial language and sketch interpretations. The following collections are considered abstract and therefore acceptable as is for skolem resolution: `Event`, `FunctionalSystem`, `Individual`, `PartiallyTangible`, `Set-Mathematical`, `SetOrCollection`, `SpatialThing`, `Thing`. Constants were defined as numbers or instances of the Cyc collection: `ScalarOrVectorInterval`. Instances of this collection include things like: `Now`, `Many-Quant`, `Few-Quant`, `Billions`, etc. These lists are not exhaustive, but simply a starting point based on inputs to the model. They point to one of the difficulties in interpreting cross-domain analogies that was mentioned in Chapter 3: representing terms at the right level of abstraction and knowing when a term transcends domain.

If there are candidate inferences that have been fully resolved, they are accepted as fact and stored in the target domain. The match can then be extended in an attempt to arrive at more inferences. Accepting inferences and extending the match makes it possible to build larger mappings with higher order candidate inferences. Match extending continues until there are no more candidate inferences that can be resolved, or until a maximum depth is reached (determined by the maximum depth of expressions in the base domain). When match extension and inference evaluation finishes, the final target description is stored into the KB so that it can be used for future reasoning.

4.3 Evaluation

To evaluate this model, I adapted 11 analogies from the FAR guide:

1) Rechargeable battery analogy for ATP
2) City analogy for a cell
3) Earth analogy for cell components
4) Supermarket analogy of a biological classification system
5) Lock and key analogy for enzyme action
6) Fluid mosaic analogy for cell membranes
7) Homeostasis is like walking up a down escalator
8) Web analogy for the interdependence of organisms
9) Clothespin analogy for the structure of DNA
10) Water circuit analogy for electric current
11) Water pressure analogy for voltage

In the FAR guide, each analogy consists of an explanation and a table of correspondences. For each analogy, I manually translated the language into simplified English to support interpretation with EANLU and sketched a spatial representation of the analogy into CogSketch. For each analogy, the model takes a sketch file and a text file as input and produces microtheory which contains newly acquired facts about the target domain.

4.3.1. Model Input

For building the text passages, each analogy needed to be translated from the table and text format in the book to complete sentences with simple syntax. Sentences that introduced correspondences were used to account for terms that were listed on the same row. For example, the enzyme/key analogy shown in Figure 1 has the following simplified English representation:

> *An enzyme is like a key. A substrate molecule is like a lock. The key has ridges. The ridges have a unique shape. The enzyme binding site is like the ridges. The enzyme binding site has a unique chemical makeup. The key only unlocks specific locks. The enzyme reacts only with specific substrate molecules. The key unlocks the lock. The enzyme breaks apart the substrate molecules. After unlocking a lock, the key is unchanged. The key can be used over and over again. Enzyme action is like unlocking a*

lock. The enzyme is unchanged by the chemical reaction. The enzyme can be used over and over again.

This passage is longer than the table representation in Figure 1 because the table representation in the FAR guide is often more concise than a passage that explains the analogy. For example, some rows only have the names of entities that are supposed to correspond to each other. For input to EANLU, these were translated into complete sentences, e.g. "An enzyme is like a key." Additionally, some longer sentences were split into multiple shorter ones (but more words overall) to enable interpretation in EANLU.

All but one of the 11 analogies evaluated here were presented with a spatial representation of some kind. Seven analogies were presented in the FAR guide with spatial representations (either diagrams, sketches, or photos) of the base and target domain (either separately or together in a blended representation, as shown in Figure 7). For those analogies, I sketched those representations into CogSketch and used conceptual labeling and glyph grouping to convey accurate conceptual information. As noted earlier, CogSketch does not perform sketch recognition. The kinds of things depicted in each sketch are given explicitly through labels from the Cyc knowledge base. This allows CogSketch to reason about the entities drawn in the sketch with respect to the Cyc ontology. It also allows CogSketch to reason about part-whole relationships via glyph grouping. For example, in the enzyme/key sketch (Figure 8), the enzyme and enzyme activation site are two separate glyphs, with labels, and they are grouped to convey that the activation site is part of the enzyme. The conceptual labeling and manual glyph segmentation (and optional grouping) therefore provide a great deal of conceptual information on their own. The remaining three analogies with spatial representations only had pictures for the base domain. For the analogies with only base spatial information, I supplemented the sketch with a target representation of my own so that there was enough conceptual information about the target domain to assist with the overall interpretation (the ablation experiments in this chapter will illustrate why this was necessary given the current state of this model). Lastly, one of the analogies, the cell/city analogy, had no spatial representations at all. For this analogy, I supplemented the text passage with a CogSketch sketch of a city and a cell. The full set of text passages and sketch descriptions can be found in APPENDIX A: Instructional Analogies.

As an evaluation measure, I tested the model's accuracy on a set of gold standard queries that I wrote based on the presentation of the analogies in the FAR guide. Essentially, these queries address the following question: if the model successfully interpreted the analogy, what should it know about the target domain? The expected knowledge was manually translated into Cyc queries. Table 9 shows the gold standard queries for the enzyme/key analogy in natural language and in the Cyc representational language.

Enzyme active sites have a unique chemical makeup.

```
(partTypes EnzymeMolecule
                  (CollectionIntersectionFn
                   (TheSet EnzymeBindingSite
                           (ThingDescribableAsFn Unique-TheWord Adjective))))
```

Enzymes only react with specific substrates.

```
(relationAllExistsAndOnly chemicalReactants
   (SubcollectionOfWithRelationToTypeFn
       (SubcollectionOfWithRelationToTypeFn ChemicalReaction chemicalReactants
       SubstrateMolecule) catalyst EnzymeMolecule)
         (CollectionIntersectionFn
           (TheSet SubstrateMolecule
              (ThingDescribableAsFn Specific-TheWord Adjective))))
```

Enzyme action breaks apart substrate molecules

```
(relationAllExists subEvents EnzymeActivationEvent
      (SubcollectionOfWithRelationToTypeFn
       (SubcollectionOfWithRelationToTypeFn BreakingEvent doneBy EnzymeMolecule)
       objectOfStateChange SubstrateMolecule))
```

The enzyme comes out of the reaction unchanged.

```
 (relationExistsAll unchangedActors EnzymeActivationEvent EnzymeMolecule)
```

When enzymes bind to substrate molecules, it enables the substrate to break apart.

```
(enables-SitTypeSitType
        (SubcollectionOfWithRelationToFn
         (CollectionIntersectionFn
          (TheSet
           (SubcollectionOfWithRelationToTypeFn Situation subEvents
                                                ChemicalReaction)
           (SubcollectionOfWithRelationToTypeFn Situation subEvents
                                                EnzymeActivationEvent) Situation))
                holdsIn
         (relationExistsExists objectsIntersect EnzymeMolecule
                               SubstrateMolecule))
        (SubcollectionOfWithRelationToTypeFn
         (SubcollectionOfWithRelationToTypeFn BreakingEvent doneBy
                                              EnzymeMolecule) objectOfStateChange
                SubstrateMolecule))
```

Enzymes can be reused.

```
(relationAll repeatedEvent EnzymeActivationEvent)
```

Table 9: Gold standard queries in natural language and in the Cyc representational language for the enzyme/key analogy.

4.3.2. Model Performance

The full model successfully processed all 11 analogies.

Table 10 shows the sizes of the analogies in terms of sentences in text input and number of facts at various stages of processing. The text passages varied in size but ranged between 4 and 18 sentences. Notably, the multimodal interpretation is not simply a union of the facts in each modality. Because SME is used to align the representations, entities are merged and redundant facts are avoided. Similarly, the multimodal interpretation is also not a union of the base and target cases, since there are often facts that cannot be classified as belonging to either the base or target domain (Table 11). After facts are classified as belonging to the base or target domain, they are generalized to type-level statements. The sizes of the base and target cases after type-level summarization are shown in Table 12.

Analogy Name	**Sentences in text input**	**Facts in text interp.**	**Facts in sketch interp.**	**Facts in multimodal interp.**
ATP	15	181	121	258
Cell (City)	15	160	231	321
Cell (Earth)	5	38	165	163
Classification	17	188	415	580
Enzyme	15	121	187	292
Membrane	9	98	74	144
Homeostasis	10	97	75	150
Ecosystem	4	42	250	254
DNA	12	85	374	354
Circuits	18	162	130	250
Voltage	7	51	63	90

Table 10: Summary of text and case sizes for 9 biology analogies and 2 electricity analogies.

Analogy Name	Base Facts	Target Facts	Ambiguous Facts
ATP	87	124	47
Cell (City)	153	128	40
Cell (Earth)	58	36	69
Classification	261	202	117
Enzyme	127	125	40
Membrane	54	83	7
Homeostasis	66	60	24
Ecosystem	60	164	30
DNA	138	163	53
Circuits	112	78	60
Voltage	54	26	10

Table 11: Number of facts classified as belonging to the base domain, belonging to the target domain, and being ambiguous.

Analogy Name	Type-level Base Facts	Type-level Target Facts
ATP	54	77
Cell (City)	86	88
Cell (Earth)	31	8
Classification	108	47
Enzyme	45	115
Membrane	13	35
Homeostasis	38	38
Ecosystem	15	228
DNA	23	83
Circuits	47	27
Voltage	23	9

Table 12: Number of facts in base and target domains after type-level summarization and analogical mapping.

The full model achieved an average accuracy of 83% across all 11 analogies. Ablation experiments compared the full model to three other conditions:

- Full model without the use of background knowledge
- Full model without the base domain
- Full model without the sketched input

Removing the ability to import qualitative background knowledge in the form of rule macro predicates and qualitative model fragments caused overall accuracy to drop to 59% (Figure 10). This means that background knowledge plays an important role in this model's ability to build new knowledge about the target domain. However, at 59%, this result also means that much of the knowledge that was tested in the gold standard queries was either explicitly stated in the analogy's text or sketch, was transferred from explicitly stated base knowledge, or was inferred during the visual conceptual elaboration. This is true for analogies that deal with part-whole relations, like the cell/earth analogy, where the part-whole relations are given explicitly in the sketch via relation arrows, or the cell/city analogy where part-whole relations are inferred from glyph grouping. However, as noted in Chapter 3, analogies that are less verbose might require learners to fill the gaps with background knowledge. The cell/earth is one such analogy, and its accuracy drops from 100% with the full model to 72% when no background knowledge is used. This suggests that when it comes to the importance of background knowledge, there is variation within this set of analogies and that the presentation of the analogy (complete sentences vs. correspondences only) may be a clue that background knowledge is important. As expected, the performance was worse in the no background condition for the three analogies that relied on qualitative modeling for interpretation (homeostasis, ecosystem, and voltage).

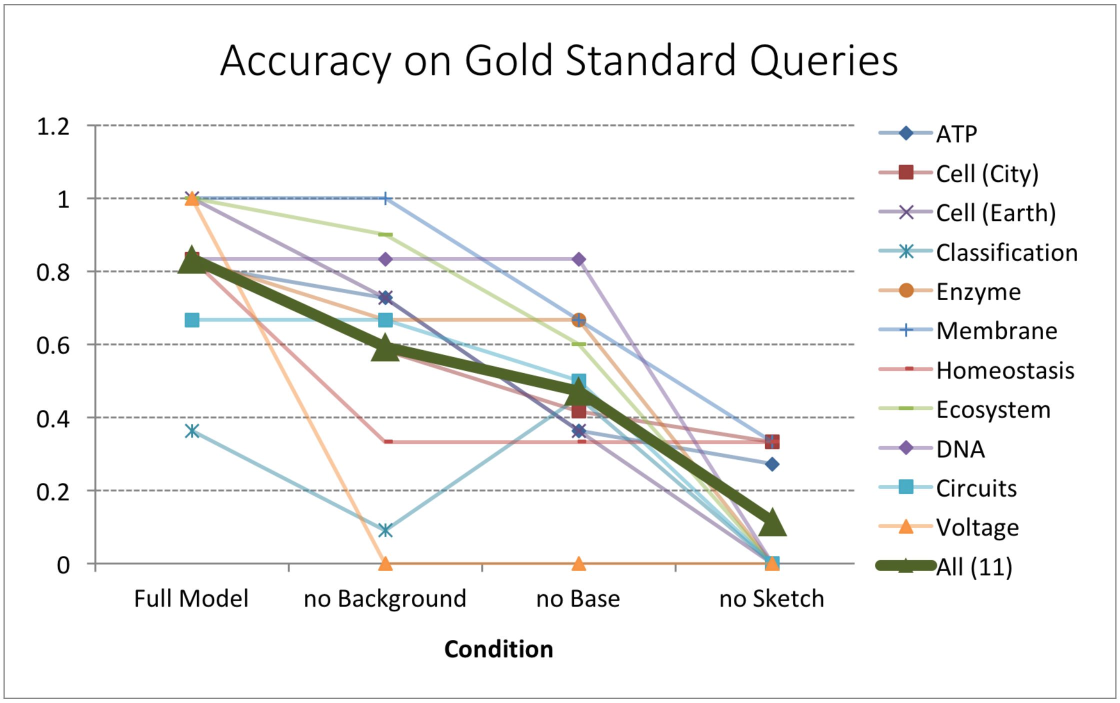


Figure 10: Results from ablation experiments by analogy. The bold line represents average accuracy across all analogies.

Analogy Name	# Gold Standard Queries	Full Model (# correct)	No Background (# correct)	No Base (# correct)	No Sketch (# correct)
ATP	11	9	8	4	3
Cell (City)	12	10	7	5	4
Cell (Earth)	11	11	8	4	0
Classification	11	4	1	5	0
Enzyme	6	5	4	4	0
Membrane	6	6	6	4	2
Homeostasis	6	5	2	2	2
Ecosystem	10	10	9	6	0
DNA	12	10	10	10	0
Circuits	6	4	4	3	0
Voltage	1	1	0	0	0

Table 13: Performance of individual analogies on each of the four experimental conditions.

The second ablation condition disabled the analogy entirely, making it both impossible to use background information from the base domain and explicit information about the base domain. This effect caused a greater drop in accuracy, to 47%. This indicates, as expected, that background and explicit knowledge about the base domain is important for interpreting the analogy. Nevertheless, a considerable amount of target knowledge (almost half) is still explicitly stated in the analogy. Interestingly, the performance on one of the analogies (supermarket analogy for classification) improved in the no base domain condition. This analogy was unlike the others in that it did not have a base and target domain where individual entities could be aligned to each other. The analogy compares the organizational structure of a supermarket with the organizational structure of biological taxonomies. In the other analogies, entities were typically introduced individually in the base and target domain (e.g. the enzyme is like a key; DNA is like a clothespin structure) and therefore represented individually in the text and sketch representations. In the classification analogy, however, groups (i.e. species and genus) and sections (e.g. dairy section, produce section) are represented with multiple diagram elements, making a clean 1:1 mapping difficult. Notably, the full model performed the worst on this analogy when compared to all of the others (4/11 queries correct; 36%). This is one case where perhaps rerepresentation is needed to accurately interpret the analogy.

The final condition removed the sketch entirely from the process. Background knowledge and explicit base knowledge were still used in this condition; it just relied entirely on the natural language input. The model performed at its worst in this condition with an accuracy of 12%. This indicates that in this model, the conceptual labels provided in the sketch play a critical role in interpretation. The conceptual labels from the sketch come directly from the Cyc ontology, meaning that there is a direct connection between the labels and the semantic interpretation choices provided by EANLU. Thus, the conceptual labels from the sketch are mostly responsible for making the right disambiguation choices.

4.4 Discussion

The experiments in this chapter support claim 1 of this thesis. They demonstrate that a combination of sketched input, simplified natural language understanding, and structure-mapping can be used to capture qualitative science knowledge. Structure-mapping is used in two main ways. First, it is used to merge sketch and text representations. It is also used to build the cross-domain analogy, which is effective at transferring knowledge from the base to the target domain based on the ablation conditions where the background knowledge and the base domain were excluded. As mentioned earlier in this chapter, structure-mapping is also used to detect important differences between states in a sketch, but this was only relevant for one out of the 11 analogies.

The ablation experiments also support the claim that background knowledge and visual representations play an important role interpreting instructional analogies. The qualitative analysis in Chapter 3 also supports this claim.

One important limitation of this model is how base domain and background knowledge is used. The ablation experiments indicated that out of the three sources of information (background knowledge, explicit base knowledge, sketch knowledge), sketch knowledge had the greatest impact on interpretation accuracy. This result cannot be used to suggest claims about the way people interpret analogies. The reason for this has to do with broader challenges in knowledge representation. For a symbolic reasoning system, such as the one built for this thesis, there is no difference between asserting a fact and having the system understand it (as far as *understand* means in this system). In this model, there is no difference between *storing* a fact and *learning* it because of the limits of the evaluation. To test the system's knowledge, it is simply queried, which is the equivalent of having someone recite a fact back to you. Better evaluations of knowledge, such as *how* it is used in other types of reasoning like question answering, would likely provide a more accurate measure of how well a system understands a concept and therefore how important background knowledge is.

Another limitation is the reliance on manual labeling of sketches. This is a very useful simplification because open-domain sketch recognition is an unsolved problem and because labeling sketches and diagrams (e.g. with written labels or spoken language) is commonplace, especially in science instruction where it has been shown to help students learn new content (Mason et al., 2013). However, this means that the reasons sketches are important for this system are not the same as the reasons visual representations might be important for people. Future experiments, such as using an oracle for language disambiguation or ablating visual conceptual reasoning only, could provide a more detailed explanation for how and why spatial representations are important.

Despite these limitations, building systems that can understand multimodal instructional analogies, even if not in exactly the same ways as people do, is still useful from an educational and interactive perspective. Interpretation abilities are needed for learning by reading and for using analogies to explain phenomena to a human collaborator or student.

Chapter 5: Instructional Analogies for Question Answering

The previous chapter demonstrated that qualitative science knowledge can be captured from multimodal instructional analogies. The evaluation of that knowledge, however, was closely tied to the analogies themselves. This chapter shows the extent to which that knowledge can be used to answer science questions that were developed independently from the instructional analogies used for learning.

5.1 Problem

Open-domain question answering has been an AI goal for decades and it has seen resurgence since the Watson Jeopardy! Challenge (Ferrucci et al., 2010). AI researchers have suggested using elementary and middle school science exams as one new benchmark for progress toward human-level AI (Clark, 2015). The reason these exams were suggested was because they require a great deal of qualitative and common sense knowledge.

Instructional analogies may be a good precursor to solving these questions because (i) they model how people are first introduced to science topics, and (ii) they are good tools for building qualitative knowledge, which is needed to answer many of these questions.

5.2 Approach

To test the utility of qualitative knowledge captured through instructional analogies, I gathered science questions from the NY state regents 4th grade, 8th grade and living environment exams[2] and the Massachusetts comprehensive assessment system biology and physical sciences exams[3]. Out of 86 questions that were related to human biology or electrical energy, I identified 14 questions that involved topics that overlapped with the topics of the analogies used in the previous chapter and did not involve graph understanding or target domain diagram recognition. Not all analogies were represented in the full question set. Some topics, like enzyme activation, were not addressed. Other topics, like ecosystems and interdependence, were represented on multiple questions. The 14 questions required knowledge about ecosystems, cell structure and function, DNA structure, homeostasis, and electrical circuits.

[2] http://www.nysedregents.org/
[3] http://www.doe.mass.edu/mcas/

To limit the scope of this question answering experiment, I made two major simplifications. First, rather than attempt natural language understanding on questions, I manually translated them from English to CycL. Each question was translated into a query. Most questions (9) could be formulated as a query with an open variable, where the answer options were candidate bindings for that variable. The rest of the questions (5) were *true statement* questions, where each answer option is a ground query (i.e. a query with no variables). If the question involved a scenario, the properties of that scenario were translated into CycL and stored in the microtheory for that individual question. Second, I only considered diagram questions when those questions did not require target domain recognition (e.g. interpreting an unlabeled image of a cell) or graph understanding. Out of the 14 questions examined, 3 involved diagrams but only 1 used a diagram that was required to answer the question. This diagram showed a partially complete circuit with labels. To make the diagram usable for question answering, I sketched it into CogSketch using the labels provided on the test. The sketch was then declared a visual aid for the question. When the model attempts to answer a question, it searches for a visual aid for the question. If it finds one, which only happened once for this set of questions, it uses the same visual conceptual reasoning (described in 4.2.2.1) that was used for learning from instructional analogies to extract additional information from the sketch, and stores the resulting facts in the question's microtheory. This way, when queries are executed to answer the question, the information from the visual aid is visible from the question's logical environment.

My approach for question answering required reasoning about question types, computing similarity between statements, reasoning about object properties, reasoning about example situations, and using abduction. Each of these abilities and the order in which they are executed in this question answering (QA) system are described next.

Object Property Question

A function of cell membranes in humans is the

a) synthesis of the amino acids
b) production of energy
c) replication of genetic material
d) recognition of certain chemicals

```
(queryForQuestion
  NYSRE-2014-LE-01
 (relationExistsAll doneBy ?fn
                    CellMembrane))

(multipleChoiceSingleOptionList
  NYSRE-2014-LE-01
  (TheList (MakingFn AminoAcid) 1))
(multipleChoiceSingleOptionList
  NYSRE-2014-LE-01
  (TheList (MakingFn EnergyStuff) 2))
(multipleChoiceSingleOptionList
  NYSRE-2014-LE-01
  (TheList Replication-DNA 3))
(multipleChoiceSingleOptionList
  NYSRE-2014-LE-01
  (TheList
   (IdentifyingAsTypeFn ChemicalObject) 4))

(correctAnswerChoice NYSRE-2014-LE-01 4)
```

Figure 11: An example question in natural language and in CycL.

5.2.1. Detecting Question Types

To build a simple question answering system, I needed a small set of heuristics that were sufficient to understand how the qualitative science knowledge could be used, rather than a large set for broad test coverage. To achieve this, I used the analysis provided by Clark and colleagues (Clark et al., 2013) to inform the strategies I implemented. They identified knowledge requirements for answering questions on elementary science tests and eight different (but not mutually exclusive) question categories. Only some of the categories identified by Clark and colleagues were present in the current set of 14: object property questions, example situation questions, and questions that involved causality and processes.

Object property questions typically ask about the properties of a class rather than a specific object. Within CycL, such queries are expressed with rule macro predicates that use universal quantification, e.g. `relationExistsAll` or `relationAllExists`. Such questions tend to be about the collection that is universally quantified. For example, the question in Figure 11 is an object property question about cell membranes. If the question query involves a universally quantified rule macro predicate, or it is a true statement question where a particular collection is mentioned in more than one option, the QA system assumes it is an object property question.

Example situation questions have knowledge associated with them that are not about the multiple choice question ontology. A simple check in the question's microtheory for facts other than those used to represent the question itself (i.e. predicates other than `correctAnswerChoice`, `multipleChoiceSingleOptionList`, and `queryForQuestion`) is a good indicator that the question involves an example situation. Lastly, there were questions that involved causality and processes, but they tended to also be example situation questions, so they were simply treated as such.

5.2.2. Interchangeability of Statements

To deal with non-identical, but semantically and syntactically similar statements, I developed a very simple interchangeability measure that was inspired by the loose speak interpreter (Fan et al., 2009). It compares two statements and generates an interchangeability score if they are related to each other using Cyc's collection and predicate hierarchies. Identical terms receive a score of 2. Terms that are not related receive a score of 0. Terms that are related receive a score that is inversely proportional to the distance between them in the `genls` (for collections) or `genlPreds` (for predicates) hierarchy. Common supercollections are also used to relate collections even if one collection is not a direct specialization of the other. Terms that represent qualitative quantities (e.g. `Many-Quant, Thousands-Quant`) are grouped into three general categories: small (e.g. `Few-Quant`), medium (e.g. `Dozens-Quant`), and large (e.g. `Millions-Quant`). The interchangeability score between two qualitative quantities is the closeness between them on the small, medium, large spectrum. Terms that are quantitative values (e.g. 37, 10) are given a score of 0 unless they are equal, since the difference between the values is not meaningful without a frame of reference for scale. Interchangeability of statements is the average interchangeability of their terms, unless any terms are unrelated, in which case the interchangeability of the statements is zero. For example, the statements `(causes-SitTypeSitType Poaching Extinction)` and `(enables-Generic HumanActivity Extinction)` receive a score of 1.1, because the predicates are related via `genlPreds` hierarchy, `Poaching` is a subcollection of `HumanActivity`, and both statements have identical third arguments. The statements `(causes-SitTypeSitType Poaching Extinction)` and `(enables-Generic EatingEvent Extinction)` receive a score of 0 because Poaching and EatingEvent are not connected by a subcollection relation. This is a very crude measure of relatedness but is useful for evaluating statements that have been identified as relevant by some other means.

5.2.3. Object Properties via Subparts

One heuristic for solving an object property question is to check for the property in known sub part types. If a question has been determined to be an object property question, then there is a specific collection associated with it. The question in Figure 11, for example, is about the collection `CellMembrane`. If the property we seek is known for any sub part types of `CellMembrane` (e.g. `LipidBilayer`), then we can guess that it also holds for `CellMembrane`. If the property we seek is not known for any known sub part types, then this heuristic formulates alternative expressions for the failed sub part query and stores them away as potentially relevant patterns that can be explored if no other reasoning path leads to an answer.

The system answers the question shown in Figure 12 by reasoning about subparts. Rephrased in terms of CycL, this question asks, “All cells have thousands of ?x” which can be represented with the following rule macro predicate statements:

```
(relationAllExistsRange physicalParts Cell ?x Thousands-Quant)
```

This fact is nowhere to be found in the KB, so the system begins by looking at the subparts of Cell. It then executes queries like

```
(relationAllExistsRange physicalParts CellNucleus ?x Thousands-Quant),
(relationAllExistsRange physicalParts Mitochondrion ?x Thousands-Quant),
```

and so on. It still doesn't find an answer, but it stores several possibly relevant patterns in working memory. These patterns substitute answer options into the variable position, move variables, and introduce new ones if the expression is long enough (to avoid a fully open and expensive query). One of the patterns that gets stored is

```
(relationAllExistsRange physicalParts ?x Gene-HereditaryUnit ?y)
```

When queried, this pattern returns

```
(relationAllExistsRange physicalParts CellNucleus Gene-HereditaryUnit Many-Quant)
```

This fact exists in the KB because it was learned from the cell/earth analogy. The part-for-whole substitution is undone, and we now have the fact:

```
(relationAllExistsRange physicalParts Cell Gene-HereditaryUnit Many-Quant)
```

Because `Many-Quant` and `Thousands-Quant` bear some resemblance as computed by the interchangeability algorithm, and all other terms in the statement are equal to one of the answer options, this statement serves as evidence for the first answer option:

```
(relationAllExistsRange physicalParts Cell Gene-HereditaryUnit Thousands-Quant)
```

Object Property Question (#1)

A single human body cell typically contains thousands of

```
(relationAllExistsRange
  physicalParts  Cell
  ?cell-part Thousands-Quant))
```

a) Genes — `Gene-HereditaryUnit`

b) Nuclei — `CellNucleus`

c) Chloroplasts — `Chloroplast`

d) Bacteria — `Bacterium`

Ground Query for Answer Option 1

```
(relationAllExistsRange physicalParts Cell
    Gene-HereditaryUnit Thousands-Quant)
```

Ground Query for sub parts

```
(relationAllExistsRange
    physicalParts CellNucleus
    Gene-HereditaryUnit Thousands-Quant)
(relationAllExistsRange
    physicalParts Mitochondrion
    Gene-HereditaryUnit Thousands-Quant)
...
```

Query variants

```
(relationAllExistsRange
    physicalParts CellNucleus
    Gene-HereditaryUnit ?q)
(relationAllExistsRange
    physicalParts ?whole
    Gene-HereditaryUnit ?q)
...
```

Result (one of many)

```
(relationAllExistsRange
    physicalParts CellNucleus
    Gene-HereditaryUnit Many-Quant)
```

Undo part-for-whole substitution

```
(relationAllExistsRange
    physicalParts Cell
    Gene-HereditaryUnit Many-Quant)
```

Evidence for answer option 1

Figure 12: A subset of the queries that are executed for question 1. Ground queries are first created for all answer options. Since none succeed for this question, the QA system attempts queries with subpart types in place of the original collection. Since those also fail, the QA system attempts query variants, with variables introduced to increase the chances of returning a result. One of the results that is returned is that Nuclei have many genes. This is similar to the ground query for answer option 1, and is therefore used as evidence for that answer.

5.2.4. Object Properties via Base Domain Speculation

Another approach for answering object property questions is to think of the object in terms of a prior analogy. Given an object property question about some collection C and some property P, if there exists a known analogical mapping where C mapped to C' and C' has property P, then the system assumes that C also has that property. This is a risky heuristic and is therefore assigned a high cost so that the solver only attempts this if finding an answer in the original domain fails. If C' does not have property P, then variants of the original query, with C' in place of C, are stored as potentially relevant query patterns that might be useful as a last resort.

For example, the system uses base domain speculation to answer the question shown in Figure 11. The question asks about the function of cell membranes. From a prior analogy (i.e. the cell/city analogy), the system finds that at one point, the collection `CellMembrane` corresponded to the collection `Border`. The system can search the knowledge base for functions of borders to see if there are functions that are similar (i.e. have a high interchangeability score) to any of the answer options. For this question, the ground queries are executed with the border collection replacing the cell membrane collection. These queries do not return any results, so query variants (i.e. with different open variables) are explored. One of the results that is found involves the relationship between borders and border check events:

```
(relationExistsAll preActors BorderCheckEvent Border)
```

When this fact is transformed back into the target domain, it becomes

```
(relationExistsAll preActors BorderCheckEvent CellMembrane)
```

This intermediate fact doesn't make very much sense because the system does not create an analogy skolem for `BorderCheckEvent`. However, this is part of the advantage of base domain speculation. It allows the system to make big inference leaps that are guided by prior analogies. The system proceeds by computing interchangeability scores between all the facts found from query variants and the answer options. The fact involving `BorderCheckEvent` stands out because it is interchangeable with answer option 4:

```
(relationExistsAll doneBy (IdentifyingAsTypeFn ChemicalObject) CellMembrane)

(relationExistsAll preActors BorderCheckEvent CellMembrane)
```

These two facts return a non-zero interchangeability score because `BorderCheckEvent` is a specialization (i.e. subcollection of) `Inspecting` and `Identifying`, which are supercollections of `(IdentifyingAsTypeFn ChemicalObject)`. Additionally, doneBy is a specialization of `preActors`. The rest of the terms are identical, so the interchangeability returns a non-zero value, indicating that the second fact can be used as evidence to support answer option 4. This causes the involvement of borders in border check events to be used as evidence that cell membranes identify chemical objects.

5.2.5. Example Situation Analysis

Example situation questions set up a scenario for the reader and ask something about that scenario. In the question representation used here, that means that the question's microtheory has extra knowledge in it. The way to answer these questions usually requires filling in missing pieces with background knowledge. Cyc's rule macro predicates are very useful for this and are just the types of representations that are learned from the instructional analogy interpretation model.

To apply background knowledge to scenarios, I implemented forward inference methods for rule macro predicates (see examples in Table 1) along with the predicates that represent qualitative proportionalities (i.e. qprop, qprop-). These methods search the local context for instances of universally quantified collections and existentially quantified collections and asserts the binary relation between them. If no existential instance is found and there are no other potential candidates (i.e. instances whose collections are unknown), then a speculative instance is created. In this way, the application of background knowledge enables the creation of new terms that are assumed to be there even if they were not mentioned explicitly. For rule macro predicates, the universal and existential quantifiers are explicit based on the argument position. For qualitative reasoning predicates, I chose to treat them as if the cause was universally quantified and the effect was existentially quantified. For example, given a type level negative qualitative proportionality between the rate of change in an ecosystem (the cause) and the stability of that ecosystem (the effect), then the application method for this predicate will seek instances of changes in an ecosystem. If any are found, then an instance level negative qualitative proportionality between the rate of that change and the stability of the ecosystem are asserted. This is how some of questions related to ecosystems and interdependence are solved.

Example Situation Question (#4)

Human population growth has led to a reduction in the populations of predators throughout natural ecosystems across the United States. Scientists consider the loss of these predators to have a

```
(isa this-ecosystem Ecosystem)
(consumerInEcosystem this-ecosystem these-humans)
(consumerInEcosystem this-ecosystem other-predators)
(eatsWillingly other-predators other-prey)
(isa human-pop-growth IncreaseEvent)
(isa human-pop-growth IntrinsicStateChangeEvent)
(objectOfStateChange human-pop-growth this-ecosystem)
(objectOfStateChange human-pop-growth other-predators)
(isa predator-decline DecreaseEvent)
(isa predator-decline IntrinsicStateChangeEvent)
(objectOfStateChange predator-decline this-ecosystem)
(causes-EventEvent human-pop-growth predator-decline)
```

a) positive effect, because an increase in their prey helps to maintain stability in the ecosystem

```
(and (causes-Event predator-decline ?prey-increase)
(isa ?prey-increase IncreaseEvent)
(positivelyInfluencedBy
       ((QPQuantityFn Stability)
          this-ecosystem)
       (RateFn ?prey-increase)))
```

b) positive effect, because the predators usually eliminate the species they prey on

```
(and (causes-Event predator-decline ?ext)
     (isa ?ext Extinction)
     (positivelyInfluencedBy
       ((QPQuantityFn Stability)
            this-ecosystem)
       (RateFn ?ext)))
```

c) negative effect, because predators have always made up a large portion of our food supply

```
(and (eatsWillingly these-humans ?food-supply)
     (negativelyInfluencedBy
        (AmountOfFn ?food-supply)
        (RateFn predator-decline)))
```

d) negative effect, because predators have an important role in maintaining stable ecosystems

```
(negativelyInfluencedBy
   ((QPQuantityFn Stability) this-ecosystem)
   (RateFn predator-decline))
```

Figure 13: An example question about ecosystems. The right hand column shows the CycL representation of the scenario and the answer options.

For example, one of the questions (

Figure 13) presents an example situation about human population affecting predator populations. The first row of the table shows the scenario as it is described in the question (left) and the CycL representation for that scenario (right). The answer options in English and in CycL are shown below the scenario description. The system answers this question by applying background knowledge into the example situation. Background knowledge is retrieved by searching the KB for rule macro predicates and qualitative proportionalities that involve collections mentioned in the scenario. In this case, the collections mentioned in the scenario are: `Ecosystem`, `IntrinsicStateChangeEvent`, `IncreaseEvent`, and `DecreaseEvent`. One of the qualitative proportionalities found in the KB was captured from the analogy about ecosystems from Chapter 4:

```
(qprop- ((QPQuantityFn Stability) Ecosystem)
        (RateFn (IntrinsicStateChangeOfFn Ecosystem)))
```

This states that the rate of change in an ecosystem negatively influences the stability of that ecosystem. The forward inference methods developed for this system treat qualitative proportionalities as being type-level statements where the cause is universally quantified and the effect is existentially quantified:

```
(relationExistsAll qprop- ((QPQuantityFn Stability) Ecosystem)
                          (RateFn (IntrinsiceStateChangeOfFn Ecosystem))
```

This statement means that for all ecosystem changes, there exists an ecosystem that is negatively affected by it. Note that the rule macro predicate for representing this qualitative proportionality is actually underspecified because it does not say that the change to a *specific* ecosystem negatively affects the stability of *the same* ecosystem. However, the forward inference methods attempt to resolve this issue by hypothesizing which entities (if any) in the local context are the best candidates for the existential variable. To apply this knowledge to the current question scenario, the system first looks for instances of `(IntrinsicStateChangeOfFn Ecosystem)`. The question does not explicitly state that human population growth and the decline of predators are instances of (IntrinsicStateChangeOfFn Ecosystem), but it does say that they are both instances of IntrinsicStateChangeEvent and that the objectOfStateChange is an instance of Ecosystem (Figure 13). The system uses this knowledge and the knowledge of IntrinsicStateChangOfFn as a collection denoting function in the KB to infer that human population growth and the decline of predators are indeed instances of (IntrinsicStateChangeOfFn Ecosystem). Next, the system looks for something in the scenario that might be an instance of ((QPQuantityFn Stability) Ecosystem). It first searches for instances of ((QPQuantityFn Stability) Ecosystem). It finds none, so it proceeds by searching for instances of Ecosystem and finds the one described in this question scenario. Since there is only one ecosystem in this question scenario, the system assumes that this is the entity that the background knowledge should be applied to. If there were multiple instances of ecosystem, the system would use the number of common relations between each instance of Ecosystem and each instance of (IntrinsicStateChangeOfFn Ecosystem) to guess which instances to choose. If no instances of Ecosystem were found, then the system would create a speculative instance of Ecosystem. For this question, the type-level qualitative proportionality from background knowledge is applied to the scenario, which results in the following instance-level statements:

```
(isa human-pop-growth (IntrinsicStateChangeOfFn Ecosystem))
(isa predator-decline (IntrinsicStateChangeOfFn Ecosystem))
(qprop- ((QPQuantityFn Stability) this-ecosystem) (RateFn human-pop-growth))
(qprop- ((QPQuantityFn Stability) this-ecosystem) (RateFn predator-decline)))
```

The resulting statements are asserted into the question scenario. Unlike statements in background knowledge, these statements are about specific entities rather than collections. Now the question answering process continues with more knowledge than it started with. When the system directly queries each of the answer options, option D succeeds because pre-existing rules (from the QP ontology) enable the system to infer `positivelyInfluencedBy` and `negativelyInfluencedBy` statements based on `qprop` and `qprop-` statements, respectively.

5.2.6. Abductive Hypotheses

Abduction is potentially useful for answering object property questions and scenario questions. In both cases, this system looks for known facts that might explain or support any of the answer options. This was implemented by writing rules to capture the requirements for particular functions and the existence of part-whole relationships to explain the collection membership of a particular item (Figure 14).

For function requirements, rules state that part-whole relationships between entities could suggest that the part is the direct object in an action completed by the whole. This was implemented for three classes of functions (based on the types of functions that were observed in the 11 instructional analogies used in Chapter 4 and the question set): making things available, changing things, and taking care of things (e.g. protecting, storing). The rationale is, if some collection C typically has sub parts C', then that can be used as evidence to suggest that instances of C perform actions on instances of C', such as making instances of C' available, taking care of instances of C' and changing instances of C'. For example, batteries *have* energy. Knowing this might suggest that batteries convert energy, make energy available, or store energy. For example, the question in Figure 15 asks about the function of the cell nucleus. From the cell/city analogy, the system knows that one of the functions of the cell nucleus is that it controls the cell. However, this fact is not useful for this particular question. Instead, the system checks if any of the answer options can be assumed via abduction. Options 2 and 3 can be assumed if the things made available or stored are known to be parts of the cell nucleus. From the cell/earth analogy, the system knows that genes tend to be parts of cell nuclei and therefore assumes that option 3 is the correct answer.

For explaining the collection membership of a particular item, one rule was used to state that given some term P that is part of some other term W, and a type-level part-whole relationship between W-TYPE and P-TYPE, then there is an abductive hypothesis that P is an instance of P-TYPE. This is useful for questions that ask about parts of things in a particular situation (e.g. question 13, Table 14). When given a situation where there is an object (e.g. a circuit) with parts that are unlabeled, the system assumes that the unlabeled part belongs to a collection that is known to be a subpart type of the original object (e.g. battery or wire).

Explaining functions:

```
(abductiveHypothesis (relationExistsAll ?role ?fn ?whole-type)
    (and (unifies ?fn (event-fn? ?part-type))
        (or (resultGenl ?event-fn InstrinsicStateChangeEvent)
                (resultGenl ?event-fn MakingSomethingAvailable)
                (resultGenl ?event-fn TakingCareOfSomething))))
        (partTypes-transitive ?whole-type ?part-type))))
```

Explaining collection membership:

```
(abductiveHypothesis (isa ?term ?part-type)
    (and (physicalParts ?whole ?term)
        (isa ?whole ?whole-type)
        (partTypes-transitive ?whole-type ?part-type)))
```

Figure 14: Abductive hypothesis statements used to assume functions and collection membership. These are implemented in horn clauses, with the first abductive hypothesis statement broken into three separate horn clauses to handle the disjunction. In English, the first states that if something has subparts, then a reasonable hypothesis is that it can make the subparts available, change the subparts, or take care of them. For example, knowing have batteries *have* energy suggests that batteries make energy available. The second states that if there is something that has subparts and it is known to have subparts of a particular type, then we can assume that the subparts are of that subpart type. For example, if there is a circuit with a hidden part, and we know that circuits tend to have wires, then a reasonable hypothesis is that the hidden part might be a wire.

Object Property Question (#8)	
Which of the following is a function of the nucleus in organism 2?	`(relationExistsAll performedBy ?fn CellNucleus)`
a) absorbing sunlight	`(SubcollectionOfWithRelationToTypeFn AbsorptionEvent objectActedOn Sunlight)`
b) releasing usable energy	`(MakingAbstractAvailableFn EnergyStuff)`
c) storing genetic material	`(StoringFn Gene-HereditaryUnit)`
d) producing food molecules	`(MakingFn Food)`

Figure 15: An object property question that about the function of the cell nucleus.

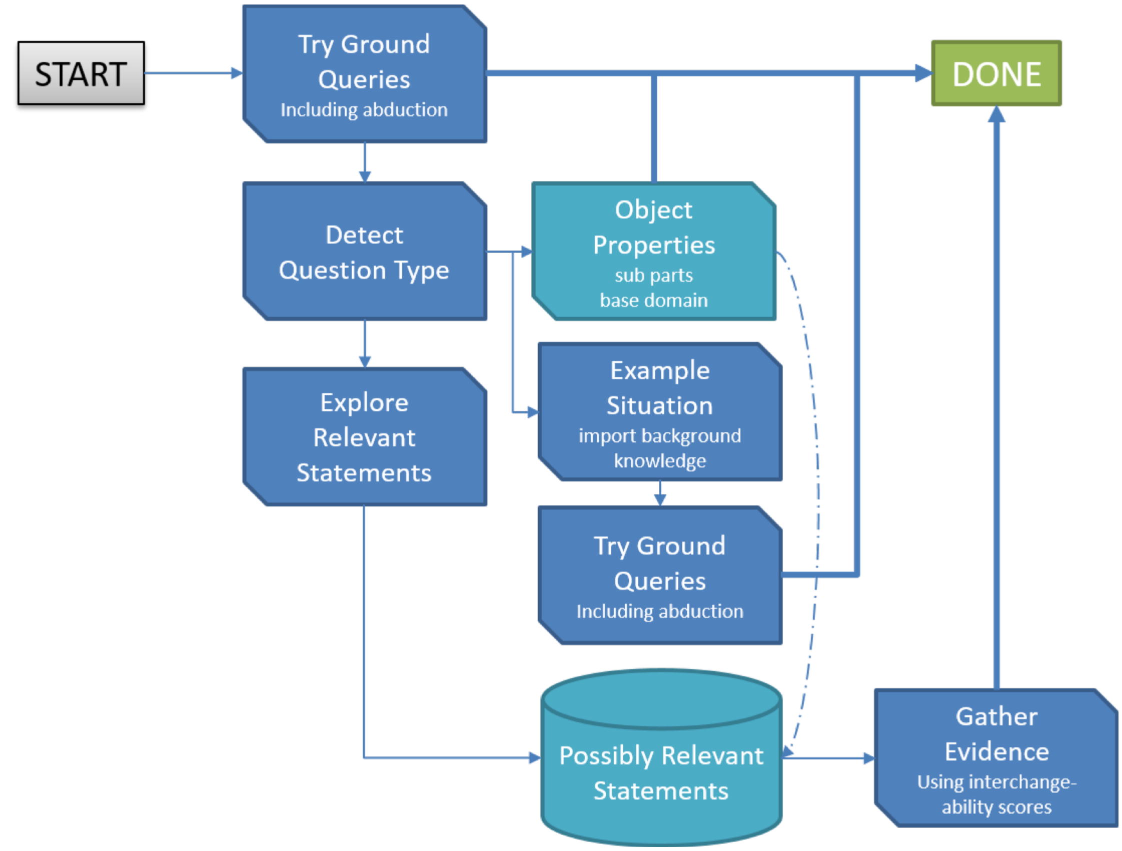


Figure 16: An overview of how multiple choice questions are answered. The first step executes the ground queries directly. If an answer is found (following the bold arrow), then the process is done. If not, it continues with a basic analysis of the question to determine if it is about object properties or an example situation. For each, different strategies are employed. The object property strategy uses subparts and base domain entities to come up with variants of the original query by placing variables in different argument positions. If an answer is not found directly from sub parts or base domain entities, the system examines the query variants and used them to compute evidence for each of the answer options.

5.2.7. Solving Sequence

The sequence for attempting to answer a question is illustrated in Figure 16. The first, most basic approach is to execute the ground queries for the question directly. The ground queries for each question are either a unification of the answer options and the open question query (as would be the case for the question shown in Figure 11), or each answer option if the question is a true question statement (since each answer option is a fully ground statement). This step also includes abduction. So, if the ground query is explicitly known in the KB or if it can be assumed via abduction, then the query will succeed. If any of these queries succeed, then the system returns the answer (following the bold arrow in Figure 16). If not, it proceeds to the next step, which involves detecting the question type.

If a question type is detected, the corresponding strategy is triggered. As described above, these strategies include looking at subpart types and base domain items for object properties, and importing background knowledge for example situation questions. For object properties, if the property in question is found in either a sub part type or in a corresponding base term, then the system has found its answer. If not, it generates variants of the subpart and base domain queries by inserting variables at different argument positions. These non-ground expressions are marked as possibly relevant query patterns, which are not executed, but stored away as a last resort if other reasoning paths fail. For situation questions, background knowledge in the form of rule macro predicates and qualitative proportionalities are imported into the question's reasoning environment. As mentioned earlier, applying background knowledge results in the existence of new statements, and sometimes in the existence of new terms. This means that the ground queries for the question should be attempted again, in case the answer was brought in from the background. If any of the ground queries succeed, the system is done. If not, it continues to the final heuristic, which looks at the possibly relevant query patterns stored away earlier.

The query patterns that have been marked as possibly relevant will include statements about subpart types and corresponding base concepts. It will also include variations of true statement options that have failed and variations of the original ground queries. This is essentially a purely syntactic way of finding statements that may be useful, but are not found by direct queries. For each statement that is returned by the patterns, the interchangeability score is computed between it and all of the answer options. The scores for each answer option are summed as an estimate of the system's confidence in that answer choice. Once all the statements are evaluated, the answer choice with the highest confidence is chosen as the answer. The question answering process fails if the query variants do not return any results or the results bear no resemblance to any of the answer options (i.e. the interchangeability scores are all zero).

5.3 Evaluation

Using the heuristics described above, the knowledge captured from the instructional analogies in Chapter 4 was sufficient for answering 11 out of 14 questions. This included 4 out of 5 questions on ecosystems, 5 out of 6 questions on the cell and its parts, and 2 out of 2 questions on circuits (Table 14).

#	Question Prompt -- Correct Answer	Correct?	Approach
1	A single human body cell typically contains thousands of -- genes	y	object property via sub parts + interchangeability
2	A function of cell membranes is humans is the -- recognition of chemicals	y	object property via base domain + interchangeability
3	Which statement best describes the organelles in a cell? -- they work together	y	direct query + interchangeability
4	Human population growth has led to a reduction in the populations of predators throughout natural ecosystems across the United States. Scientists consider the loss of these predators to have a -- negative effect on stability	y	import background QP knowledge + inference
5	A variety of pear tree, known as Bradford, was originally introduced into the eastern United States in the 1960s. Today, this tree is crowding out other plants in these states. This situation best illustrates -- negative effect on stability	y	import background QP knowledge + inference
6	If the producers in a food web were removed, which of the following changes would most likely occur? – collapse because ecosystems require producers	y	direct query
7	Within a prey population, which of the following is most immediately affected by the arrival of a new predator? -- population of predator influences death rate	n	import background QP knowledge, but didn't have QP relationship impacting population
8	Which of the following is a function of the nucleus in organism 2? -- storing genetic material	y	abduction
9	In the three organisms, what are synthesized by the ribosomes? -- protein	y	direct query + interchangeability
10	Which of the following best describes the producers in a terrestrial food web? -- producers convert solar energy to chemical energy	y	direct query
11	Which of the following is the best example of the human body maintaining homeostasis? -- breathing during exercise	n	import background knowledge, but didn't have a way of inferring process type
12	In a sample of double-stranded DNA, 30% of the nitrogenous bases are thymine. What percentage of the nitrogenous bases in the sample are adenine? -- 0.3	n	import background knowledge, but no explicit equals relation
13	Jamal wants to make an electrical circuit, but he only has the objects shown below. Which of the following must Jamal also have to make an electrical circuit? -- battery	y	abduction
14	The diagram below shows a project that a student made to test an electrical circuit. Part of the electrical circuit is underneath the board. When the student connects the two nails using a wire, the bulb lights up. Which of the following must be underneath the board? -- battery and wires	y	abduction

Table 14: Questions used to evaluate qualitative science knowledge.

A variety of strategies were used to arrive at answers. In some cases, direct queries sufficed. Question 10, for example, is a true statement question where one of the options described the function of producers: that they convert solar energy to chemical energy. In CycL, this is formulated as a rule macro predicate using a primary actor role relation:

```
(relationExistsAll
   doneBy
   (SubcollectionOfWithRelationToTypeFn
    (SubcollectionOfWithRelationToTypeFn
          IntrinsicStateChangeEvent toState ChemicalEnergy)
    objectOfStateChange SolarEnergy)
   Autotroph)
```

This fact was captured from by interpreting the ecosystem analogy from Chapter 4, so a direct query was sufficient for answering the question.

Abduction was very useful for scenario questions where there was a missing or unlabeled entity and the question asked for a collection for that entity (questions 13, 14). The second abduction rule in Figure 14 was responsible for solving these questions. Since batteries and wires are known to be parts of circuits, it was assumed that the answer options that mentioned those parts were correct. However, the background knowledge was not definitive in terms of *how many* of each part type a circuit had, so it could have incorrectly answered another battery. But, since items that were already in the scenario were not also answer options, the system was able to correctly choose the right answer. Clearly, the background knowledge is not as deep as it should be, but for these questions, it is sufficient. Essentially, the question is asking the reader: what are the required parts of this object? Whatever is missing is the answer.

Background knowledge was useful for two of the questions about ecosystems and interdependence. The ecosystem/web analogy in Chapter 4 was interpreted using qualitative modeling of an ad hoc student web (i.e. a situation where students gather around, connected by strings to mimic an ecosystem). The primary inference is that changes within the web decrease stability. The causal relationship between ecosystem changes, and the stability of the ecosystem was captured as a result of interpreting this analogy. When presented with a scenario question about an ecosystem and some change happening within it (i.e. questions 4, 5, 7), the system infers that the change is negatively influencing ecosystem stability. This relationship is very simple, yet it is sufficient for answering questions 4 and 5.

As noted earlier, object property questions could be solved by looking at subpart types and by speculating about known base domains. Detecting object properties via subparts was used to answer question 1 (Figure 12) and detecting object properties via base domain speculation was used to answer question 2 (Figure 11).

5.4 Discussion

The experiment in this chapter supports claim 2 of this thesis. It demonstrates that qualitative science knowledge captured from multimodal instructional analogies can be used in combination with very basic question answering strategies to answer questions on middle school science exams. The question answering strategies developed for this experiment along with background knowledge and ontologies within Cyc, were used to answering 11 out of 14 questions. The impact of the specific content learned from the instructional analogies in Chapter 4 can be evaluated by making that content unavailable during question answering. After removing the knowledge learned from those analogies (but leaving in all other biology and electricity knowledge in ResearchCyc), the QA system was unable to answer any of the 14 questions examined here. This demonstrates that while the ontologies in Cyc are extremely useful and necessary, they need to be complemented with general, non-technical statements about the domain, like the ones that were captured from instructional analogies. They also need to be complemented with QA strategies like the ones presented in this chapter.

There are, however, several limitations to this claim. The first has to do with coverage. Out of the 86 test questions that were found to involve human biology or electricity, only 14 had answers that overlapped with the target knowledge of the analogies. Around 16% coverage is modest, but also not surprising when one considers all the other topics that appear on these exams. In the full set of 86 questions on human biology and electrical energy, there were many questions on genetics, causal reasoning about experiments, specific biological taxa, class definitions for specific types of cells, properties of everyday objects (e.g. conductance) and many other topics not addressed by any of the analogies in the FAR guide. However, it's not the case that the FAR guide topics are rare. They include very basic fundamental ideas in biology and electricity, it is just that it would be very unusual for 11 analogies to cover a very large set of questions at all the levels of detail that they are asked. The analogies address basic conceptual knowledge about several topics with the expectation that more detailed knowledge is added later, which is not currently done in this interpretation model. The analogies do not dive into technical detail and they assume that the learner has a significant amount of commonsense knowledge. The commonsense knowledge assumption is fine for people, but a big problem for intelligent systems. In short, there are topics on both ends of the technical spectrum (from basic commonsense knowledge to technical knowledge of specific processes) that are missed by these instructional analogies.

The second limitation has to do with how qualitative science knowledge is actually used. If it cannot be deployed properly, then it's not worth very much. The three questions that the QA system could not answer point to some of the reasoning abilities that are lacking in this system, and are taken for granted when people use qualitative knowledge to answer questions. Question 7 was a situation question, and asked what is immediately affected by the arrival of a new predator. Based on the knowledge gained from the ecosystem/web analogy, there is information in the KB about producers, consumers, that predators eat prey, and that changes in the ecosystem negatively impact stability. What's missing the very basic causal information between eating, death, and population. Background knowledge was imported correctly, but there was no causal knowledge about the impact predators have on the population of their prey. A more sophisticated model of the dynamics in an ecosystem would be needed to capture that. Those models would also need to be tied with our most general models of how the world works in order to be useful. Question 11, for example, provides four situations and asks which one is the best example of homeostasis. Answering this question requires an understanding of the notion of competing influences or feedback systems. The simple forward inference method implemented in this QA system is insufficient for this question because instead of finding quantities and asserting qualitative relationships about them, this question requires detecting relationships and asserting the process or situation type. This is akin to model formulation in general, something that was used for interpreting analogies but not for answering questions about them. Lastly, question 12 was about the structure of an example DNA molecule. From the DNA/clothespin analogy, the system knew that adenine always bonds with thymine and that guanine always bonds with cytosine. The connection between that knowledge and the fact that in a given molecule the number of adenine molecules should be *equal* to the number of thymine molecules was missing. These questions illustrate the rich interaction between qualitative science knowledge and qualitative commonsense knowledge that is needed to pass this test.

Another limitation to this claim is the absence of natural language understanding in question interpretation. To limit the scope of this thesis, I chose to manually translate the questions from English to CycL. This proved to be very difficult because there are numerous ways to represent things using CycL, and very often, the statements defined to represent knowledge are underspecified or not completely accurate (Fan et al., 2009). However, as seen in the analysis of analogies in the FAR guide, sometimes the loose use of language is a feature not a bug. Inanimate things or processes are sometimes described using language that would suggest they are living agents. For instance, organelles don't really engage in *coordinated group activities* in the sense that Cyc defines that type of activity. This is by design because the analogy uses a level of abstraction that is amenable to cross-domain mappings. Using this level of abstraction, while at the same time adhering to the representational constraints (e.g. predicate argument type constraints) within Cyc, remains a challenge.

Chapter 6: Related Work

Analogical reasoning in AI systems has a rich history with many different applications. Researchers often differ in how they define and use analogical reasoning, with some developing methods for solving traditional analogy word problems (i.e. A:B::C:D), others using it to perform case-based reasoning, and others who use it as a general measure of relational similarity.

6.1 Problem Solving and Question Answering

Most analogical problem solvers work by transferring prior knowledge to a new scenario. Early work by Burstein (1985) analyzed human tutoring protocols to propose a model for learning by analogy. This model was implemented in a system that used stored planning knowledge to incrementally generate hypotheses for how to solve computer programming (i.e. variable assignment) problems. The PRODIGY planning architecture used a derivational analogy engine (Veloso & Carbonell, 1993) to learn control knowledge for future situations. When new planning problems arose, PRODIGY/ANALOGY used prior similar problems to limit the search for new solutions. The general approach of using prior experiences or examples to solve or plan in new situations is typical in case-based reasoning systems (Kolodner, 2006).

A slightly different approach to analogical problem solving is to use prior knowledge to construct new knowledge that is abstracted away from any one scenario. In PHINEAS (Falkenhainer, 1990), new domain theories were constructed by analogy to previous observations of known phenomena. Given a behavior of some phenomenon from an unknown domain, PHINEAS retrieved an analogous behavior from a known domain. By constructing an analogy between the two behaviors, the system was able to detect entity correspondences, and transfer model fragments from the known domain to the unknown domain. As a result, the analogy not only provided useful information for the current scenario, but also abstract information in the form of model fragments that could be applied to new situations in the previously unknown domain. A similar cross-domain approach was used by Klenk's domain transfer via analogy (DTA) system (Klenk & Forbus, 2013). In DTA, analogies were constructed between worked solutions (i.e. problems with the solution steps explicitly identified) to create a mapping between two different domains, e.g. rotational kinematics and translational kinematics. Given a worked solution to a problem in one domain and a library of worked solutions from another, DTA used analogical retrieval to find an analogous worked solution from library. The entity correspondences between the two worked solutions were used to create a mapping between the two different domains. Thus, equation schemas from the one domain could be projected to the other, and vice versa. DTA is very similar to PHINEAS in the sense that observations from known and unknown domains are compared to arrive at more general between-domain mappings.

More recently, models of analogy have been used to aid in open-domain question answering and traditional analogy word problems. IBM's Watson used structure-mapping between individual questions and potentially useful text passages as a measure of semantic similarity (Murdock, 2011). When structure-mapping was used on more abstract representations than the individual lexical items within a given question, Watson's ability to identify text passages that provided evidence for particular answers improved (Murdock, 2011). Boteanu and Chernova (2015) developed an approach for answering traditional analogy word problems (e.g. A:B::C:D) that used concepts in semantic networks. Given two concepts in the network, their similarity was computed based on relational pathways between them. The similarity of those pathways was the proportion of common relations. These relational pathways were also used to generate explanations for any answers generated, which often contained salient relational similarities. In a domain-specific, but more complex problem topic, Chaudhri et al. (2014) used a large scale knowledge base to answer compare and contrast problems between topics in biology. Analogical mappings between topics were used to generate tables where information in both topics was aligned, in order to highlight important similarities and differences. In each of these systems, the notion of analogical reasoning is implemented in different ways, but the purpose in each is to either provide some measure of similarity between two descriptions or to identify meaningful similarities and differences.

6.2 Knowledge Capture

The most direct application of analogical reasoning to knowledge capture requires understanding explicit analogies from natural language. Barbella and Forbus' (2011) analogical dialogue acts model provides a framework for translating utterances into directives for building analogical mappings. Their model has been integrated into EANLU and was used in the interpretation model in Chapter 4.

Analogy has also been combined with crowdsourcing as a means to capture common sense knowledge (Chklovski, 2003). Chklovski's system, Learner, uses multiple analogies over concepts that have been previously entered by human volunteers to pose questions about new topics that are chosen by the human volunteer. Mappings between the new topic and previously entered topics are used to identify properties that are known about similar topics but not known for the current own. Rather than transferring the value of the property to the new topic, the property is formulated as a question, which is posed to the human volunteer. In terms of structure-mapping theory, this is akin to using crowdsourcing to accept, reject, or clarify candidate inferences between topics. Speer et al. (2008) used a non-structural approach to analogical reasoning and applied it to the ConceptNet semantic network. They use singular value decomposition to reduce knowledge from ConceptNet along key feature dimensions. Feature overlap can be easily computed via dot product, but this comes at the loss of structured relational information.

In the area of multimodal reasoning, Ferguson's JUXTA system (Ferguson & Forbus, 1998) was an early approach for interpreting and critiquing diagrams with juxtaposed scenarios. In JUXTA, analogy was used to detect differences in the spatial representation of the diagram, and label them based on the diagram's caption. The multimodal knowledge capture system (MMKCAP) from Lockwood and Forbus (2009) used structure-mapping to align representations from different modalities. Lockwood's system read simplified English versions of passages and sketched versions of diagrams from a physics textbook on basic machines. Given that passages typically refer to their accompanying diagrams and vice versa, structure-mapping was used to align information from the passages with information from the diagram.

The analogical mapping indicated how corresponding items could co-refer and therefore provided a template for how both information sources could be merged into an integrated case. Integrated cases for basic machines from the first chapter were captured using this approach. The resulting cases were then used to answer questions from that chapter. The multimodal integrated approach in Chapter 4 extends Lockwood's work because it uses automatic language disambiguation and event interpretations to add structural similarity between the sketch and text interpretations.

6.3 Tutoring Systems

In comparison to analogical reasoning in general, there have been relatively few attempts at incorporating analogical reasoning into intelligent educational software. For teaching conceptual knowledge, the use of analogies has been explored in three systems. Murray et al. (1990) designed a tutoring system based on the idea of bridging analogies (Clement, 1993), which are designed to gradually change misconceptions by comparing a misunderstood scenario with a well-understood, intuitive scenario, called an anchor. If the student cannot remedy the misconception by comparison to the anchor, then a bridge scenario is summoned to facilitate inference projection from the anchor to the bridge and eventually to the initially misunderstood scenario. In this system, the scenarios which were used for instruction were determined empirically. They were organized in a conceptual network, such that at the end of each comparison, the system used the structure of the network to determine which analog should be used next. Each comparison activity consisted of a forced choice question and a confidence estimate from the student. If the student answered the question incorrectly, an analogy between the initial situation and an intuitive anchor was invoked. When the student answers the question correctly, analogs are selected along the graph going closer to the original scenario, thereby gradually encouraging the student to project inferences from the anchor to the originally misunderstood situation. Thus, the main instructional strategy of this tutor is in its selection of analogs as determined by the scenario network. There is no fine-grained feedback during the comparison process. By contrast, another analogy tutoring system used hierarchical plan operators to provide domain-specific, fine-grained feedback for analogies about the circulatory system (Lulis et al., 2004; Lulis, 2005). These plans were modeled after analyses of tutoring dialogues between students and expert tutors (Lulis & Evens, 2003) and each analogy required different entry-level plans. This means that although the system was able to provide detailed feedback in natural language, the teaching strategies were not domain general. More recently, Alizadeh et al. (2015) conducted a study that characterized computer science (i.e. data structures) tutoring dialogues and Harsley et al. (2016) used those corpus analyses to incorporate analogies into the ChiQat tutoring system. This is a promising area of research, although all tutoring systems to date that have incorporated analogies focus on one domain, usually for a restricted set of topics, and are not able to handle novel analogies.

Given the importance of analogies for case-based reasoning, a natural extension is to build a case-based tutoring system. This is the rationale behind Sketch Worksheets (Yin et al., 2010), which is a sketch-based tutoring system that provides on-demand feedback on sketches. Each sketch worksheet assignment has a pre-defined solution associated with it, so that a student can get advice on their sketch with respect to that solution. Structure-mapping is used to align and compare the student's sketch with the pre-defined solution. Differences identified by the analogy are used to determine what feedback should be given to the student. Note that unlike explicit analogy tutors (e.g. CIRCSIIM, ChiQat), Sketch Worksheets does not use analogy as an explicit instructional tool. Instead, it uses it as a means to evaluate a new sketch in terms of a familiar one to provide automatic feedback.

6.4 Summary

Analogical reasoning has been used in a variety of ways in AI systems. Case-based reasoning systems primarily use analogy as a means to transfer knowledge from prior situations to new contexts for planning or problem solving. Some problem solving systems, like PHINEAS and DTA, have explored analogy as a way to transfer knowledge between domains to generate abstract knowledge (or schemas) that can then be applied to new situations in the same domain. Knowledge capture systems range from language systems that try to interpret explicit analogies, to feature-based methods of reducing semantic network size and complexity. Tutoring systems have been proposed to use analogy as a means for explaining instructional content as well as interpreting student input. In all three areas, analogy is used as a method for aligning complex representations, measuring overall similarity, and transferring knowledge from one topic to another.

Currently no systems use multimodal instructional analogies for knowledge capture or problem solving. Only Lockwood's MMKCAP and Klenk's DTA used a combination of analogical reasoning and spatial reasoning. DTA explored cross-domain analogies to generate equation schemas, so it did not deal with the type of instructional analogies that are used to teach novices about basic science concepts. The focus of MMKCAP was also not on instructional analogies, but rather on how to align representations from different modalities. Most of the question answering systems used analogy as a way of assessing within-domain similarity (e.g. as in Watson) or identifying and organizing similarities and differences (e.g. as in Inquire, Chaudhri et al. (2014)). Using analogical reasoning with common sense domains to answer questions in a technical domain has not yet been explored.

Building a system that can interpret multimodal instructional analogies is novel and useful. Novelty is demonstrated by the lack of AI systems with the ability to build knowledge from the types of analogies that are used to teach beginning science topics. In terms of the content expressed in analogies, the model that is most closely related to this thesis is Barbella and Forbus' (2011) model of analogical dialogue acts. This thesis extends that work by using multimodal instructional analogies (rather than just text passages) to build qualitative knowledge about several topics. This thesis also builds upon the MMKCAP system (Lockwood & Forbus, 2009) by using analogy to combine language and visual information, but unlike MMKCAP, circumvents the need for manual language disambiguation. The resulting model of multimodal analogy interpretation opens the door to new kinds of knowledge capture – via reading or via communication with a human collaborator – and to new kinds of intelligent tutoring systems that use multimodal instructional analogies to teach basic science knowledge.

Chapter 7: Conclusion

The analyses in this thesis address the following questions about instructional analogies for teaching middle school science:

What are these analogies about and how are they used?

What are the reasoning requirements for interpreting them?

Do they provide knowledge that is useful to an intelligent system?

The analysis in Chapter 3 illustrates that analogies can be used for a wide range of topics in introductory science, including topics in biology, chemistry, earth/space science, and physics. These analogies appeal to commonsense notions of change, behavior, possession (including part-whole relationships), and relative size. Most of the analogies in the FAR guide focus on functions and behaviors of things and their parts. Other analogies convey information about relative size or magnitude, which is helpful for communicating ideas about very large or very small scales. The overwhelming majority of analogies use spatial representations or visual input of some kind. The analogies use base domains that most people are familiar with, like human activities in a city and unlocking locks. Interestingly, some analogies create ad hoc base domains, often in the form of physical demonstrations, to convey a new topic. For example, walking up an escalator that is moving downward is a very specific source of knowledge for learning homeostasis. So is a web of students connected by string to teach about ecosystems. Students are much less likely to have real experiences with ad hoc base domains, but they can still be useful because they ultimately rely on students' understanding of the world. Spatial representations such as diagrams and pictures or general visual input from physical demonstrations or role-playing activities are also used as tools to recruit that background knowledge. In short, instructional analogies are used for a variety of topics and they almost always involve some kind of visual stimulation. They differ in what similarities matter (i.e. function and/or structure), and in how much background knowledge is needed.

Interpreting instructional analogies requires natural language understanding, significant background knowledge about the physical world, and spatial reasoning. The goal of the experiments in Chapter 4 was to build a model of interpretation that was able to capture qualitative science knowledge about new domains. The model consisted of natural language understanding (of simplified English), sketch understanding (albeit with conceptual labels), and structure-mapping that was used to create multimodal representations and to transfer knowledge from the base domain to the target domain. Background knowledge and explicit base domain knowledge made contributions to the overall performance of the model, but the presence of sketched input had the greatest impact. This is due to the sketch representation's critical role in natural language disambiguation. However, spatial information also plays a role in overall alignment, especially when there are many containment spatial relations and/or many arrows to trigger visual conceptual reasoning.

Lastly, the system presented in Chapter 5 demonstrated that the qualitative science knowledge captured in Chapter 4 was useful for answering questions from middle school science exams. Consistent with previous analyses of middle school science exams (Clark et al., 2013) many questions required type-level knowledge (i.e. object type properties) and the ability to project background knowledge into example scenarios. Of the 14 questions that were examined, only one involved quantitative values (Table 14, question 12). Even still, this question did not require quantitative calculations or a numerical simulation. It simply required a type level equality relationship (although this was one problem the QA system did not solve). All the other questions required knowledge of the general parts and functions of classes of objects or the ability to import qualitative causal information into an example situation.

7.1 Claims Revisited

Claim 1: Structure-mapping can be used to build qualitative knowledge from multimodal instructional analogies.

1) Multimodal instructional analogies use visual representations to facilitate interpretation.
2) Instructional analogies use background knowledge to facilitate interpretation.
3) Multimodal integration and analogy interpretation can be achieved using structure-mapping.

Claim 1 was supported by the experiments in Chapter 4, which showed that structure mapping could be used both as a way to integrate multimodal information and as a way to build cross-domain analogies such that the interpretation resulted in the intended target knowledge. Ablation experiments illustrated that the use of background knowledge, explicit base-domain knowledge, and especially visual-conceptual knowledge from the sketch were all substantial contributors to the model's performance.

It is important to note, however, that the evidence gathered in Chapter 4 cannot be taken as a strong reflection of the relative importance of background knowledge and spatial reasoning in people.

As mentioned in Chapter 4, the interpretation model does not completely account for why background and explicit base domain knowledge is so important for people. The value of commonsense domains is that they are richly connected to other experiences and conceptual knowledge. The base knowledge that the interpretation model uses is connected to ontologies in Cyc and explicit base domain knowledge from the analogy, but these connections are nothing compared to the interconnectedness of human knowledge. Without comparable interconnectedness and flexibility of background knowledge, I do not expect that the importance of base domain transfer can be captured accurately by this interpretation model. Modeling background and commonsense knowledge is an entire research area in itself, so it is not surprising that shortcomings in that area negatively impact the model's fidelity to human reasoning.

Claim 2: Qualitative knowledge captured via multimodal instructional analogies can be used to answer questions.

The QA system described in Chapter 5 was developed to test this claim. Not surprisingly, the breadth of topics in the biology and electrical energy questions in the exams was not fully covered by the instructional analogies from Chapter 4. However, for the set of 14 exam questions where the answers overlapped at least partially with the intended target knowledge of the analogies, the QA system was able to correctly answer 11 questions. During question answering, the system had access to the knowledge learned from analogies and all of the human physiology and universal vocabulary information in the Cyc KB. The qualitative science knowledge learned from the analogies played a critical role in question answering, since the system could not answer any of the questions when it did not use knowledge captured from the instructional analogies in Chapter 4. This illustrates the importance of type-level, qualitative knowledge in answering these questions. Since many of the questions involved example situations, this result also illustrates that the type-level knowledge can be deployed into instance-level descriptions, indicating its utility for reasoning about novel situations.

7.2 Open Questions and Future Work

One of the biggest challenges to interpreting multimodal instructional analogies and question answering is natural language understanding. Some of the challenges encountered in the construction of this interpretation model include: disambiguation, reference resolution, and generics. For disambiguation, conceptually labeled sketches were a crucial. One can argue that this simply side-steps the disambiguation issue by being heavily influenced by a manually disambiguated sketch. Better disambiguation heuristics are needed to reduce the model's dependence on the conceptual labels form the sketch. Additionally, instructional analogies can be presented without spatial representations, so language-only disambiguation is still an unfulfilled requirement. Experiments with narrative functions (McFate et al., 2014) and analogical word sense disambiguation (Barbella & Forbus, 2013) could help improve this process. As noted in the analysis of the FAR guide analogies, choosing the right level of abstraction for representing functions and behaviors is critical. Alternatively, it could be better to avoid irreversible semantic interpretation choices altogether. An informative experiment would be to attempt to construct a cross-domain analogy without making semantic interpretation choices, and using commonalities to provide evidence for particular semantic interpretations. If goal inference or correspondence information were available (e.g. from analogical dialogue acts), those could guide the interpretation choices even further. For reference resolution, challenges arose when the same object was referred to differently. Creating deterministic rules to handle these cases is difficult. In some cases, they should corefer, e.g. rechargeable battery and charged battery. In other cases, they should be distinct, e.g. rechargeable battery and drained battery. The importance of the distinction depends highly on context. In the case of the ATP/battery analogy, the difference between a rechargeable battery and a charged one can be ignored, while the difference between a rechargeable battery and a drained one cannot. Another issue is the problem of generics. The interpretation model for this thesis takes an aggressive approach to generics, assuming that most relations mentioned in an analogy can be generalized somehow, over the collections of primary actors in an event, over the collections of wholes in part-whole relationships, or over the causes in qualitative causal influences. This happens to work well in the context of instructional analogies for recall, but less so for precision. As would be expected of a novel cross-domain analogy, the system overproduces and is very liberal when it comes to accepting inferences. This is useful for initializing a target domain if knowledge refinement is postponed. Having overly general (or underspecified) type-level statements is very useful because it means that the knowledge is applicable in a wide range of situations. However, it also increases the likelihood of creating false knowledge. Maintaining a balance between overly general and overly specific statements is an open problem. A better understanding of how to interpret generics more precisely in the context of instructional analogies would greatly improve interpretation. This model also needs to be evaluated on a larger set of analogies, both in new domains to broaden coverage, but also on existing topics to explore how multiple analogies would help refine overproduced knowledge.

Another potential area for future work is interactivity for incremental natural language understanding and incremental analogical matching. The interpretation model built for this thesis is fairly linear. An alternative approach would be to have a human collaborator in the loop for at least some parts of knowledge capture. In some cases of language interpretation, there are competing choices that are equally valid, but one less favorable due to the context of the analogy. If the interpretation model had the ability to pose questions to a human collaborator, it could engage in active learning and possibly generate models for what levels of abstraction are favorable for instructional analogies. A similar approach could be used for multimodal integration, but from early experiments, it is not clear that it would be beneficial. Early on, I experimented with an incremental version of multimodal integration, where instead of fully interpreting the text and merging it with the sketch, I began with the sketch information and incrementally added knowledge from each sentence, merging entities for each addition. For simple situations, this addressed issues in coreference as well: if there was one nucleus in the sketch, and “the nucleus” was mentioned in three different sentences, each time a sentence was added, the nuclei would merge and the final representation would have one nucleus. This ignored the nuance of reference resolution, however, as it also would merge events that had different actors. For example, consider these two sentences from the cell/city analogy:

The power station provides electricity. The mitochondrion provides chemical energy.

Interpreting the first sentence would result in an instance of `MakingSomethingAvailable`, where the power station is making electricity available. The second sentence would result in another instance of `MakingSomethingAvailable` where the mitochondrion makes chemical energy available. A simple incremental merging approach would result in these two event instances being merged, even though they involve different actors. On the other hand, this incremental approach may be useful for cross-domain analogies, so that inferences could be evaluated incrementally. The model currently uses incremental matching to incorporate inferences and extend cross-domain matches. Having a human, or some self-evaluation step, in the loop for this process could greatly improve the matching because it would prevent incorrect inferences from extending the match in an unintended way. Consequently, overproduction would be reduced, improving the overall quality of the final target representations. Exploring techniques like these, or others that make the interpretation model more incremental could be beneficial to the model’s transparency and flexibility.

More investigations are also needed to assess the generality of this approach and of the characterization of instructional analogies in Chapter 3. Analogies are used in many other domains beyond the ones explored here. For example, the analogies analyzed by Richland et al. (2007) involved procedural knowledge for solving mathematics problems. Using the categories identified by Alfieri et al. (2013), the 11 analogies explored here all involved conceptual knowledge (rather than strictly procedural or perceptual knowledge). How well the interpretations strategies identified here work with other types of analogies, i.e. procedural and perceptual, is still an open question. Larger corpus analyses could reveal how far the visual conceptual relations described in 4.2.2.1 cover spatial representations used in instructional analogies. Similarly, such corpus analyses could be used to expand the set of very abstract terms (e.g. Event, PartiallyTangible, Individual) and determine if it is even possible to arrive at a set of abstract terms that can be used for a very wide range of analogy topics.

This work leads to an exciting intersection of analogical reasoning, spatial reasoning, and intelligent tutoring. One of the ways in which multimodal reasoning could be useful is for multimodal instruction. Intelligent systems that have the ability to understand visual representations paired with text are needed to capture knowledge from reading and from interacting with other people. Such systems would be endowed with the types of qualitative knowledge that would improve their ability to understand the world and to communicate with people. One of the ways in which instructional analogies are powerful is through Socratic tutoring dialogues (for examples of analogies used this way, see (Alizadeh et al., 2015; Lulis & Evens, 2003)). As mentioned in related work, a few researchers have worked on tutoring systems that use analogies as an explicit instructional tool. But, those systems are mostly scripted from corpus analyses of human tutoring dialogues. This is an important first step, but these systems cannot interpret novel analogies or propose new ones. An intelligent system that could engage in rich dialogues with a student and support novel analogies in multiple domains would be a huge advance in the state of the art in intelligent tutoring systems. This cannot be achieved without improvements to general purpose natural language understanding and without general purpose cognitive models of analogical reasoning like structure-mapping.

REFERENCES

Alfieri, L., Nokes-Malach, T. J., & Schunn, C. D. (2013). Learning Through Case Comparisons: A Meta-Analytic Review. *Educational Psychologist, 48*(2), 87-113. doi:10.1080/00461520.2013.775712

Alizadeh, M., Di Eugenio, B., Harsley, R., Green, N., Fossati, D., & AlZoubi, O. (2015). A Study of Analogy in Computer Science Tutorial Dialogues. *Proceedings of the 7th International Conference on Computer Supported Education, 53*(19.2), 6.6.

Barbella, D., & Forbus, K. (2011). *Analogical Dialogue Acts: Supporting Learning by Reading Analogies in Instructional Texts.* Paper presented at the Twenty-Fifth AAAI Conference on Artificial Intelligence, San Francisco, CA.

Barbella, D., & Forbus, K. D. (2013). Analogical Word Sense Disambiguation. *Advances in Cognitive Systems, 2*, 297-315.

Blanchette, I., & Dunbar, K. (2002). Representational change and analogy: How analogical inferences alter target representations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*(4), 672.

Boteanu, A., & Chernova, S. (2015). *Solving and Explaining Analogy Questions Using Semantic Networks.* Paper presented at the AAAI.

Burstein, M. H. (1985). *Learning by Reasoning from Multiple Analogies.* Yale University.

Chang, M. D., Wetzel, J., & Forbus, K. D. (2014). *Spatial Reasoning in Comparative Analyses of Physics Diagrams.* Paper presented at the International Conference on Spatial Cognition, Bremen, Germany.

Chaudhri, V. K., Heymans, S. J., Overholtzer, A., Spaulding, A., & Wessel, M. (2014). *Large-Scale Analogical Reasoning.* Paper presented at the Twenty-Eighth AAAI Conference on Artificial Intelligence.

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and Representation of Physics Problems by Experts and Novices. *Cognitive science, 5*(2), 121-152. doi:DOI 10.1207/s15516709cog0502_2

Chklovski, T. (2003). *Learner: a system for acquiring commonsense knowledge by analogy.* Paper presented at the Proceedings of the 2nd international conference on Knowledge capture.

Clark, P. (2015). *Elementary School Science and Math Tests as a Driver for AI: Take the Aristo Challenge!* Paper presented at the AAAI.

Clark, P., Harrison, P., & Balasubramanian, N. (2013). *A study of the knowledge base requirements for passing an elementary science test.* Paper presented at the Proceedings of the 2013 workshop on Automated knowledge base construction.

Clement, C. A., & Gentner, D. (1991). Systematicity as a selection constraint in analogical mapping. *Cognitive science, 15*(1), 89-132.

Clement, J. (1993). Using bridging analogies and anchoring intuitions to deal with students' preconceptions in physics. *Journal of research in science teaching, 30*(10), 1241-1257.

Cohn, A. G., Bennett, B., Gooday, J., & Gotts, N. M. (1997). Qualitative spatial representation and reasoning with the region connection calculus. *Geoinformatica, 1*(3), 275-316.

Falkenhainer, B. (1990). A Unified Approach to Explanation and Theory Formation. In J. Shrager & P. Langley (Eds.), *Computational models of scientific discovery and theory formation* (pp. 157-196). Los Altos, CA: Morgan Kaufman.

Falkenhainer, B., Forbus, K. D., & Gentner, D. (1989). The structure-mapping engine: Algorithm and examples. *Artificial intelligence, 41*(1), 1-63.

Fan, J., Barker, K., & Porter, B. (2009). Automatic interpretation of loosely encoded input. *Artificial intelligence, 173*(2), 197-220.

Ferguson, R. W., & Forbus, K. D. (1998). Telling juxtapositions: Using repetition and alignable difference in diagram understanding. *Advances in analogy research*, 109-117.

Ferrucci, D., Brown, E., Chu-Carroll, J., Fan, J., Gondek, D., Kalyanpur, A. A., . . . Welty, C. (2010). Building Watson: An Overview of the DeepQA Project. *AI magazine, 31*(3), 59-79.

Forbus, K., Usher, J., Lovett, A., Lockwood, K., & Wetzel, J. (2011). Cogsketch: Sketch understanding for cognitive science research and for education. *Topics in Cognitive Science, 3*(4), 648-666.

Forbus, K. D. (1984). Qualitative process theory. *Artificial intelligence, 24*(1), 85-168.

Forbus, K. D., Gentner, D., & Ferguson, R. W. (1994). Incremental Structure-Mapping. *Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society*, 313-318.

Forbus, K. D., Klenk, M., & Hinrichs, T. (2009). Companion cognitive systems: Design goals and lessons learned so far. *IEEE Intelligent systems, 24*(4), 36-46.

Forbus, K. D., Usher, J., & Tomai, E. (2005). *Analogical learning of visual/conceptual relationships in sketches.* Paper presented at the Proceedings of the National Conference on Artificial Intelligence.

Friedman, S. E., & Forbus, K. D. (2011). *Repairing incorrect knowledge with model formulation and metareasoning.* Paper presented at the Proceedings of the Twenty-Second international joint conference on Artificial Intelligence.

Gadgil, S., Nokes-Malach, T. J., & Chi, M. T. (2012). Effectiveness of holistic mental model confrontation in driving conceptual change. *Learning and Instruction, 22*(1), 47-61.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive science, 7*(2), 155-170.

Gentner, D. (2010). Bootstrapping the mind: Analogical processes and symbol systems. *Cognitive science, 34*(5), 752-775.

Gentner, D., Levine, S. C., Ping, R., Isaia, A., Dhillon, S., Bradley, C., & Honke, G. (2015). Rapid Learning in a Children's Museum via Analogical Comparison. *Cognitive science*, 1-17.

Gentner, D., Loewenstein, J., & Thompson, L. (2003). Learning and transfer: A general role for analogical encoding. *Journal of Educational Psychology, 95*(2), 393-408. doi:10.1037/0022-0663.95.2.393

Gentner, D., & Medina, J. (1998). Similarity and the Development of Rules. *Cognition, 65*(2), 263-297.

Gentner, D., & Namy, L. L. (1999). Comparison in the development of categories. *Cognitive Development, 14*(4), 487-513.

Gentner, D., & Ratterman, M. J. (1991). Language and the career of similarity. In S. A. Gelman & J. P. Byrnes (Eds.), *Perspectives on language and thought: Interrelations in development* (pp. 225-275). London: Cambridge University Press.

Gentner, D., & Toupin, C. (1986). Systematicity and Surface Similarity in the Development of Analogy. *Cognitive science, 10*(3), 277-300. doi:DOI 10.1207/s15516709cog1003_2

Gick, M. L., & Holyoak, K. J. (1980). Analogical Problem-Solving. *Cognitive psychology, 12*(3), 306-355. doi:Doi 10.1016/0010-0285(80)90013-4

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive psychology, 15*(1), 1-38.

Glynn, S. M. (2008). Making science concepts meaningful to students: Teaching with analogies. In S. Mikelskis-Seifert, U. Ringelband, & M. Brückmann (Eds.), *Four decades of research in science education: From curriculum development to quality improvement* (pp. 113-125). Münster, Germany: Waxman.

Goldstone, R. L., & Medin, D. L. (1994). Time-Course of Comparison. *Journal of Experimental Psychology-Learning Memory and Cognition, 20*(1), 29-50. doi:Doi 10.1037/0278-7393.20.1.29

Harrison, A. G., & Coll, R. K. (Eds.). (2007). *Using Analogies in Middle and Secondary Science Classrooms: The FAR Guide – An Interesting Way to Teach With Analogies*. Thousand Oaks, CA: SAGE Publications.

Harsley, R., Green, N., Alizadeh, M., Acharya, S., Fossati, D., Di Eugenio, B., & AlZoubi, O. (2016). *Incorporating Analogies and Worked Out Examples as Pedagogical Strategies in a Computer Science Tutoring System.* Paper presented at the Proceedings of the 47th ACM Technical Symposium on Computing Science Education.

Hestenes, D., & Halloun, I. (1995). Interpreting the force concept inventory. *The Physics Teacher, 33*(8), 502-506.

Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force concept inventory. *The Physics Teacher, 30*, 141.

Holyoak, K. J., & Richland, L. E. (2014). Using Analogies as a Basis for Teaching Cognitive Readiness. In H. O'Neil, R. Perez, & E. Baker (Eds.), *Teaching and Measuring Cognitive Readiness* (pp. 223-238). New York, NY: Springer Publishing Company.

Klenk, M., & Forbus, K. (2009). Analogical model formulation for transfer learning in AP Physics. *Artificial intelligence, 173*(18), 1615-1638.

Klenk, M., & Forbus, K. (2013). Exploiting persistent mappings in cross-domain analogical learning of physical domains. *Artificial intelligence, 195*, 398-417. doi:10.1016/j.artint.2012.11.002

Klenk, M., Forbus, K., Tomai, E., & Kim, H. (2011). Using analogical model formulation with sketches to solve Bennett Mechanical Comprehension Test problems. *Journal of Experimental & Theoretical Artificial Intelligence, 23*(3), 299-327.

Kolodner, J. L. (2006). Case-Based Reasoning. *Cambridge Handbook of the Learning Sciences*, 225-242. doi:Book_Doi 10.2277/ 0521607779

Kotovsky, L., & Gentner, D. (1996). Comparison and Categorization in the Development of Relational Similarity. *Child Development, 67*, 2797-2822.

Kurtz, K. J., Miao, C. H., & Gentner, D. (2001). Learning by analogical bootstrapping. *Journal of the Learning Sciences, 10*(4), 417-446. doi:DOI 10.1207/S15327809JLS1004new_2

Lakoff, G., & Johnson, M. (2003). *Metaphors we live by*. Chicago: University of Chicago Press.

Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive science, 11*(1), 65-100.

Leslie, S.-J. (2008). Generics: Cognition and acquisition. *Philosophical Review, 117*(1), 1-47.

Lockwood, K., & Forbus, K. (2009). *Multimodal knowledge capture from text and diagrams.* Paper presented at the Fifth international conference on Knowledge capture.

Loewenstein, J., Thompson, L., & Gentner, D. (1999). Analogical encoding facilitates knowledge transfer in negotiation. *Psychonomic Bulletin & Review, 6*(4), 586-597.

Lulis, E., & Evens, M. (2003). *The use of analogies in human tutoring dialogues.* Paper presented at the Natural language generation in spoken and written dialogue. AAAI Symposium, Stanford.

Lulis, E., Evens, M., & Michael, J. (2004). *Implementing analogies in an electronic tutoring system.* Paper presented at the Intelligent Tutoring Systems.

Lulis, E. A. (2005). *Implementing analogies in an intelligent tutoring system by modeling human tutoring sessions.* Illinois Institute of Technology.

Markman, A. B., & Gentner, D. (1993). Structural alignment during similarity comparisons. *Cognitive psychology, 25*(4), 431-467.

Mason, L. (2004). Fostering understanding by structural alignment as a route to analogical learning. *Instructional Science, 32*(4), 293-318. doi:Doi 10.1023/B:Truc.0000026512.88700.32

Mason, L., Pluchino, P., & Tornatora, M. C. (2013). Effects of Picture Labeling on Science Text Processing and Learning: Evidence From Eye Movements. *Reading Research Quarterly, 48*(2), 199-214. doi:10.1002/rrq.41

McFate, C. J., Forbus, K. D., & Hinrichs, T. (2014). *Using Narrative Function to Extract Qualitative Information from Natural Language Texts.* Paper presented at the Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence, Quebec City.

Murdock, J. W. (2011). Structure mapping for jeopardy! clues *Case-Based Reasoning Research and Development* (pp. 6-10): Springer.

Murray, T., Schultz, K., Brown, D., & Clement, J. (1990). An Analogy-Based Computer Tutor for Remediating Physics Misconceptions. *Interactive Learning Environments, 1*(2), 79-101.

Namy, L. L., & Gentner, D. (2002). Making a silk purse out of two sow's ears: young children's use of comparison in category learning. *Journal of Experimental Psychology: General, 131*(1), 5.

Nokes-Malach, T. J., VanLehn, K., Belenky, D. M., Lichtenstein, M., & Cox, G. (2013). Coordinating principles and examples through analogy and self-explanation. *European Journal of Psychology of Education, 28*(4), 1237-1263. doi:10.1007/s10212-012-0164-z

Richland, L. E., Zur, O., & Holyoak, K. J. (2007). Cognitive Supports for Analogies in the Mathematics Classroom. *Science, 316*(5828), 1128-1129. doi:10.1126/science.1142103

Speer, R., Havasi, C., & Lieberman, H. (2008). *AnalogySpace: Reducing the Dimensionality of Common Sense Knowledge.* Paper presented at the AAAI.

Spiro, R. J., Feltovich, P. J., Coulson, R. L., & Anderson, D. K. (1989). Multiple Analogies for Complex Concepts - Antidotes for Analogy-Induced Misconception in Advanced Knowledge Acquisition. *Similarity and Analogical Reasoning*, 498-531. doi:Doi 10.1017/Cbo9780511529863.023

Tomai, E., & Forbus, K. D. (2009). *EA NLU: Practical Language Understanding for Cognitive Modeling.* Paper presented at the Proceedings of the 22nd International Florida Artificial Intelligence Research Society Conference, Sanibel Island, Florida.

Veloso, M. M., & Carbonell, J. G. (1993). Derivational Analogy in Prodigy - Automating Case Acquisition, Storage, and Utilization. *Machine Learning, 10*(3), 249-278. doi:Doi 10.1023/A:1022686910523

Wai, J., Lubinski, D., & Benbow, C. P. (2009). Spatial ability for STEM domains: Aligning over 50 years of cumulative psychological knowledge solidifies its importance. *Journal of Educational Psychology, 101*(4), 817.

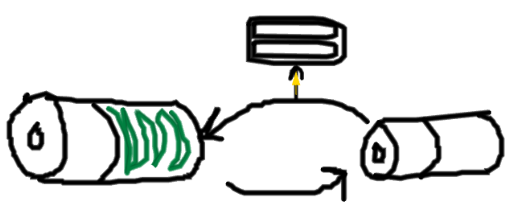
Yin, P., Forbus, K. D., Usher, J. M., Sageman, B., & Jee, B. D. (2010). *Sketch Worksheets: A Sketch-Based Educational Software System.* Paper presented at the Innovative Applications of Artificial Intelligence (IAAI).

Zeitoun, H. H. (1984). Teaching Scientific Analogies: a proposed model. *Research in Science & Technological Education, 2*(2), 107-125. doi:10.1080/0263514840020203

APPENDIX A: Instructional Analogies

ATP / battery

Text:
ATP is like a charged battery. The charged battery has electrical energy. The charged battery releases electrical energy to electronic devices. The charged battery changes into a drained battery when electrical energy is used. The charged battery can be used over and over again. A battery charger changes the flat battery into the charged battery. The ATP has chemical energy. The ATP releases chemical energy to cell parts. The ATP changes into ADP when chemical energy is used. The ATP can be used over and over again. Mitochondria changes ADP into ATP. The ADP is like the drained battery. The charged battery provides energy gradually, but the ATP provides energy immediately. A phosphate breaks off from the ATP, but nothing breaks off from the battery. Devices usually use two batteries, but cells use many ATP molecules.



Sketch objects:
```
(isa using-atp IntrinsicStateChangeEvent)
(isa ADP AdenosineDiphosphate)
(isa charging-the-battery ChargingABattery)
(isa ATP AdenosineTriphosphate)
(isa atp-energy EnergyStuff)
(isa battery-charger BatteryCharger)
(isa making-atp IntrinsicStateChangeEvent)
(isa phosphate Phosphate)
```

```
(isa battery-energy EnergyStuff)
(isa charged-battery ChargedBattery)
(isa drained-battery FlatBattery)
(isa Object-17 BreakingEvent)
(isa using-battery IntrinsicStateChangeEvent))
```

Gold Standard Queries in English, CycL:

Four rightmost columns show which queries succeeded in the four experimental conditions in Chapter 4.

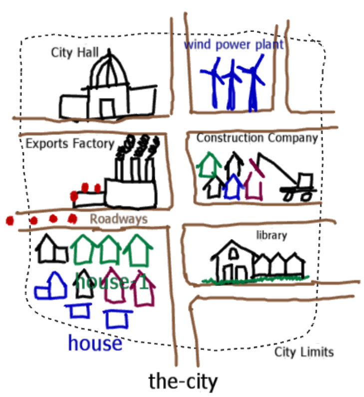
		Successfully captured by Model?			
Query in English	**Query in CycL**	**Full Model**	**No Back-ground**	**No Base**	**No Sketch**
ATP is energy sufficient.	`(relationAllExists possessiveRelation AdenosineTriphosphate EnergyStuff)`	✓	✓	✓	NO
ATP releases energy to parts/places in the cell.	`(relationExistsAll performedBy (SubcollectionOfWithRelationToTypeFn (SubcollectionOfWithRelationToTypeFn MakingSomethingAvailable target CellPart) transferredObject EnergyStuff) AdenosineTriphosphate)`	✓	✓	✓	NO
ADP converts to ATP.	`(relationAllExists toState (IntrinsicStateChangeOfFn AdenosineDiphosphate) AdenosineTriphosphate)`	✓	✓	NO	NO
ATP converts to ADP.	`(relationAllExists toState (IntrinsicStateChangeOfFn AdenosineTriphosphate) AdenosineDiphosphate)`	✓	✓	NO	NO
ATP can be used over and over again.	`(relationAll repeatedEvent (UsingInstanceFn AdenosineTriphosphate))`	✓	✓	NO	NO
Mitochondria are the sites where ADP gets changed to ATP.	`(relationExistsAll doneBy (IntrinsicStateChangeOfFn AdenosineDiphosphate) Mitochondrion)`	✓	✓	NO	NO
Battery's energy is released gradually, but ATP energy is released	`(relationAllInstance conceptuallyRelated (SubcollectionOfWithRelationToTypeFn (SubcollectionOfWithRelationToTypeFn MakingSomethingAvailable objectActedOn EnergyStuff) performedBy AdenosineTriphosphate) Immediately)`	✓	✓	✓	NO

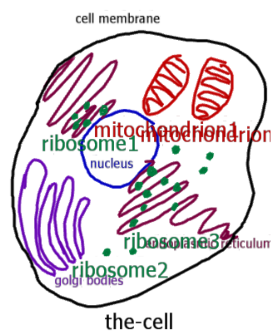
immediately.					
A phosphate breaks away from ATP.	`(relationAllExists from-UnderspecifiedLocation (SubcollectionOfWithRelationToTypeFn BreakingEvent objectOfStateChange Phosphate) AdenosineTriphosphate)`	✓	✓	✓	✓
Many ATP molecules are used.	`(relationAllInstance qualitativeExtent (SubcollectionOfWithRelationFromTypeFn (SubcollectionOfWithRelationFromTypeFn Molecule compoundNoun AdenosineTriphosphate) instrument-Generic (UsingInstanceFn (SubcollectionOfWithRelationFromTypeFn Molecule compoundNoun AdenosineTriphosphate))) Many-Quant)`	NO	NO	NO	NO
Energy is required to change ADP into ATP.	`(relationAllExists requires-Underspecified (IntrinsicStateChangeOfFn AdenosineDiphosphate) EnergyStuff)`	NO	NO	NO	NO
Energy is released when ATP changed into ADP.	`(relationAllExists releases-Underspecified (IntrinsicStateChangeOfFn AdenosineTriphosphate) EnergyStuff)))`	✓	NO	NO	✓

Cell/city

Text:

A cell is like a city. The nucleus is like the city government. The city government controls the city. The nucleus controls the cell. The mitochondrion is like the power station. The power station provides electricity. The mitochondrion provides chemical energy. Construction companies build houses. The ribosomes make proteins. Things in the city move along the road. The endoplasmic reticulum is like the road. Golgi bodies export substances outside of the cell. Factories export things outside of the city. The city government changes direction after elections and is very adaptable. Unlike the city, the nucleus always controls the cell.





Sketch objects:

```
(isa mitochondrion1 Mitochondrion)
(isa endoplasmic-reticulum EndoplasmicReticulum)
(isa membrane CellMembrane)
(isa ribosome3 Ribosome)
(isa construction-companies ConstructionCompany)
 (isa city-limits Border)
(isa golgi-bodies GolgiApparatus)
 (isa roads Roadway)
 (isa roads TransportationPathSystem)
(isa house-1 House-Modern)
(isa factory FactoryBuilding)
 (isa the-cell Cell)
 (isa house-3 House-Modern)
(isa windmills WindPowerPlant)
(isa library SchoolBuilding)
 (isa the-city City)
 (isa nucleus CellNucleus)
 (isa city-hall CityGovernment)
(isa ribosome1 Ribosome)
```

```
 (isa mitochondrion2 Mitochondrion)
(isa ribosome2 Ribosome))
```

Gold Standard Queries in English, CycL:

Four rightmost columns show which queries succeeded in the four experimental conditions in Chapter 4.

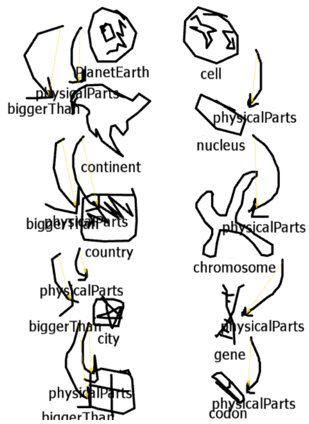
		Successfully captured by Model?			
Query in English	**Query in CycL**	**Full Model**	**No Back-ground**	**No Base**	**No Sketch**
Cells have cell membranes.	`(partTypes Cell CellMembrane)`	✓	✓	✓	NO
Cells have cell nuclei.	`(partTypes Cell CellNucleus)`	✓	✓	✓	NO
Cells have ribosomes.	`(partTypes Cell Ribosome)`	✓	NO	NO	NO
Cells have endoplasmic reticulum.	`(partTypes Cell EndoplasmicReticulum)`	✓	✓	✓	NO
Cells have mitochondria.	`(partTypes Cell Mitochondrion)`	✓	✓	✓	NO
Cell nuclei control cells.	`(relationExistsAll performedBy (SubcollectionOfWithRelationToTypeFn ControllingSomething objectControlled Cell) CellNucleus)`	✓	✓	✓	✓
Mitochondria provide energy.	`(relationExistsAll performedBy (MakingAbstractAvailableFn EnergyStuff) Mitochondrion)`	✓	✓	NO	NO
Ribosomes construct proteins.	`(relationExistsAll performedBy (MakingFn ProteinStuff) (SetOfTypeFn Ribosome))`	✓	✓	NO	NO
The endoplasmic reticulum transports things.	`(relationExistsAll doneBy Movement-TranslationEvent EndoplasmicReticulum)`	NO	NO	NO	NO
Cell functions are interdependent.	`(relationAllExists requires-Underspecified (TypicalBehaviorOfTypeFn Cell) (PartTypeFn Cell Individual))`	NO	NO	NO	✓
Cells communicate	`(relationAllExists communicatesWith Cell Cell)`	✓	NO	NO	✓

with other cells.					
Individual cell parts make up a functional system.	`(relationExistsAll members FunctionalSystem (PartTypeFn Cell Individual))))`	✓	NO	NO	✓

Cell/Earth

Text:

A cell is like planet Earth. A cell nucleus is like a continent. A chromosome is like a country. A gene is like a city. A codon is like a street address.



Sketch objects:

```
(isa codon Codon-MolecularSegment)
(isa street-address StreetAddress)
(isa gene Gene-HereditaryUnit)
(isa continent Continent)
(isa cell Cell)
(isa city City)
(isa chromosome Chromosome)
(isa nucleus CellNucleus)
(isa country Country)
(isa PlanetEarth Planet)
```

Gold Standard Queries in English, CycL:

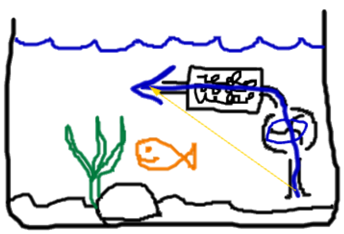
Four rightmost columns show which queries succeeded in the four experimental conditions in Chapter 4.

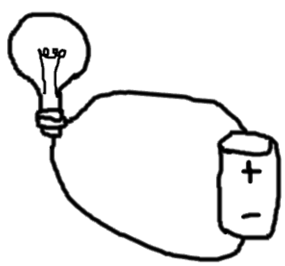
		Successfully captured by Model?			
Query in English	**Query in CycL**	**Full Model**	**No Back-ground**	**No Base**	**No Sketch**
Cells are bigger than cell nuclei.	`(relationAllExists biggerThan Cell CellNucleus)`	✓	✓	NO	NO
Cell nuclei are bigger than	`(relationAllExists biggerThan CellNucleus Chromosome)`	✓	✓	NO	NO

chromosomes.					
Chromosomes are bigger than genes.	`(relationAllExists biggerThan Chromosome Gene-HereditaryUnit)`	✓	✓	NO	NO
Genes are bigger than codons.	`(relationAllExists biggerThan Gene-HereditaryUnit Codon-MolecularSegment)`	✓	✓	✓	NO
Cells have cell nuclei.	`(relationAllExists physicalParts Cell CellNucleus)`	✓	✓	✓	NO
Cell nuclei have chromosomes.	`(relationAllExists physicalParts CellNucleus Chromosome)`	✓	✓	✓	NO
Chromosomes have genes.	`(relationAllExists physicalParts Chromosome Gene-HereditaryUnit)`	✓	✓	✓	NO
Genes have codons.	`(relationAllExists physicalParts Gene-HereditaryUnit Codon-MolecularSegment)`	✓	✓	NO	NO
Cell nuclei have many chromosomes.	`(relationAllExistsRange physicalParts CellNucleus Chromosome Many-Quant)`	✓	NO	NO	NO
Chromosomes have many genes.	`(relationAllExistsRange physicalParts Chromosome Gene-HereditaryUnit Many-Quant)`	✓	NO	NO	NO
Genes have many codons.	`(relationAllExistsRange physicalParts Gene-HereditaryUnit Codon-MolecularSegment Many-Quant)`	✓	NO	NO	NO

Circuit/Aquarium

Text:
A simple series circuit is like an aquarium. The electric current is like the water. The wire is like the pipe. The wire carries the electricity. The pipe carries the water. The battery uses voltage to push electrons. The pump uses pressure to push the water. Voltage is like pressure. Thin wire in the light bulb resists electric current. The filter resists water flow. Electric current is conserved. No water is lost. Water is a material liquid. Electricity is a flow of charge in an electric field. Water can flow in an incomplete cycle. Electricity always needs a complete circuit. Water flow depends on the pressure of the pump. Electric current depends on the entire circuit.





Sketch objects:
```
((isa water-pipe Pipe-GenericConduit) (isa water-pump Pump-Generic) (isa the-filter
Filter)
 (isa Object-302 Aquarium-Container) (isa circuit-wire-2 Wire) (isa the-battery
Battery)
 (isa the-light-bulb LightBulbIncandescent) (isa pump-pressure Pressure)
 (isa water-arrow DirectionOfMovement) (isa aquarium-water Water) (isa aquarium-water
Liquid-StateOfMatter)
 (isa aquarium-container Container) (isa aquarium-container Physob) (isa sandy-bits
SandySoilRegion)
 (isa circuit-wire-1 Wire) (isa Object-305 SeriesCircuit) (isa fishface Goldfish))
```

Gold Standard Queries in English, CycL:

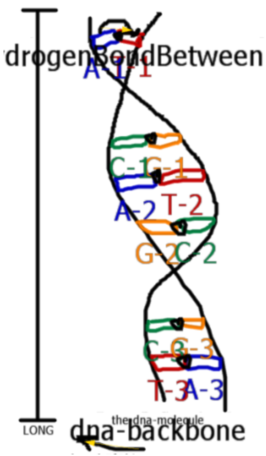
Four rightmost columns show which queries succeeded in the four experimental conditions in Chapter 4.

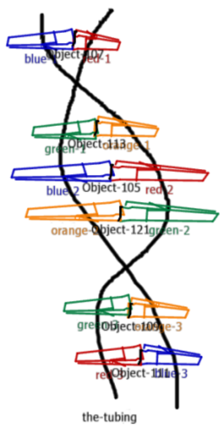
		Successfully captured by Model?			
Query in English	**Query in CycL**	**Full Model**	**No Back-ground**	**No Base**	**No Sketch**
Batteries push electrons (voltage).	`(relationExistsAll providerOfMotiveForce (SubcollectionOfWithRelationToTypeFn PushingAnObject objectActedOn Electron) Battery)`	✓	✓	✓	NO
Flow of electricity depends on a complete circuit.	`(dependsOn-Underspecified Electricity (CollectionIntersectionFn (TheSet ElectricalCircuit (ThingDescribableAsFn Entire-TheWord Adjective-Gradable))))`	NO	NO		NO
Wires resist the flow of electricity.	`(relationExistsAll doneBy (SubcollectionOfWithRelationToFn PreventingSomething objectActedOn Electricity) Wire)`	NO	NO		NO
Circuits have wires.	`(partTypes SeriesCircuit (SetOfTypeFn Wire))`	✓	✓	✓	NO
Circuits have batteries.	`(partTypes SeriesCircuit Battery)`	✓	✓	✓	NO
Circuits have light bulbs.	`(partTypes SeriesCircuit LightBulbIncandescent)`	✓	✓		NO

DNA/pins

Text:

DNA is like a clothespin structure. Cytosine is like a green clothespin. Guanine is like an orange clothespin. Thymine is like a red clothespin. Adenine is like a blue clothespin. The sugar phosphate backbone is like the tube. Weak hydrogen bonds are like the clothespin attachments. Any of the colored clothespins will clip together. Cytosine will bond only with guanine. Thymine will bond only with adenine. The clothespin structure is short, but DNA molecules are typically very long. DNA molecules typically have thousands of base pairs.





Sketch objects:

```
((isa red-3 ClothesPin)
(isa red-3 ColoredThing) (isa C-2 Cytosine-Base)
 (isa the-tubing Pipe-GenericConduit)
(isa C-3 Cytosine-Base) (isa G-3 Guanine-Base)
 (isa dna-backbone SugarPhosphateBackbone)
 (isa T-1 Thymine-Base) (isa T-1 ChemicalObject)
 (isa blue-2 ClothesPin) (isa blue-2 ColoredThing)
 (isa orange-1 ClothesPin) (isa orange-1 ColoredThing)
 (isa green-1 ClothesPin) (isa green-1 ColoredThing)
 (isa orange-3 ClothesPin) (isa G-1 Guanine-Base)
 (isa Object-155 HydrogenBond) (isa Object-121 Attachment)
(isa Object-149 HydrogenBond)
 (isa red-2 ClothesPin) (isa red-2 ColoredThing)
 (isa T-3 Thymine-Base) (isa Object-109 Attachment)
 (isa T-2 Thymine-Base) (isa Object-107 Attachment)
 (isa Object-145 HydrogenBond)
 (isa Object-153 HydrogenBond) (isa A-1 ChemicalObject)
(isa A-1 Adenine-Base) (isa green-3 ClothesPin)
 (isa blue-1 ClothesPin) (isa blue-1 ColoredThing)
(isa green-2 ClothesPin) (isa green-2 ColoredThing)
 (isa blue-3 ClothesPin) (isa blue-3 ColoredThing)
```

```
(isa A-2 Adenine-Base) (isa Object-111 Attachment)
(isa Object-113 Attachment) (isa Object-105 Attachment)
(isa C-1 Cytosine-Base) (isa red-1 ColoredThing)
(isa red-1 ClothesPin) (isa Object-147 HydrogenBond)
(isa orange-2 ClothesPin) (isa orange-2 ColoredThing)
(isa Object-151 HydrogenBond) (isa the-dna-molecule DNAMolecule)
(isa the-dna-molecule SpatialThing-NonSituational) (isa Long Distance))
```

Gold Standard Queries in English, CycL:

Four rightmost columns show which queries succeeded in the four experimental conditions in Chapter 4.

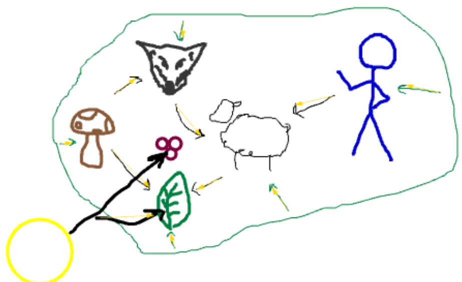
		Successfully captured by Model?			
Query in English	**Query in CycL**	**Full Model**	**No Back-ground**	**No Base**	**No Sketch**
Adenine bases are parts of DNA.	`(partTypes DNAMolecule Adenine-Base)`	✓	✓	✓	NO
Cytosine bases are parts of DNA.	`(partTypes DNAMolecule Cytosine-Base)`	✓	✓	✓	NO
Guanine bases are parts of DNA.	`(partTypes DNAMolecule Guanine-Base)`	✓	✓	✓	NO
Thymine bases are parts of DNA.	`(partTypes DNAMolecule Thymine-Base)`	✓	✓	✓	NO
DNA molecules have a sugar phosphate backbone.	`(partTypes DNAMolecule SugarPhosphateBackbone)`	✓	✓	✓	NO
Thymine bonds with Adenine.	`(relationAllExists hydrogenBondBetween Thymine-Base Adenine-Base)`	✓	✓	✓	NO
Cytosine bonds with Guanine.	`(relationAllExists hydrogenBondBetween Cytosine-Base Guanine-Base)`	✓	✓	✓	NO
Thymine only bonds with Adenine.	`(relationAllExistsAndOnly hydrogenBondBetween Thymine-Base Adenine-Base)`	✓	✓	✓	NO
Cytosine only bonds with	`(relationAllExistsAndOnly hydrogenBondBetween`	✓	✓	✓	NO

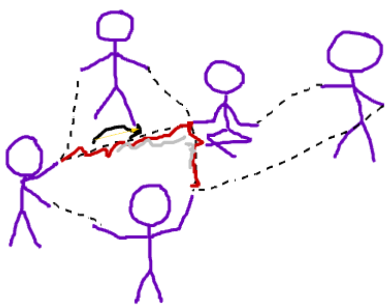
Guanine.	`Cytosine-Base Guanine-Base)`				
DNA molecules are very long.	`(relationAllInstance lengthOfObject DNAMolecule Long)`	✓	✓	✓	NO
Thymine connects with Adenine.	`(relationAllExists connectedTo-Directly Thymine-Base Adenine-Base)`	NO	NO	NO	NO
Cytosine connects with Guanine.	`(relationAllExists connectedTo-Directly Cytosine-Base Guanine-Base)`	NO	NO	NO	NO

Ecosystem / Web

Text:

An ecosystem is like a web structure. A change in the ecosystem is like movement in the structure. Any movement in the structure decreases its stability. The stability of the structure is like the stability of the ecosystem.





Sketch objects:

```
((isa connecting-string ConstructionArtifact) (isa student-2 Student)
 (isa student-2 TemporalThing) (isa string-moving MovementEvent) (isa student-1
Student)
 (isa energy-conversion IntrinsicStateChangeEvent) (isa chem-energy ChemicalEnergy)
 (isa the-ecosystem Ecosystem) (isa leaf Individual) (isa leaf Autotroph) (isa leaf
BiologicalLivingObject)
 (isa leaf Plant) (isa web-stability Stability) (isa person Animal) (isa person
Predator)
 (isa person BiologicalLivingObject) (isa person Heterotroph) (isa person Human)
 (isa the-string String-Textile) (isa the-string Physob) (isa wolf Heterotroph) (isa
wolf Predator)
 (isa wolf BiologicalLivingObject) (isa wolf Animal) (isa mushroom Fungus)
 (isa mushroom (CollectionUnionFn (TheSet Fungus Bacterium))) (isa solar SolarEnergy)
(isa sheep Animal)
 (isa sheep BiologicalLivingObject) (isa sheep Prey) (isa sheep Sheep-Domestic) (isa
sheep Heterotroph))
```

Gold Standard Queries in English, CycL:

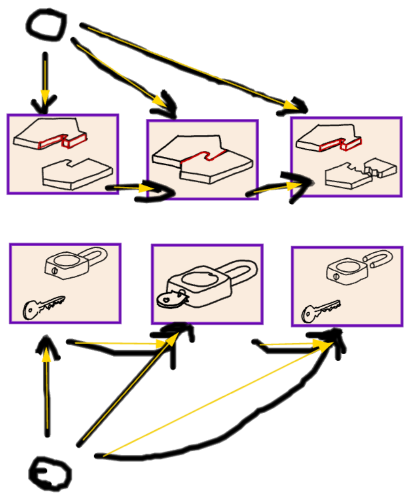
Four rightmost columns show which queries succeeded in the four experimental conditions in Chapter 4.

		Successfully captured by Model?			
Query in English	**Query in CycL**	**Full Model**	**No Back-ground**	**No Base**	**No Sketch**
Ecosystems have producers.	`(relationAllExists possessiveRelation Ecosystem Autotroph)`	✓	✓	NO	NO
Ecosystems have consumers.	`(relationAllExists possessiveRelation Ecosystem Heterotroph)`	✓	✓	✓	NO
Ecosystems have predators.	`(relationAllExists possessiveRelation Ecosystem Predator)`	✓	✓	✓	NO
Ecosystems have prey.	`(relationAllExists possessiveRelation Ecosystem Prey)`	✓	✓	✓	NO
Fungi are decomposers.	`(relationExistsAll decomposerInEcosystem Ecosystem Fungus)`	✓	✓	✓	NO
Autotrophs are producers.	`(relationExistsAll producerInEcosystem Ecosystem Autotroph)`	✓	✓	NO	NO
Heterotrophs are consumers.	`(relationExistsAll consumerInEcosystem Ecosystem Heterotroph)`	✓	✓	✓	NO
Predators eat prey.	`(relationAllExists eatsWillingly Predator Prey)`	✓	✓	✓	NO
Producers convert solar energy to chemical energy.	`(relationExistsAll doneBy (SubcollectionOfWithRelationToTypeFn (SubcollectionOfWithRelationToTypeFn IntrinsicStateChangeEvent toState ChemicalEnergy) objectOfStateChange SolarEnergy) Autotroph)`	✓	✓	NO	NO
Changes in the ecosystem decrease stability.	`(qprop- ((QPQuantityFn Stability) Ecosystem) (RateFn (IntrinsicStateChangeOfFn Ecosystem)))`	✓	NO	NO	NO

Enzyme/Key

Text:

An enzyme is like a key. A substrate molecule is like a lock. The key has ridges. The ridges have a unique shape. The enzyme binding site is like the ridges. The enzyme binding site has a unique chemical makeup. The key only unlocks specific locks. The enzyme reacts only with specific substrate molecules. The key unlocks the lock. The enzyme breaks apart the substrate molecules. After unlocking a lock, the key is unchanged. The key can be used over and over again. Enzyme action is like unlocking a lock. The enzyme is unchanged by the chemical reaction. The enzyme can be used over and over again.



Sketch objects:

```
((isa the-enzyme EnzymeMolecule) (isa the-substrate SubstrateMolecule) (isa active-
site EnzymeBindingSite))
((isa the-ridge Ridge) (isa the-lock Lock) (isa the-key KeyOfLock))
((isa Object-129 Event) (isa Object-129 ChemicalReaction) (isa Object-129
EnzymeActivationEvent)
 (isa Object-135 Event) (isa Object-135 UnlockingALock))
(isa lock-s1 Event)
 (isa lock-s2 Event)
 (isa lock-s2 Situation)  (isa enzymeaction-s3 Situation) (isa enzymeaction-s3 Event)
(isa enzymeaction-s2 Situation)
(isa lock-s3 Event)
(isa lock-s3 Situation)
(isa enzymeaction-s2 Event)
 (isa lock-s3 TemporalThing) (isa enzymeaction-s2 TemporalThing)
 (isa enzymeaction-s1 Event)
(isa enzymeaction-s1 TemporalThing)
(isa enzymeaction-s1 Situation)
(isa enzymeaction-s3 BreakingEvent)
 (isa lock-s1 TemporalThing)
(isa lock-s1 Situation)
```

```
(isa lock-s2 TemporalThing)
(isa enzymeaction-s3 TemporalThing)
```

Gold Standard Queries in English, CycL:

Four rightmost columns show which queries succeeded in the four experimental conditions in Chapter 4.

		Successfully captured by Model?			
Query in English	**Query in CycL**	**Full Model**	**No Back-ground**	**No Base**	**No Sketch**
Enzyme molecules have unique binding sites.	`(partTypes EnzymeMolecule (CollectionIntersectionFn (TheSet EnzymeBindingSite (ThingDescribableAsFn Unique-TheWord Adjective))))`	✓	✓	✓	NO
Enzymes only react with specific substrates.	`(relationAllExistsAndOnly chemicalReactants (SubcollectionOfWithRelationToTypeFn (SubcollectionOfWithRelationToTypeFn ChemicalReaction chemicalReactants SubstrateMolecule) catalyst EnzymeMolecule) (CollectionIntersectionFn (TheSet SubstrateMolecule (ThingDescribableAsFn Specific-TheWord Adjective))))`	NO	NO	NO	NO
Enzyme action breaks apart substrate molecules.	`(relationAllExists subEvents EnzymeActivationEvent (SubcollectionOfWithRelationToTypeFn (SubcollectionOfWithRelationToTypeFn BreakingEvent doneBy EnzymeMolecule) objectOfStateChange SubstrateMolecule))`	✓	✓	✓	NO
The enzyme comes out of the reaction unchanged.	`(relationExistsAll unchangedActors EnzymeActivationEvent EnzymeMolecule)`	✓	✓	✓	NO
When enzymes bind to substrate molecules, it enables the substrate to break apart.	`(enables-SitTypeSitType (SubcollectionOfWithRelationToFn (CollectionIntersectionFn (TheSet (SubcollectionOfWithRelationToTypeFn Situation subEvents ChemicalReaction) (SubcollectionOfWithRelationToTypeFn Situation subEvents`	✓	✓	✓	NO

	EnzymeActivationEvent) Situation)) holdsIn (relationExistsExists objectsIntersect EnzymeMolecule SubstrateMolecule)) (SubcollectionOfWithRelationToTypeFn (SubcollectionOfWithRelationToTypeFn BreakingEvent doneBy EnzymeMolecule) objectOfStateChange SubstrateMolecule))				
Enzymes can be reused.	(relationAll repeatedEvent EnzymeActivationEvent)	✓	NO	NO	NO

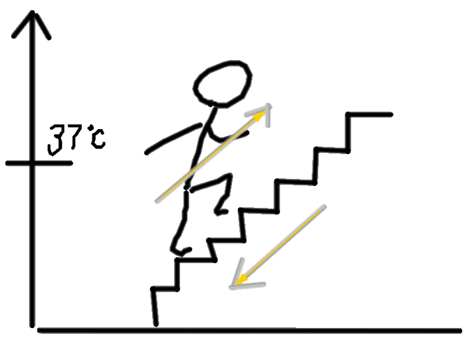
Homeostasis / Escalator

Text:

Homeostasis is like going up a down escalator. The escalator is moving down. The person is walking up. Heat is constantly escaping from the human body. Heat escaping is like the escalator moving down. The human body makes the heat by exercising. Making the heat is like walking up. Temperature is like the height of the person.

The heat loss and the heat production are balanced, so the temperature remains constant.

During homeostasis, the ideal temperature of the human body is 37 degrees Celsius.



Sketch objects:

```
((isa person-motion Walking-Generic) (isa the-escalator Physob) (isa the-escalator
Escalator)
 (isa escalator-motion MovementEvent) (isa temp-scale Temperature) (isa temp-scale
ScalarOrVectorInterval)
 (isa bob Physob) (isa bob Person) (isa esc-dir DirectionOfMovement) (isa bob-dir
DirectionOfMovement)
 (isa going-up-a-down-esc GoingUpADownEscalator) (isa temp-value
TemperatureIndicator))
```

Gold Standard Queries in English, CycL:

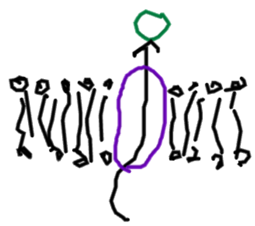
Four rightmost columns show which queries succeeded in the four experimental conditions in Chapter 4.

		Successfully captured by Model?			
Query in English	**Query in CycL**	**Full Model**	**No Back-ground**	**No Base**	**No Sketch**
Heat escapes from the body.	`(relationExistsAll doneBy (SubcollectionOfWithRelationToTypeFn LeavingAPlace fromLocation HumanBody) ThermalEnergy)`	NO	NO	NO	NO
The body produces heat.	`(relationAllExists products (SubcollectionOfWithRelationToTypeFn MakingSomething performedBy HumanBody) ThermalEnergy)`	✓	✓	✓	✓
Heat loss and heat production are balanced so temperature remains constant.	`(implies (qEqualTo (RateFn (SubcollectionOfWithRelationToTypeFn (MakingFn ThermalEnergy) beneficiary HumanBody)) (RateFn (SubcollectionOfWithRelationToTypeFn LeavingAPlace doneBy ThermalEnergy))) (qEqualTo (QPDerivativeFn (CollectionIntersectionFn (TheSet Temperature (ThingDescribableAsFn Ideal-TheWord Adjective)))) Zero))`	✓	NO	NO	NO
Body temperature (positively) depends on heat production.	`(qprop ((QPQuantityFn (CollectionIntersectionFn (TheSet Temperature (ThingDescribableAsFn Ideal-TheWord Adjective)))) HumanBody) (RateFn (SubcollectionOfWithRelationToTypeFn (MakingFn ThermalEnergy) beneficiary HumanBody)))`	✓	NO	NO	NO
Body temperature (negatively) depends on heat loss.	`(qprop- ((QPQuantityFn (CollectionIntersectionFn (TheSet Temperature (ThingDescribableAsFn Ideal-TheWord Adjective)))) HumanBody) (RateFn`	✓	NO	NO	NO

	(SubcollectionOfWithRelationToTypeFn LeavingAPlace doneBy ThermalEnergy)))				
Ideal body temperature is 37°C.	(relationAllInstance measure (CollectionIntersectionFn (TheSet Temperature (ThingDescribableAsFn Ideal-TheWord Adjective))) (DegreeCelsius 37))	✓	✓	✓	✓

Membrane / Mosaic

Text:
A membrane is like a mosaic that has liquid grout. The membrane protein is like the mosaic tiles. The lipid bilayer is like the grout. The tiles move in the grout. The membrane protein moves in the lipid bilayer. The membrane protein moves some chemical substances across the membrane. The membrane protein blocks some substances. Some tiles are made of different materials. Some membrane proteins have different functions.



Sketch objects:
```
((isa moving-thru-mosaic CausingAnotherObjectsTranslationalMotion) (isa moving-thru-
mosaic MovementEvent)
 (isa chemical-substance ChemicalObject) (isa the-fluid-mosaic Mosaic) (isa mosaic-
tile Tile)
 (isa liquid-grout Grout) (isa moving-thru-membrane MovementEvent)
 (isa moving-thru-membrane CausingAnotherObjectsTranslationalMotion) (isa membrane-
protein MembraneProtein)
 (isa lipid-bilayer LipidBilayer) (isa the-bio-membrane BiologicalMembrane))
```

Gold Standard Queries in English, CycL:

Four rightmost columns show which queries succeeded in the four experimental conditions in Chapter 4.

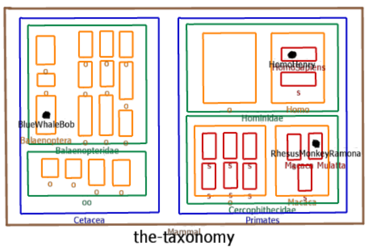
		Successfully captured by Model?			
Query in English	**Query in CycL**	**Full Model**	**No Back-ground**	**No Base**	**No Sketch**
Membrane proteins move things across the membrane.	`(relationExistsAll doneBy (SubcollectionOfWithRelationToTypeFn (SubcollectionOfWithRelationToTypeFn CausingAnotherObjectsTranslationalMotion objectActedOn (SetOfTypeFn PartiallyTangible)) across-UnderspecifiedRegion BiologicalMembrane) MembraneProtein)`	✓	✓	NO	✓
Membranes can block substances.	`(relationExistsAll barrier (SubcollectionOfWithRelationToTypeFn BarrierSituation blockedPath (SetOfTypeFn PartiallyTangible)) MembraneProtein)`	✓	✓	✓	✓
Different membrane proteins have different functions.	`(relationAllExists possessiveRelation (SetOfTypeFn MembraneProtein) (CollectionIntersectionFn (TheSet Capability (ThingDescribableAsFn Differ-TheWord Adjective-Gradable))))`	✓	✓	✓	NO
Membrane proteins move within the lipid bilayer.	`(and (relationExistsAll primaryObjectMoving MovementEvent MembraneProtein) (relationAllExists in-UnderspecifiedContainer MembraneProtein LipidBilayer))`	✓	✓	✓	NO
Membranes have membrane proteins.	`(partTypes BiologicalMembrane MembraneProtein)`	✓	✓	NO	NO
Membranes have lipid bilayers.	`(partTypes BiologicalMembrane LipidBilayer)))`	✓	✓	✓	NO

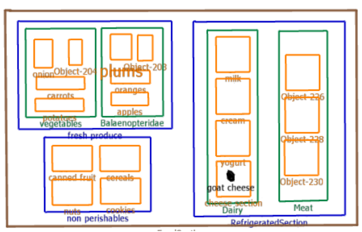
Taxonomy / Supermarket

Text:

A biological classification system is like a supermarket. Supermarkets have many sections. Customers use the sections to find the products. Large sections have many different products. Small sections have fewer products. Finding products depends on attributes of products. Some products fit into many different sections. Supermarkets use different sections. But all supermarkets have sections.

Biological classification systems have many groups. Biologists use groups to study species. Large groups have high diversity. Small groups have low diversity. Grouping species depends on attributes of species. Some species are difficult to group. Different biological classification systems have different groups. But all biological classification systems have groups.





Sketch objects:

```
((isa HomoGenus BiologicalGenus) (isa HomoGenus Group) (isa Object-76
BiologicalFamily)
 (isa Object-76 Group) (isa Object-62 Group) (isa Object-62 BiologicalSpecies)
 (isa RefrigeratedMarketCategory ExistingObjectType)
 (isa RefrigeratedMarketCategory RefrigeratedMarketCategory)
 (isa RefrigeratedMarketCategory ProductTypeByMarketCategory) (isa Object-218
GroceryMarketCategory)
 (isa Object-134 BlueWhale) (isa Object-134 Organism-Whole) (isa Object-90
BiologicalOrder)
 (isa Object-90 Group) (isa Object-241 Product) (isa Object-241 GoatCheese)
 (isa cheese-section CheeseMarketCategory) (isa DairyMarketCategory
ExistingObjectType)
 (isa DairyMarketCategory DairyMarketCategory) (isa DairyMarketCategory
ProductTypeByMarketCategory)
 (isa Object-232 NonperishableMarketCategory) (isa Object-102 BiologicalGenus) (isa
Object-102 Group)
 (isa Object-99 BiologicalGenus) (isa Object-99 Group) (isa Object-92 Group)
```

```
 (isa Object-92 BiologicalFamily) (isa Object-124 BiologicalGenus) (isa Object-124
Group)
 (isa Object-113 Group) (isa Object-113 BiologicalGenus) (isa HominidaeFamily Group)
 (isa HominidaeFamily BiologicalFamily) (isa HominidaeFamily
QAClarifyingCollectionType)
 (isa homo-henry Organism-Whole) (isa homo-henry HomoSapiens) (isa Object-239
NutsMarketCategory)
 (isa Object-111 BiologicalGenus) (isa Object-111 Group) (isa HomoSapiens
OrganismClassificationType)
 (isa HomoSapiens Group) (isa HomoSapiens BiologicalSpecies) (isa Object-52
BiologicalClass)
 (isa Object-52 Group) (isa Object-114 Group) (isa Object-114 BiologicalGenus)
 (isa Object-119 BiologicalGenus) (isa Object-119 Group) (isa rhesus-ramona Monkey)
 (isa rhesus-ramona Organism-Whole) (isa MacacaMulattaSpecies BiologicalSpecies)
 (isa MacacaMulattaSpecies Group) (isa Object-72 Group) (isa Object-72
BiologicalGenus)
 (isa Object-56 BiologicalSpecies) (isa Object-56 Group) (isa Object-126
BiologicalGenus)
 (isa Object-126 Group) (isa Object-125 Group) (isa Object-125 BiologicalGenus) (isa
Object-74 Group)
 (isa Object-74 BiologicalGenus) (isa Object-118 BiologicalGenus) (isa Object-118
Group)
 (isa Object-58 BiologicalSpecies) (isa Object-58 Group) (isa yogurt-section
YogurtCategory)
 (isa Primate KEClarifyingCollectionType) (isa Primate OrganismClassificationType)
(isa Primate Group)
 (isa Primate BiologicalOrder) (isa Object-94 BiologicalFamily) (isa Object-94 Group)
 (isa Object-60 BiologicalSpecies) (isa Object-60 Group) (isa Object-101
BiologicalGenus)
 (isa Object-101 Group) (isa Object-112 Group) (isa Object-112 BiologicalGenus)
 (isa Object-234 FreshFruitMarketCategory) (isa Object-109 BiologicalGenus) (isa
Object-109 Group)
 (isa homo-genus-2 Group) (isa homo-genus-2 BiologicalGenus) (isa Object-68 Group)
 (isa Object-68 BiologicalSpecies) (isa Object-70 BiologicalSpecies) (isa Object-70
Group)
 (isa Object-64 Group) (isa Object-64 BiologicalSpecies) (isa Object-66 Group)
 (isa Object-66 BiologicalSpecies) (isa homo-species-2 BiologicalSpecies) (isa homo-
species-2 Group)
 (isa Object-100 BiologicalGenus) (isa Object-196 FreshVegetablesMarketCategory)
 (isa Object-110 BiologicalGenus) (isa Object-110 Group) (isa Object-120
BiologicalGenus)
 (isa Object-120 Group) (isa milk-section MilkMarketCategory) (isa cream-category
CreamMarketCategory)
 (isa the-taxonomy Taxonomy) (isa the-store GroceryStore))
```

Gold Standard Queries in English, CycL:

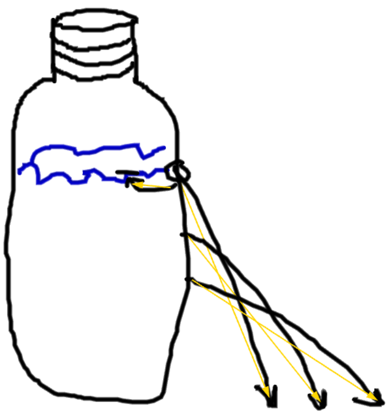
Four rightmost columns show which queries succeeded in the four experimental conditions in Chapter 4.

		Successfully captured by Model?			
Query in English	**Query in CycL**	**Full Model**	**No Back-ground**	**No Base**	**No Sketch**
Taxonomies are classification systems.	`(genls Taxonomy ClassificationSystem)`	✓	✓	✓	NO
Larger groups have greater diversity.	`(relationAllInstance possessiveRelation (SubcollectionOfWithRelationToFn Group sizeParameterOfObject (RelativeGenericValueFn sizeParameterOfObject Group highAmountOf)) (HighAmountFn Diversity))`	NO	NO	NO	NO
Smaller groups have lower diversity.	`(relationAllInstance possessiveRelation (SubcollectionOfWithRelationToFn Group sizeParameterOfObject (RelativeGenericValueFn sizeParameterOfObject Group veryLowToLowAmountOf)) (LowAmountFn Diversity))`	NO	NO	NO	NO
Grouping an organism depends on its properties.	`(relationAllExists dependsOn-Underspecified (GroupingOfFn Organism-Whole) (SubcollectionOfWithRelationFromTypeFn AttributeCharacteristicOfAnEntity possessiveRelation BiologicalSpecies))`	NO	NO	NO	NO
Some species are difficult to classify.	`(relationExistsInstance degreeOfDifficulty (GroupingFn BiologicalSpecies) Difficult)`	NO	NO	NO	NO
Taxonomies have many different groups.	`(relationAllExistsRange possessiveRelation Taxonomy (SetOfTypeFn Group) Many-Quant)`	NO	NO	✓	NO
Order is more specific than class.	`(relationAllExists possessiveRelation BiologicalClass BiologicalOrder)`	✓	NO	✓	NO
Family is more specific than order.	`(relationAllExists possessiveRelation BiologicalOrder BiologicalFamily)`	✓	NO	✓	NO

Genus is more specific than family.	`(relationAllExists possessiveRelation BiologicalFamily BiologicalGenus)`	✓	NO	✓	NO
Species is more specific than genus.	`(relationAllExists possessiveRelation BiologicalGenus BiologicalSpecies)`	NO	NO	NO	NO
An organism can belong to many different groups.	`(relationAllExistsRange member Organism-Whole Group Many-Quant)`	NO	NO	NO	NO

Voltage / Pressure

Text:
A battery is like a hole in a bottle. The water flows through the hole. The greater the depth, the greater the water flow rate. Electricity flows through the battery. The battery has voltage. Voltage is like depth. The greater the voltage, the greater the flow of electricity.



Sketch objects:
```
((isa Object-46 DirectionOfMovement) (isa hole-in-bottle PolyDimensionalThing)
 (isa hole-in-bottle FluidConduit) (isa the-bottle Container) (isa the-bottle Bottle)
(isa the-bottle Physob) (isa depth-of-hole Depth) (isa Object-44 DirectionOfMovement)
(isa the-water Liquid-StateOfMatter) (isa Object-43 DirectionOfMovement))
```

Gold Standard Queries in English, CycL:

Four rightmost columns show which queries succeeded in the four experimental conditions in Chapter 4.

		Successfully captured by Model?			
Query in English	**Query in CycL**	**Full Model**	**No Back-ground**	**No Base**	**No Sketch**
The flow of electricity (positively) depends on the battery voltage.	`(qprop (RateFn (FlowFn Electricity)) ((QPQuantityFn Voltage) Battery))`	✓	NO	NO	NO

APPENDIX B: Science Questions

NYSRE-2014-8SCI-19	1
NYSRE-2014-LE-01	2
NYSRE-2014-LE-04	3
NYSRE-2014-LE-21	4
NYSRE-2014-LE-28	5
MCAS-2014-BIO-03	6
MCAS-2014-BIO-07	7
MCAS-2014-BIO-09	8
MCAS-2014-BIO-10	9
MCAS-2014-BIO-17	10
MCAS-2014-BIO-35	11
MCAS-2014-BIO-39	12
MCAS-K8-EE-14	13
MCAS-K8-EE-20	14

1. A single human body cell typically contains thousands of
    a. genes
    b. nuclei
    c. chloroplasts
    d. bacteria
2. A function of cell membranes is humans is the
    a. synthesis of the amino acids
    b. production of energy
    c. replication of genetic material
    d. recognition of certain chemicals
3. Which statement best describes the organelles in a cell?
    a. All organelles are involved directly with communication between cells.
    b. Organelles must work together and their activities must be coordinated
    c. Organelles function only when there is a disruption in homeostasis
    d. Each organelle must function independently of the others in order to maintain homeostasis
4. Human population growth has led to a reduction in the populations of predators throughout natural ecosystems across the United States. Scientists consider the loss of these predators to have a
    a. positive effect, because an increase in their prey helps to maintain stability in the ecosystem
    b. positive effect, because the predators usually eliminate the species they prey on

c. negative effect, because predators have always made up a large portion of our food supply
d. negative effect, because predators have an important role in maintaining stable ecosystems

5. A variety of pear tree, known as Bradford, was originally introduced into the eastern United States in the 1960s. Today, this tree is crowding out other plants in these states. This situation best illustrates
   a. an unintentional negative effect of altering an ecosystem
   b. how a foreign species is controlled in the eastern United States
   c. that the introduction of a foreign species does not affect food webs
   d. that serious environmental consequences can be avoided by importing a foreign species
6. If the producers in a food web were removed, which of the following changes would most likely occur?
   a. The entire food web would collapse over time.
   b. The food web would depend on the decomposers for energy.
   c. The consumers would begin making energy for the food web.
   d. The populations of the remaining organisms in the food web would increase.
7. Within a prey population, which of the following is most immediately affected by the arrival of a new predator?
   a. death rate
   b. evolution rate
   c. immigration rate
   d. maturation rate
8. Which of the following is a function of the nucleus in organism 2?
   a. absorbing sunlight
   b. releasing usable energy
   c. storing genetic material
   d. producing food molecules
9. In the three organisms, what are synthesized by the ribosomes?
   a. Carbohydrates
   b. Lipids
   c. nucleic acids
   d. proteins
10. Which of the following best describes the producers in a terrestrial food web?
    a. They are at the highest trophic level.
    b. They are not affected by decomposers.
    c. They convert solar energy to chemical energy.
    d. They obtain all their nutrients and energy from consumers.
11. Which of the following is the best example of the human body maintaining homeostasis?
    a. The heart beats using cardiac muscle

   b. The breathing rate increases during exercise.
   c. The nose and ear contain cartilage for flexibility.
   d. The digestive system uses enzymes to break down food.
12. In a sample of double-stranded DNA, 30% of the nitrogenous bases are thymine. What percentage of the nitrogenous bases in the sample are adenine?
   a. 20%
   b. 30%
   c. 60%
   d. 70%
13. Jamal wants to make an electrical circuit, but he only has the objects shown below. Which of the following must Jamal also have to make an electrical circuit?
   a. a motor
   b. a switch
   c. a bar magnet
   d. a power source
14. The diagram below shows a project that a student made to test an electrical circuit. Part of the electrical circuit is underneath the board. When the student connect the two nails using a wire, the bulb lights up. Which of the following must be underneath the board?
   a. a magnet and a switch
   b. a switch and some wires
   c. a magnet and a power source
   d. a power source and some wires

```
;; NYSRE-2014-8SCI-19
;; A single human body cell typically contains thousands of
(queryForQuestion NYSRE-2014-8SCI-19 (relationAllExistsRange physicalParts Cell
?cell-part Thousands-Quant))
(multipleChoiceSingleOptionList NYSRE-2014-8SCI-19 (TheList Gene-HereditaryUnit 1))
(multipleChoiceSingleOptionList NYSRE-2014-8SCI-19 (TheList CellNucleus 2))
(multipleChoiceSingleOptionList NYSRE-2014-8SCI-19 (TheList Chloroplast 3))
(multipleChoiceSingleOptionList NYSRE-2014-8SCI-19 (TheList Bacterium 4))
(correctAnswerChoice NYSRE-2014-8SCI-19 1)
```

```
;; NYSRE-2014-LE-01
;; A function of cell membranes is humans is the
```

```
(queryForQuestion NYSRE-2014-LE-01 (relationExistsAll doneBy ?functional-event
CellMembrane))
(multipleChoiceSingleOptionList NYSRE-2014-LE-01 (TheList (MakingFn AminoAcid) 1))
(multipleChoiceSingleOptionList NYSRE-2014-LE-01 (TheList (MakingFn EnergyStuff) 2))
(multipleChoiceSingleOptionList NYSRE-2014-LE-01 (TheList Replication-DNA 3))
(multipleChoiceSingleOptionList NYSRE-2014-LE-01 (TheList (IdentifyingAsTypeFn
ChemicalObject) 4))
(correctAnswerChoice NYSRE-2014-LE-01 4)


;; NYSRE-2014-LE-04
;; Which statement best describes the organelles in a cell?
(queryForQuestion NYSRE-2014-LE-04 (trueQuestionOption NYSRE-2014-LE-04 ?n))
(multipleChoiceSingleOptionList
 NYSRE-2014-LE-04
 (TheList (relationExistsAll
           doneBy
           (SubcollectionOfWithRelationToTypeFn
            InformationTransferPhysicalEvent
            informationDestination
            Cell)
           Organelle)
          1))
(multipleChoiceSingleOptionList
 NYSRE-2014-LE-04
 (TheList (relationExistsAll members FunctionalSystem Organelle) 2))
(multipleChoiceSingleOptionList
 NYSRE-2014-LE-04
 (TheList (responsibleFor-TypeType (PreventingFn Homeostasis) (ActorPlaysRoleFn
Organelle doneBy)) 3))
(multipleChoiceSingleOptionList
 NYSRE-2014-LE-04
 (TheList (responsibleFor-TypeType
           (ActorPlaysRoleFn Organelle doneBy)
           Homeostasis)
          4))
(correctAnswerChoice NYSRE-2014-LE-04 2)



(in-microtheory (ATQuestionMtFn NYSRE-2014-LE-21) :exclude-globals t)
(isa (ATQuestionMtFn NYSRE-2014-LE-21) (InformationThingAboutFn NYSRE-2014-LE-21))
(isa this-ecosystem Ecosystem)
(consumerInEcosystem this-ecosystem these-humans)
(consumerInEcosystem this-ecosystem other-predators)
```

```
(eatsWillingly other-predators other-prey)
(isa human-pop-growth IncreaseEvent)
(isa human-pop-growth IntrinsicStateChangeEvent)
(objectOfStateChange human-pop-growth this-ecosystem)
(objectOfStateChange human-pop-growth other-predators)
(isa predator-decline DecreaseEvent)
(isa predator-decline IntrinsicStateChangeEvent)
(objectOfStateChange predator-decline this-ecosystem)
(causes-EventEvent human-pop-growth predator-decline)
;; Question info
(in-microtheory AnalogyTutorEvaluationQuestionsMt)
(queryForQuestion NYSRE-2014-LE-21 (trueQuestionOption NYSRE-2014-LE-21 ?n))
(multipleChoiceSingleOptionList NYSRE-2014-LE-21 (TheList (and (causes-Event
predator-decline ?prey-increase)
                                                              (isa ?prey-increase
IncreaseEvent)
                                                              ;; something about
increasing prey population

(positivelyInfluencedBy ((QPQuantityFn Stability) this-ecosystem)

(RateFn ?prey-increase)))
                                                         1))
(multipleChoicesingleOptionList NYSRE-2014-LE-21 (TheList (and (causes-Event
predator-decline ?ext)
                                                              (isa ?ext Extinction)

(positivelyInfluencedBy
                                                               ((QPQuantityFn
Stability) this-ecosystem)
                                                               (RateFn ?ext)))
                                                         2))
(multipleChoicesingleOptionList NYSRE-2014-LE-21 (TheList (and (eatsWillingly these-
humans ?food-supply)

(negativelyInfluencedBy
                                                               (AmountOfFn ?food-
supply)
                                                               (RateFn predator-
decline)))
                                                         3))
(multipleChoicesingleOptionList NYSRE-2014-LE-21 (TheList (negativelyInfluencedBy
                                                          ((QPQuantityFn Stability)
this-ecosystem)
                                                          (RateFn predator-decline))
                                                         4))
```

```
(correctAnswerChoice NYSRE-2014-LE-21 4)



;;
;; NYSRE-2014-LE-28
 (in-microtheory (ATQuestionMtFn NYSRE-2014-LE-28))
(isa (ATQuestionMtFn NYSRE-2014-LE-28) (InformationThingAboutFn NYSRE-2014-LE-28))
(isa BradfordTree OrganismClassificationType)
(genls BradfordTree Tree-ThePlant)
(isa this-ecosystem Ecosystem)
(producerInEcosystem this-ecosystem these-bradford-trees)
(producerInEcosystem this-ecosystem other-plants)
(isa introduction-of-btree IncreaseEvent)
(isa introduction-of-btree IntrinsicStateChangeEvent)
(objectOfStateChange introduction-of-btree this-ecosystem)
(objectOfStateChange introduction-of-btree other-plants)
;; Question info
(in-microtheory AnalogyTutorEvaluationQuestionsMt)
(queryForQuestion NYSRE-2014-LE-28 (trueQuestionOption NYSRE-2014-LE-28 ?n))
(multipleChoiceSingleOptionList NYSRE-2014-LE-28 (TheList (negativelyInfluencedBy
                                                           ((QPQuantityFn Stability)
this-ecosystem)
                                                           (RateFn introduction-of-
btree))
                                                          1))
(multipleChoicesingleOptionList NYSRE-2014-LE-28 (TheList (objectControlled
?controlling-btrees these-bradford-trees)
                                                          2))
(multipleChoicesingleOptionList NYSRE-2014-LE-28 (TheList (falseSentence
                                                           (influencedBy
                                                            ((QPQuantityFn Stability)
this-ecosystem)
                                                            (RateFn introduction-of-
btree)))
                                                          3))
(multipleChoicesingleOptionList NYSRE-2014-LE-28 (TheList (preventedProp
                                                           introduction-of-btree
                                                           (negativelyInfluencedBy
                                                            ((QPQuantityFn Stability)
this-ecosystem)
                                                            (RateFn introduction-of-
btree)))
                                                          4))

(correctAnswerChoice NYSRE-2014-LE-28 1)
```

```
;;
;; MCAS-2014-BIO-03
 (in-microtheory (ATQuestionMtFn MCAS-2014-BIO-03))
(isa (ATQuestionMtFn MCAS-2014-BIO-03) (InformationThingAboutFn MCAS-2014-BIO-03))
(isa this-ecosystem Ecosystem)
(producerInEcosystem this-ecosystem these-producers)
(isa these-producers Autotroph)
;; question info
(in-microtheory AnalogyTutorEvaluationQuestionsMt)
(queryForQuestion MCAS-2014-BIO-03 (trueQuestionOption MCAS-2014-BIO-03 ?n))
(multipleChoiceSingleOptionList
 MCAS-2014-BIO-03
 (TheList (relationAllExists
           requires-Underspecified
           Ecosystem
           Autotroph)
          1))
(multipleChoiceSingleOptionList
 MCAS-2014-BIO-03
 (TheList (relationExistsExists
           producerInEcosystem
           Ecosystem
           (SubcollectionOfWithRelationFromTypeFn BiologicalLivingObject
decomposerInEcosystem Ecosystem))
          2))
(multipleChoiceSingleOptionList
 MCAS-2014-BIO-03
 (TheList (relationExistsExists
           producerInEcosystem
           Ecosystem
           Heterotroph)
          3))
(multipleChoiceSingleOptionList
 MCAS-2014-BIO-03
 (TheList (negativelyInfluencedBy
           ((QPQuantityFn Population) this-ecosystem)
           ((QPQuantityFn Population) these-producers))
          4))
(correctAnswerChoice MCAS-2014-BIO-03 1)
```

```
;;
;; MCAS-2014-BIO-07
 (in-microtheory (ATQuestionMtFn MCAS-2014-BIO-07))
(isa (ATQuestionMtFn MCAS-2014-BIO-07) (InformationThingAboutFn MCAS-2014-BIO-07))
(isa this-ecosystem Ecosystem)
(isa predator (SetOfTypeFn BiologicalLivingObject))
(predatorInEcosystem this-ecosystem predator)
(isa prey (SetOfTypeFn BiologicalLivingObject))
(preyInEcosystem this-ecosystem prey)
(eatsWillingly predator prey)
;; question info:
(in-microtheory AnalogyTutorEvaluationQuestionsMt)
(queryForQuestion MCAS-2014-BIO-07 (influencedBy ?x ((QPQuantityFn Population)
predator)))
(multipleChoiceSingleOptionList MCAS-2014-BIO-07 (TheList (RateFn (DeathFn prey)) 1))
(multipleChoiceSingleOptionList MCAS-2014-BIO-07 (TheList (RateFn
(SubcollectionOfWithRelationToFn Evolution doneBy prey)) 2))
(multipleChoiceSingleOptionList MCAS-2014-BIO-07 (TheList (RateFn
(SubcollectionOfWithRelationToFn Immigration doneBy prey)) 3))
(multipleChoiceSingleOptionList MCAS-2014-BIO-07 (TheList (RateFn
(SubcollectionOfWithRelationToFn Maturing doneBy prey)) 4))
(correctAnswerChoice MCAS-2014-BIO-07 1)
```

```
;;
;; MCAS-2014-BIO-09
;; Which of the following is a function of the nucleus in organism 2?
(visualAidForQuestion MCAS-2014-BIO-09 (VisualAidForQuestionMtFn MCAS-2014-BIO-08-
11))
(queryForQuestion MCAS-2014-BIO-09 (relationExistsAll performedBy ?functional-event
CellNucleus))
(multipleChoiceSingleOptionList MCAS-2014-BIO-09 (TheList
(SubcollectionOfWithRelationToTypeFn AbsorptionEvent objectActedOn Sunlight) 1))
(multipleChoiceSingleOptionList MCAS-2014-BIO-09 (TheList (MakingAbstractAvailableFn
EnergyStuff) 2))
(multipleChoiceSingleOptionList MCAS-2014-BIO-09 (TheList (StoringFn Gene-
HereditaryUnit) 3))
(multipleChoiceSingleOptionList MCAS-2014-BIO-09 (TheList (MakingFn Food) 4))
(correctAnswerChoice MCAS-2014-BIO-09 3)
```

```
;;
;; MCAS-2014-BIO-10
 (queryForQuestion MCAS-2014-BIO-10 (relationExistsAll performedBy (MakingFn
?product) Ribosome))
(multipleChoiceSingleOptionList MCAS-2014-BIO-10 (TheList CarbohydrateStuff 1))
(multipleChoiceSingleOptionList MCAS-2014-BIO-10 (TheList Lipid 2))
(multipleChoiceSingleOptionList MCAS-2014-BIO-10 (TheList DNAMolecule 3))
(multipleChoiceSingleOptionList MCAS-2014-BIO-10 (TheList ProteinStuff 4))
(correctAnswerChoice MCAS-2014-BIO-10 4)



;; MCAS-2014-BIO-17
 (queryForQuestion MCAS-2014-BIO-17 (trueQuestionOption MCAS-2014-BIO-17 ?n))
(multipleChoiceSingleOptionList MCAS-2014-BIO-17 (TheList (genls Autotroph
Heterotroph) 1))
(multipleChoiceSingleOptionList
 MCAS-2014-BIO-17
 (TheList
  (falseSentence (dependsOn-Underspecified
                  Autotroph
                  (SubcollectionOfWithRelationFromTypeFn BiologicalLivingObject
decomposerInEcosystem Ecosystem)))
  2))
(multipleChoiceSingleOptionList
 MCAS-2014-BIO-17
 (TheList
  (relationExistsAll
   doneBy
   (SubcollectionOfWithRelationToTypeFn
    (SubcollectionOfWithRelationToTypeFn IntrinsicStateChangeEvent toState
ChemicalEnergy)
    objectOfStateChange
    SolarEnergy)
   Autotroph)
  3))
(multipleChoiceSingleOptionList
 MCAS-2014-BIO-17
 (TheList
  (relationExistsAll
   from-UnderspecifiedLocation
   (SubcollectionOfwithRelationToTypeFn
    (MakingAvailableFn Food)
```

```
    doneBy
    (SubcollectionOfWithRelationFromTypeFn BiologicalLivingObject consumerInEcosystem
Ecosystem))
   Autotroph)
  4))
(correctAnswerChoice MCAS-2014-BIO-17 3)
```

```
;; MCAS-2014-BIO-35
 (in-microtheory (ATQuestionMtFn MCAS-2014-BIO-35))
(isa (ATQuestionMtFn MCAS-2014-BIO-35) (InformationThingAboutFn MCAS-2014-BIO-35))
;; situation 1
(isa heart-beating HeartBeating)
(instrument-Generic heart-beating muscle)
(isa muscle CardiacMuscle)
;; situation 2
(isa breathing Breathing)
(isa exercise Exercising)
(causes-EventEvent exercise breathing)
(positivelyInfluencedBy (RateFn breathing) (RateFn exercise))
;; situation 3
(isa nose Nose)
(isa ear Ear)
(isa nose-and-ear-makeup Configuration)
(holdsIn nose-and-ear-makeup
         (possessiveRelation nose nose-cartilage))
(holdsIn nose-and-ear-makeup
         (possessiveRelation ear ear-cartilage))
(positivelyInfluencedBy
 ((QPQuantityFn Flexibility) ear)
 (AmountOfFn ear-cartilage))
(positivelyInfluencedBy
 ((QPQuantityFn Flexibility) nose)
 (AmountOfFn nose-cartilage))
;; situation 4
(isa e EnzymeMolecule)
(isa d DigestiveSystem)
(isa breaking-down-food DigestionEvent)
(doneBy breaking-down-food d)
(instrument-Generic breaking-down-food e)
;; question info
(in-microtheory AnalogyTutorEvaluationQuestionsMt)
```

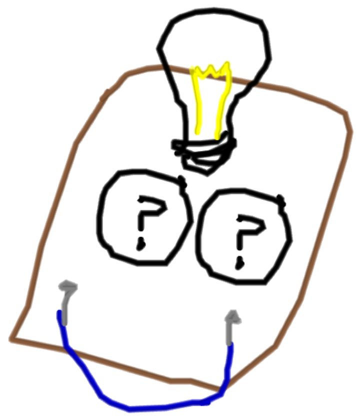
```
(queryForQuestion MCAS-2014-BIO-35 (isa ?x BiologicalHomeostasis))
(multipleChoiceSingleOptionList MCAS-2014-BIO-35 (TheList heart-beating 1))
(multipleChoiceSingleOptionList MCAS-2014-BIO-35 (TheList breathing 2))
(multipleChoiceSingleOptionList MCAS-2014-BIO-35 (TheList nose-and-ear-makeup 3))
(multipleChoiceSingleOptionList MCAS-2014-BIO-35 (TheList breaking-down-food 4))
```

```
(in-microtheory (ATQuestionMtFn MCAS-2014-BIO-39))
(isa (ATQuestionMtFn MCAS-2014-BIO-39) (InformationThingAboutFn MCAS-2014-BIO-39))
(isa sample-dna DNAMolecule)
(isa thymine-bases (SetOfTypeFn Thymine-Base))
(isa adenine-bases (SetOfTypeFn Adenine-Base))
(isa guanine-bases (SetOfTypeFn Guanine-Base))
(isa cytosine-bases (SetOfTypeFn Cytosine-Base))
(physicalParts sample-dna thymine-bases)
(physicalParts sample-dna adenine-bases)
(physicalParts sample-dna guanine-bases)
(physicalParts sample-dna cytosine-bases)
(percentOfIndividual thymine-bases sample-dna (Percent 30))

(queryForQuestion MCAS-2014-BIO-39 (percentOfIndividual adenine-bases sample-dna
(Percent ?x)))
(multipleChoiceSingleOptionList MCAS-2014-BIO-39 (TheList 20 1))
(multipleChoiceSingleOptionList MCAS-2014-BIO-39 (TheList 30 2))
(multipleChoiceSingleOptionList MCAS-2014-BIO-39 (TheList 60 3))
(multipleChoiceSingleOptionList MCAS-2014-BIO-39 (TheList 70 4))
(correctAnswerChoice MCAS-2014-BIO-39 2)
```

```
;;
;; MCAS-K8-EE-14
;; Jamal wants to make an electrical circuit, but he only has the objects shown
below.
;; Which of the following must Jamal also have to make an electrical circuit?
;; Scenario info
(in-microtheory (ATQuestionMtFn MCAS-K8-EE-14))
(isa (ATQuestionMtFn MCAS-K8-EE-14) (InformationThingAboutFn MCAS-K8-EE-14))
(isa circuit-materials (SetOfTypeFn ElectricalComponent))
(isa hypothetical-circuit SeriesCircuit)
(isa jamals-wire Wire)
(physicalParts hypothetical-circuit jamals-wire)
```

```
(isa jamals-bulb LightBulbIncandescent)
(physicalParts hypothetical-circuit jamals-bulb)
(physicalParts hypothetical-circuit missing-piece)
;; End scenario info
;; Question info
(in-microtheory AnalogyTutorEvaluationQuestionsMt)
(queryForQuestion MCAS-K8-EE-14 (isa missing-piece ?missing-type))
(multipleChoiceSingleOptionList MCAS-K8-EE-14 (TheList ElectricalMotor 1))
(multipleChoiceSingleOptionList MCAS-K8-EE-14 (TheList ElectricalSwitch 2))
(multipleChoiceSingleOptionList MCAS-K8-EE-14 (TheList Magnet 3))
(multipleChoiceSingleOptionList MCAS-K8-EE-14 (TheList Battery 4))
(correctAnswerChoice MCAS-K8-EE-14 4)


;;
;; MCAS-K8-EE-20
;; scenario info
(in-microtheory (ATQuestionMtFn MCAS-K8-EE-20))
(isa (ATQuestionMtFn MCAS-K8-EE-20) (InformationThingAboutFn MCAS-K8-EE-20))
(visualAidForQuestion MCAS-K8-EE-20 (VisualAidForQuestionMtFn MCAS-K8-EE-20))
;; question info
(in-microtheory AnalogyTutorEvaluationQuestionsMt)
(queryForQuestion MCAS-K8-EE-20 (and (isa hidden-thing1 ?hidden-type1)
                                     (isa hidden-thing2 ?hidden-type2)))
(multipleChoiceSingleOptionList MCAS-K8-EE-20 (TheList (TheSet Magnet
ElectricalSwitch) 1))
(multipleChoiceSingleOptionList MCAS-K8-EE-20 (TheList (TheSet ElectricalSwitch
(SetOfTypeFn Wire)) 2))
(multipleChoiceSingleOptionList MCAS-K8-EE-20 (TheList (TheSet Magnet Battery) 3))
(multipleChoiceSingleOptionList MCAS-K8-EE-20 (TheList (TheSet Battery (SetOfTypeFn
Wire)) 4))
(correctAnswerChoice MCAS-K8-EE-20 4)
```

Visual aid for question: `MCAS-K8-EE-20 4`



Sketch Objects:

```
(isa diagram-nail2 Nail)
(isa diagram-board Board-PieceOfWood)
(isa diagram-bulb LightBulbIncandescent)
(isa student-circuit SeriesCircuit)
(isa diagram-wire Wire)
(isa student-circuit ElectricalCircuit)
(isa diagram-nail1 Nail)
```

Sketch Objects with no labels:

```
hidden-thing1
hidden-thing2
```