

Spatial Reasoning in Comparative Analyses of Physics Diagrams

Maria D. Chang, Jon W. Wetzel, and Kenneth D. Forbus

Qualitative Reasoning Group, Northwestern University, Evanston, IL, USA
{maria.chang, jw}@u.northwestern.edu,
forbus@northwestern.edu

Abstract. Spatial reasoning plays a critical role in STEM problem solving. Physics assessments, for example, are rich in diagrams and pictures, which help people understand concrete physical scenarios and abstract aspects of physical reasoning. In this paper we describe a system that analyzes sketched diagrams to solve qualitative physics problems from a popular physics textbook. Causal models describing each problem are formulated via visual and conceptual analyses of the sketched diagrams. We use a combination of qualitative and quantitative reasoning to solve vector addition, tension, and gravitation ranking problems in the introductory chapters of the book.

Keywords: Spatial reasoning, spatial problem solving, qualitative reasoning, conceptual physics, diagrammatic reasoning.

1 Introduction

In spatial domains like science, technology, engineering and mathematics (i.e. STEM fields), problem solving frequently requires a combination of spatial and conceptual reasoning. For this reason, spatial representations (e.g. diagrams and sketches) are pervasive in science instruction. Spatial representations act as tools of communication and, unlike photographs, may convey a mix of concrete and abstract, non-literal information [1]. Externalizing this information has the added benefit of lightening working memory load, making spatial inference easier, and promoting new ideas [2, 3]. Actively generating spatial representations through sketching has been shown to increase engagement and facilitate learning [4]. Taking advantage of these spatial representations requires considerable spatial skills. Indeed, data from over 50 years of psychological research indicate that spatial skills are a strong predictor of success in STEM fields [5].

The importance of spatial representations in physics problem solving is illustrated by the use of diagrams in assessment tools and the emphasis on drawing free-body diagrams in introductory physics courses. For example, in an analysis of advanced placement (AP) physics tests in the US, 48% of problems had diagrams and 58% of those (about 28% of the total) could not be solved without information provided by the diagram [6]. Diagrams are especially common in qualitative physics problems:

two-thirds of the problems in the force concept inventory [7] involve diagrams. The presence of diagrams in the force concept inventory is important because qualitative physics problems have been shown to be a better probe for conceptual knowledge than quantitative problem solving.

The use of drawings and diagrams in science instruction presents a challenge to researchers interested in developing intelligent tutoring systems. From a computational perspective, using spatial representations to solve problems requires domain-general spatial reasoning capabilities, domain-specific knowledge, and models for how to combine both types of information. This is especially important in the domain of conceptual physics, where problems need to be solved via spatial reasoning and often in the absence of numerical values. It is therefore important for the next generation of STEM tutoring systems to support spatial reasoning for solving both quantitative and qualitative problems.

This paper describes a system that solves conceptual physics ranking problems from a popular physics text book [8]. Each problem is represented with a sketch. The visual and conceptual information depicted in each sketch is used to formulate causal models about them. Our system uses qualitative reasoning over causal models and spatial reasoning over the sketches to make judgments about quantities even when the exact values of the quantities are unknown.

2 Background

This section describes the pre-existing systems and techniques that we use for understanding sketches and reasoning about physical systems. We use CogSketch to collect and analyze physics sketches. We use qualitative mechanics and qualitative process theory to formulate and reason about causal models of physical scenarios and we use differential qualitative analysis to determine ordinal relationships between quantities.

2.1 CogSketch

To capture the spatial and conceptual information of the diagrammatic problems, we use CogSketch [9]¹, our domain-independent sketch understanding system. CogSketch provides an interface for sketching that automatically computes qualitative spatial and conceptual representations of sketches. CogSketch analyzes ink drawn by the user to generate topological relations (RCC8 [10]) and positional relations between objects (e.g. rightOf, above). Quantities from spatial computations (e.g. geometric distance) may be computed on demand. Ink may also be segmented into individual edges, which can be used to generate shape and edge level representations [11].

¹ CogSketch is freely available from:
www.qrg.northwestern.edu/software/cogsketch

Users manually label their sketches so that CogSketch can tie conceptual information to the spatial information from drawn ink. The labeling interface provided by CogSketch is intended to mimic the qualitative and communicative nature of open-domain sketching. When people sketch with each other, they frequently use natural language or gestures to communicate conceptual information about their drawing, rather than requiring others to interpret the ink in isolation. Similarly, CogSketch interprets users' ink with respect to the conceptual information they provide. In contrast, sketch recognition systems rely on users drawing elements from a small vocabulary of visual symbols (e.g. letters, electronic components). But symbol recognition is not enough to capture content in STEM domains because the mapping between shapes and conceptual entities is many to many. For example, in physics a circle might represent a ball, an orbit, or a disk. Thus, CogSketch's interface enables users to explicitly tell it what they mean.

In CogSketch, the basic building block of a sketch is a *glyph*. A *glyph* is a collection of ink strokes intended to represent something visually. Users define glyphs by telling CogSketch when they are done drawing something, and they can split and merge ink to edit glyphs as desired. Once the user draws a glyph, they provide it with a *conceptual label* from an OpenCyc-derived knowledge base (KB)², which can indicate that the glyph represents an *entity*, *relation*, or *annotation*. An *entity* is an instance of a concept, like a rigid object or a spring. A *relation* can be used to represent a relationship between two things, for instance, an arrow indicating that one object orbits another. An *annotation* provides additional information about another object. Annotations are of particular interest to this work because they can be used to represent quantities in diagrams. Quantitative values can be associated with annotations, but they are optional. This means that quantities can be represented at different levels of detail. For instance, in a free body diagram with two forces acting on an object, force vector annotation arrows can be drawn to represent the two forces (Figure 1). In the absence of quantitative information (i.e. force magnitude), the forces simply have some qualitative direction. This is still useful, since even without quantitative information, the qualitative direction of the net force may be estimated if the two forces are not in opposing directions. However, if force magnitudes are known, they can be entered using the same labeling interface.

Each sketch can contain multiple *subsketches*. The representations for each subsketch reside in their own logical environment, which inherits information from the sketch as a whole. Thus, subsketches can be thought of as independent states or scenarios within one sketch.

Because CogSketch generates relational representations for sketches, a sketch can be compared to other sketches using the *structure mapping engine* (SME) [12], which is a computational model of analogical comparison. SME takes two structured descriptions, a *base* and a *target*, and computes one or more mappings between them. For the purposes of this experiment, the most important aspect of the analogical mapping is the set of *correspondences*, which indicate how things in one description match up to things in another. Since subsketches may be used to represent indepen-

² <http://www.cyc.com/platform/opencyc>

dent scenarios, SME is used to frame comparative analyses between sketched scenarios. For instance, in two situations with multiple objects, there may be many potential matches between objects and quantities. An SME mapping indicates how items in each subsketch should match to each other, which is used to guide comparative analyses. SME is built into CogSketch, and has been used in an educational software application to give feedback to students [13].

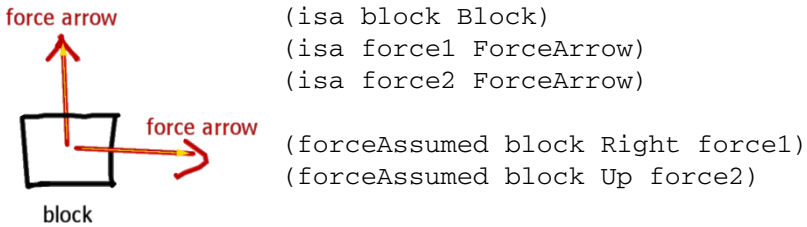


Fig. 1. An illustration of force arrow annotations. Quantitative values (e.g. force magnitude) may be included with annotations, but they are optional. A subset of the representations generated by CogSketch is shown on the right.

2.2 Qualitative and Quantitative Physical Reasoning

One of challenges of solving conceptual physics problems is combining qualitative and quantitative reasoning. Some problems may require arithmetic operations on numerical values, while others require reasoning about quantities for which no numerical values are known. Consider for example determining the net force on an object where the magnitudes of all forces are known. If all the forces are along the same axis, computing the magnitude of the net force can be reduced to addition and subtraction. Alternatively, consider a question that asks about the relative magnitude of the gravitational forces on the Moon versus Jupiter. It can be determined that the gravitational force on Jupiter is greater than on the moon because the mass of Jupiter is greater than than the mass of the moon. This inference is independent of the magnitudes of those forces and the exact masses of the Moon and Jupiter. Both types of reasoning are captured by our system.

For performing qualitative reasoning on force vectors and detecting forces between objects, we use a model of qualitative mechanics that is built into CogSketch [14]. The model is based on the work of Nielsen and Kim [15, 16] but has been adapted to capture rigid body mechanics of 2D objects in hand-drawn sketches. Vectors are represented using qualitative directions (e.g. right, left, up, down) and qualitative values (e.g. -, 0, +) to enable qualitative calculations about net force and the propagation of forces between object surfaces. While qualitative rigid body mechanics is useful for purely qualitative scenarios, there are often cases where quantities need to taken into account.

For reasoning about quantities explicitly, we use Qualitative Process Theory [17], which allows us to represent causal systems depicted in sketches. In QP theory,

physical phenomena are represented with continuous processes and quantities. Quantities represent parameters of objects (e.g. mass, velocity), while processes change quantities over time (e.g. acceleration changes velocity). Under the sole mechanism assumption [17], continuous quantities may only be changed by physical processes. Importantly, quantities can be reasoned about even if their exact numerical values are unknown. Quantities may be causally influenced directly or indirectly. *Direct influences* express contributions of rates of processes to the derivatives of quantities that they directly affect. More precisely, a direct influence means that the derivative of the quantity being influenced is equal to the sum of all direct influences (positive and negative) on it. Direct influences express the direct causal effects of continuous processes. *Indirect influences* (also called *qualitative proportionalities*) indicate instantaneous causal relationships between parameters. For example, Newton's second law, may be represented as:

(qprop acceleration Force)

(qprop- acceleration mass)

These two statements indicate that changes in force and mass cause changes in acceleration. All else being equal, the qprop indicates that an increase in force causes an increase in acceleration, whereas the qprop- indicates that an increase in mass will lead to decreased acceleration.

As a representational system, QP theory can be used to capture the causal models of a wide range of phenomena [17]. Qualitative causal models can be applied to a particular scenario via *model formulation* [18, 19]. Given a domain theory containing *model fragments* for a particular domain, a scenario can be analyzed to determine which model fragments are applicable. Each model fragment can include direct and indirect influences in its consequences, which can be used to determine how quantities are changing. This propagation of causal influences among quantities is called *influence resolution*. For example, if a force is applied to an object, that provides acceleration, which causes velocity to increase, which causes position to increase. Taken together, the model fragments that are applicable to a particular scenario represent the causal model of that scenario.

In addition to characterizing quantity changes within a single scenario, it is possible to determine quantity changes between highly similar scenarios via *differential qualitative analysis (DQA)* [17, 20]. The goal of DQA is to predict how a situation would change if some of its parameters not already constrained by its causal model were changed. DQA has been explored in sketch-based physics problem solving [21, 22] but we extend previous work by using first-principles modeling strategies and by expanding the range of visual quantities that can be assessed using DQA.

3 Approach

The goal of this work is to create the domain knowledge and reasoning capabilities needed to solve conceptual diagrammatic physics problems. To evaluate our efforts, we examined ranking problems from the first five chapters of *Conceptual Physics* [8]. Each ranking problem consists of two or more scenarios depicted in diagrams that need to be ranked along a particular quantity (Figure 2, 3). For instance, a ranking problem may involve understanding how opposite forces combine to make a net force. In some cases the problems involve basic arithmetic. In others, numerical values might be left out completely. For instance, some ranking problems require ranking scenarios by a particular quantity even if no precise values are known. To solve these problems, we developed domain theories for some of the basic concepts in the first five chapters of the book, which cover Newton's Laws. We also developed strategies for conducting differential qualitative analysis that use spatial and conceptual information depicted in sketches.

3.1 Ranking Problems

The first five chapters of *Conceptual Physics* have 27 ranking problems. Out of those 27, 12 have to do with combining vectors as well as conceptual knowledge of tension and gravity. We chose these problems as a starting point, excluding exercises about Newton's third law and problems that deal with explicit calculations about time, because we think those problems will require other aspects of QP theory and vectors beyond what we discuss here.

Each ranking problem was sketched into CogSketch so that spatial and conceptual representations could be automatically generated. The scenarios within each problem were drawn as subsketches within the sketch representing the whole problem. Every object was given one or more labels using the CogSketch labeling interface. These labels were important for objects with special properties (e.g. ropes) and for vector and quantity annotations. Vector annotations could have specific numerical values associated with them, but these were only included if they were explicitly mentioned in the problem and in the diagram. Similarly, objects were given mass annotations if mass was explicitly mentioned. There were two concepts that we had to represent by adding extra information to the sketch: hanging and co-movement. To indicate that something was hanging by something else, we drew relationship arrows to explicitly include this information in the sketch. To indicate that something moved with something else (as is necessary for some net velocity calculations), we drew relationship arrows as well. In all other cases, the sketch only contained information given explicitly in the diagram. In all cases, the sketch only contained information given explicitly in the problem text.

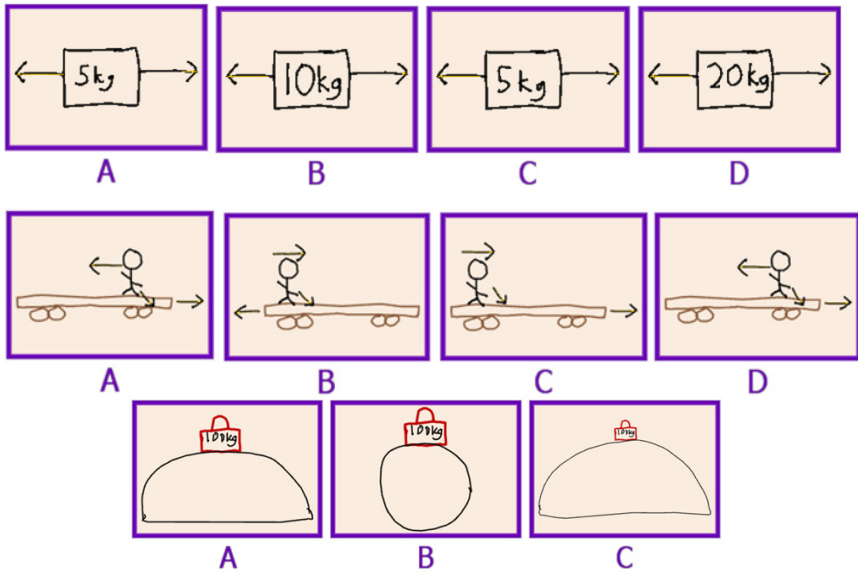


Fig. 2. Three example ranking problems with numerical values entered via CogSketch’s labeling interface. The upper problem requires ranking the net forces on the objects in the four scenarios. The middle problem requires ranking the net velocity on the person in the four scenarios. The lower problem requires ranking the mass and weight of the toolbox on Jupiter, versus Earth’s moon, versus Earth. The force and velocity arrows have explicit numerical values associated with them (e.g. 5 Newtons) and some of the objects have mass associated with them (e.g. 100 kg).

3.2 Model Fragments

The first five chapters of *Conceptual Physics* cover Newton’s three laws of motion. As a starting point, we defined model fragments to express the causal relationships between tension forces in hanging systems, gravitational forces, and vectors drawn as arrows. Model fragments can be used to represent physical processes (e.g. motion) and physical situations (e.g. inclined plane). Each model fragment has participants (i.e. entities that are involved in the causal model), conditions and consequences. In order for a model fragment to be applicable to a particular scenario, its participant constraints and conditions must be satisfied. If the model fragment is applicable, it is instantiated and its consequences are inferred. Model fragment consequences can involve assertions about quantities, including causal relationships (i.e. influences) and ordinal relationships.

A simplified example of a model fragment that describes vector addition is shown in Figure 4. The participants are the things involved: the object that the vectors describe, a set of the vectors that should be summed, and the resultant vector. Each participant must belong to a particular collection in order for the model fragment to be considered. Those collections, which come from the OpenCyc KB with our own extensions, are shown in parentheses. The participant constraints provide bindings for

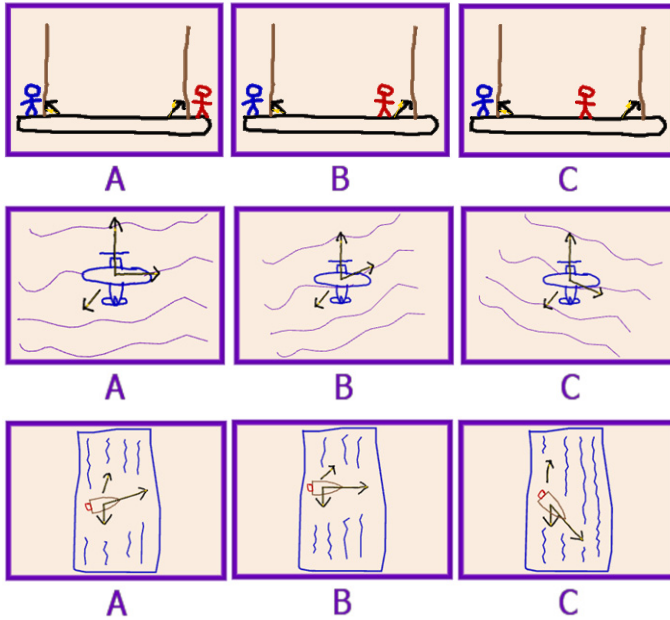


Fig. 3. Three example ranking problems without numerical values. The upper problem requires ranking the tensions in the left rope from greatest to least in the three scenarios. The middle problem requires ranking the net velocity of the airplane and the lower problem requires ranking the net velocity of the boat. No numerical values are associated with the vector arrows, so these problems must be solved visually.

potential participants of this model fragment and the conditions are criteria for instantiating the model (or considering it to be currently active). In this model, the object must be a physical object and the vectors must be represented as a set. The constraints consist of three statements. The first ensures that all the vector arrows involved in this model fragment describe the object we are interested in and that they are all the same type (to avoid summing vectors that have different meanings and different units). The second constraint calculates the sum of the vector arrows in 2D space. This operation is purely spatial and uses the *parallelogram rule* as described in the textbook: the sum of two vectors can be represented with an arrow along the diagonal of the parallelogram partially formed by the two vectors. The arrow is displayed in CogSketch and it is bound to the resultant vector variable (`?resultant-vector`) of the model fragment. With these two constraints satisfied, candidates are found for all three of the model fragment participants. The last constraint finds the quantity denoting function that corresponds with the vector type. For instance, if force arrows are being summed, then the net quantity function would be `NetForceMagFn`. Using the relationship between vector type and quantity function allows us to use one model fragment type for any kind of vector addition, rather than defining separate models for force arrows, velocity arrows, and acceleration arrows. Most importantly, the consequences of the model fragment are what give it inferential

power. In this case, the model implies that there is a positive qualitative proportionality between the magnitude of the net vector of the object and the length of the resultant vector arrow. This relationship allows the system to make judgments about vector quantities (whose exact values are unknown) via visual quantities (whose values can be computed by CogSketch). Note that this model fragment is by no means exhaustive, but this type of partial information can be very useful for solving conceptual physics problems. In our domain theories, we developed model fragments for vector addition, hanging systems with tensile objects, and conceptual knowledge of gravity.

Participants

```
?object (Physob)
?vectors (Set)
?resultant-vector (VectorInterval)
```

Participant Constraints

```
(commonVectorsForObject ?object ?vectors ?vector-type)
(visualVectorSum ?vectors ?resultant-vector)
(netQuantityFnFor ?vector-type ?net-quantity-fn))
```

Conditions

```
(hasQuantity ?object (?net-quantity-fn ?object ?axis)))
```

Consequences

```
(qEqualTo
  (?net-quantity-fn ?object ?axis)
  ((QPQuantityFn VectorMagnitude) ?resultant-vector))
(qprop
  (?net-quantity-fn ?object ?axis)
  ((QPQuantityFn Length) ?resultant-vector))
```

Fig. 4. A model fragment describing vector addition, using simplified syntax for clarity

3.3 Differential Qualitative Analysis via Causal and Spatial Reasoning

Our implementation of differential qualitative analysis (DQA) involves two main steps: (1) perform QP analysis of the scenarios, (2) compare the goal quantity across the different scenarios.

Step 1 is done using traditional model formulation and influence resolution techniques on the sketched diagrams. The domain theories are searched for applicable model fragments for each scenario in the problem. For each model fragment, the system attempts to find participants that satisfy the constraints. This usually involves conceptual reasoning about the category membership of potential participants, properties of physical scenarios, and spatial reasoning. For example, in the model fragment shown in Figure 4, the constraints involve visually computing the sum of two or more vector arrows. If the participant constraints and conditions are

satisfied for a particular model fragment, it is instantiated and its consequences (which can be causal influences or ordinal relationships between quantities) are inferred. Inferences about quantities are particularly important because they can be used to determine the ordinal relationships of quantities across different scenarios. Once models have been formulated, influence resolution determines if any of the quantities are changing and propagates those changes through the causal chain of influences. The model fragments that are active for a particular scenario represent that scenario's causal model.

In step 2, the goal of DQA is to determine how a quantity changes across two scenarios. Symbolically, this is represented by the following predicate:

```
(dqValue ?quantity ?sme-mapping ?value)
```

where `?quantity` is the goal quantity of the ranking problem (e.g. mass or net force magnitude), `?sme-mapping` is an analogical mapping between two scenarios (i.e. a base scenario and a target scenario), and `?value` is one of four possible values: *-1*, *0*, *1*, *Ambig*. The `dqValue` represents the qualitative difference between the two quantities from the base scenario to the target scenario. Therefore, a `dqValue` of *-1* means that the quantity is lesser in the base scenario than in the target scenario. A `dqValue` of *1* means that the quantity is greater in the base scenario than in the target scenario. A `dqValue` of *0* means that the quantity is unchanged, and a `dqValue` of *Ambig* means that the difference is known to be ambiguous.

There are four strategies that are used to derive the `dqValue` of a quantity with respect to the analogical mapping. The `dqValue` may be calculated via numbers, via visual quantities, via ordinal relationships, and/or via causal influences. Calculating the `dqValue` via numbers is the most basic strategy and is used when numerical values for each quantity are known in advance or are derived through some other computation. Calculating the `dqValue` via visual quantities means that the exact quantity is retrieved from a sketched annotation (e.g. a force arrow with a numerical value) or through a spatial analysis of the sketch (e.g. a distance calculation). The exact quantities that are visually derived can then be compared directly. Calculating the `dqValue` via ordinal relationships means that there is an explicit ordinal relationship between the two quantities in the logical environment. For example, the mass of Jupiter may not be known, but the mass of Jupiter relative to the mass of the Moon may be known in one of the domain theories. Lastly, calculating the `dqValue` via causal influences is used when there are one or more known causal antecedents to the quantity being examined. That is, when causal influences are inferred as a result of QP analysis, then a `dqValue` is recursively sought for the causal antecedent of the quantity. Once a `dqValue` is found, it is propagated back up the causal chain to determine the `dqValue` of the original quantity.

Using causal influences turns out to be critical for many of the problems that we analyzed and most (10 out of 12) bottom out at some kind of spatial computation. In some cases, distance was a causal antecedent of the goal quantity and calculating geometric distance between objects was required to solve the problem. In other cases, calculating the length of a resultant vector arrow was required to causally infer the `dqValue` of the goal quantity. Even in simple net force problems, where no spatially

computed quantity was an explicit antecedent to the goal quantity, other spatial information, such as the major axis of force vectors, were needed to solve the problem. Thus, both causal and spatial information played critical roles in solving these problems.

3.4 Example

To illustrate our approach, consider the problem shown in Figure 5. The scenarios show a tree stump with two ropes pulling with equal force magnitudes but different directions. The goal of the problem is to rank the net force magnitude of the stump from greatest to least.

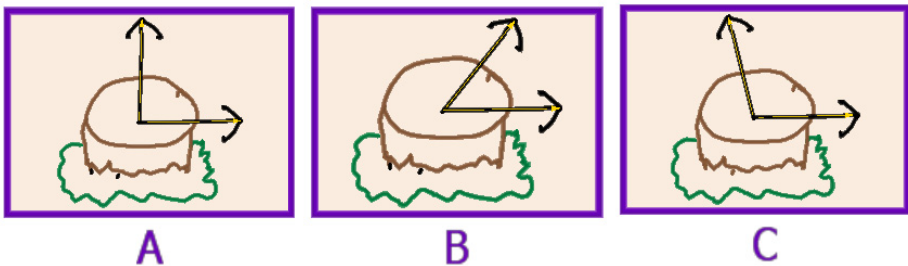


Fig. 5. An example ranking problem with three scenarios

First, the system searches for any applicable model fragments. Each of the three scenarios satisfies the constraints and conditions for the model fragment shown in Figure 4, so that model fragment is instantiated. Since one of the participants of the model fragment is a vector sum, that vector sum is calculated and represented with resultant force arrow. Once the model fragment is instantiated, its consequences are inferred and the magnitude of the net force on the stump now has a causal relationship with the length of the resultant force arrow.

Next, differential qualitative analysis begins by attempting to find differences in the net force between the three scenarios. Recall that there are four strategies for determining the dqValue of a quantity: via numerical values, visual quantities, ordinal relationships, and causal influences. The problem does not provide numerical values or ordinal relationships for the net forces, so those strategies do not succeed. Even though a resultant force arrow has been drawn, there is no numerical value associated with it, so finding a dqValue via visual quantities does not succeed either. The algorithm attempts to find differences in causal antecedents to the net force and finds one of those antecedents is the length of the resultant force arrow. DQA begins again, except this time the goal quantity is the length of the resultant force arrow. This can be measured visually, and it is found that B has the longest resultant force arrow. Since there is a positive indirect influence (qprop) between the length of the resultant arrow and the magnitude of the net force, B is consequently found to have the largest net force. This is an example of a problem that is ultimately solved using some kind of visual analysis or computation.

The problem shown in the upper portion of Figure 3, which shows a scaffold hanging by two ropes, also uses visual analysis, both in understanding the scenario before and after formulating a causal model. A model fragment describing a hanging system at mechanical equilibrium is used to solve that problem. One of the requirements of that model fragment is determining which objects are supported by the system. The fact that the scaffold hangs from the ropes is given explicitly to CogSketch via a relation arrow. Using that given information, spatial information from CogSketch, and force propagation rules in qualitative mechanics, our algorithm determines which objects are supported by the system (i.e. the two people and the scaffold itself). This characterization of the scenario enables a model fragment to be used for further inference. One of the inferences is that the tensions in the ropes are negatively influenced (via negative qualitative proportionality) by their distance to the center of mass of the whole hanging system. Using the same causal reasoning as in the previous problem, the algorithm finds that the center of mass is closest to the left rope in scenario C, and that is therefore the scenario with the greatest tension in the left rope.

4 Analysis of Problems

Using the approach described above, our system was able to solve all 12 ranking problems. Out of the 12 that were solved, 10 required spatial reasoning of some kind to solve the problem, such as a distance calculation between two objects or major axis detection.

Table 1. Summary of ranking problems and the spatial reasoning required to find a difference between the goal quantities in the scenarios

<i>Problem Category</i>	<i>N</i>	<i>Solved</i>	<i>Spatial Computation Required</i>
Vector Addition	6	6	Major axis detection Parallelogram rule for vector addition Arrow interpretation Arrow length
Gravity	2	2	N/A
Tension	4	4	Distance Approximate center of mass

In addition to the spatial computations listed in Table 1, qualitative spatial relations played an important role in the initial understanding of the sketch, before a causal model is even formulated. Surface contact detection, for example, is used by qualitative mechanics to determine how forces transfer between objects. Spatial information is also used to guide the comparison process via analogy because it shapes how items are aligned to each other.

The problems that were not analyzed in this experiment required domain models that have not yet been implemented. This includes models of friction and models that make explicit inferences over scenarios that span multiple qualitative states. For instance, problems that require estimating the velocity of an object after a certain

period of time cannot be solved by our system currently. However, we plan on using other representations (e.g. encapsulated histories [19]) to capture these types of problems.

5 Related Work

Our work is very closely related to work on everyday physical reasoning problems [22]. However, while Klenk and colleagues used case-based reasoning to infer causal models of physical scenarios, our work uses a first-principles approach to model formulation. It is likely that both approaches to model formulation would be needed to develop a system with robust qualitative physics problem solving capabilities.

Many other systems have been developed to solve physics problems, but only some of them use spatial reasoning in the problem solving process. BEATRIX [23] and Figure Understanterder [24] used information from text and diagrams, but the diagrams were created with graphical diagram tools and are therefore not subject to the same qualitative spatial reasoning requirements of hand-drawn sketches. Work by Lockwood et al. [25] answered questions about information learned from a combination of text and sketches, but comparative analyses were not explored.

In the field of intelligent tutoring systems, Atlas-Andes and AutoTutor [26-28] incorporate diagrams to help students but they are not spatially analyzed by the software to solve problems. Given the integration of qualitative and quantitative information demonstrated in this paper, it is possible to accommodate both types of representations in general problem solving and models of student problem solving.

6 Conclusion and Future Work

In this paper we have shown that CogSketch's visual and spatial representations and reasoning, combined with QP theory and Differential Qualitative Analysis, can be used to solve diagrammatic ranking problems from a popular physics textbook. Out of the 12 problems examined and solved, 10 could not be solved without spatial reasoning of some kind.

One of the challenges with creating these representations is delineating spatial and conceptual knowledge. All problems assume some level of common sense knowledge. Since we are using knowledge base contents derived from OpenCyc, there is a wealth of potentially relevant material available, but determining what concepts are relevant for a particular scenario is difficult. For example, there are concepts that help formalize the notion that a scaffold *hangs from* two ropes or that a plane *moves with* the wind, which we sketched as relationship arrows. However, we are currently working on models of cords and strings (and other flexible objects) that will be able to automatically infer that something hangs from something else based purely on visual and conceptual reasoning. This would reduce the amount of extra knowledge that is manually given in our sketches.

Another important consideration is the level of detail in model fragment types. Model fragment types should be as general as possible to reduce the number of new

models needed for new problems. For example, all four tension problems were solved with the same model of hanging systems at equilibrium. This indicates that the model is general enough to handle different situations, yet it is likely that it will need to be extended to handle more complex hanging systems (e.g. pulleys, three or more tensile objects).

The problems that have not yet been solved by the system indicate other areas for future work. A fuller domain theory of conceptual physics is needed to model many other topics and other aspects of qualitative reasoning (i.e. encapsulated histories, limit analysis) will be needed to address problems that span multiple qualitative states [19, 29, 30].

These results add to the evidence that spatial reasoning is an important factor in physics problem solving and physics instruction. Notably, the spatial reasoning requirements appear to be greater in these qualitative, conceptual problems, which has implications for tutoring systems that aim to improve conceptual physics understanding. It is therefore important to continue to develop systems that can integrate spatial information into problem solving strategies in STEM domains.

Acknowledgements. This work was supported by the Spatial Intelligence and Learning Center (SILC), an NSF Science of Learning Center (Award Number SBE-1041707).

References

1. Tversky, B.: What do sketches say about thinking. In: 2002 AAAI Spring Symposium, Sketch Understanding Workshop, Stanford University, AAAI Technical Report SS-02-08, pp. 148–151 (2002)
2. Larkin, J.H., Simon, H.A.: Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science* 11, 65–100 (1987)
3. Suwa, M., Tversky, B., Gero, J., Purcell, T.: Seeing into sketches: Regrouping parts encourages new interpretations. In: Gero, J., Tversky, B., Purcell, T. (eds.) *Visual and Spatial Reasoning in Design II*, University of Sydney, Sydney, pp. 207–219 (2001)
4. Ainsworth, S., Prain, V., Tytler, R.: Drawing to learn in science. *Science* 3, 5 (2011)
5. Wai, J., Lubinski, D., Benbow, C.P.: Spatial ability for STEM domains: Aligning over 50 years of cumulative psychological knowledge solidifies its importance. *Journal of Educational Psychology* 101, 817 (2009)
6. Chaudhri, V.K., Clark, P.E., Mishra, S., Pacheco, J., Spaulding, A.: *Aura: Capturing knowledge and answering questions on science textbooks* (2009), <http://www.ai.sri.com/pubs/files/1768.pdf>
7. Hestenes, D., Wells, M., Swackhamer, G.: Force concept inventory. *The Physics Teacher* 30, 141 (1992)
8. Hewitt, P.: *Conceptual Physics*. Pearson-Addison Wesley, Boston (2010)
9. Forbus, K., Usher, J., Lovett, A., Lockwood, K., Wetzell, J.: Cogsketch: Sketch understanding for cognitive science research and for education. *Topics in Cognitive Science* 3, 648–666 (2011)
10. Cohn, A.G., Bennett, B., Gooday, J., Gotts, N.M.: Qualitative spatial representation and reasoning with the region connection calculus. *Geoinformatica* 1, 275–316 (1997)

11. Lovett, A., Kandaswamy, S., McLure, M., Forbus, K.: Evaluating Qualitative Models of Shape Representation. In: 26th International Workshop on Qualitative Reasoning (2012)
12. Falkenhainer, B., Forbus, K.D., Gentner, D.: The structure-mapping engine: Algorithm and examples. *Artificial intelligence* 41,1-63 (1989)
13. Yin, P., Forbus, K.D., Usher, J.M., Sageman, B., Jee, B.D.: Sketch Worksheets: A Sketch-Based Educational Software System. In: *Innovative Applications of Artificial Intelligence, IAAI* (2010)
14. Wetzel, J., Forbus, K.: Integrating Open-Domain Sketch Understanding with Qualitative Two-Dimensional Rigid-Body Mechanics. In: *Proceedings of the 22nd International Workshop on Qualitative Reasoning* (2008)
15. Kim, H.: Qualitative reasoning about fluids and mechanics. Ph.D. dissertation and ILS Technical Report, Northwestern University, Evanston, IL (1993)
16. Nielson, P.E.: A qualitative approach to rigid body mechanics. Tech. Rep. No. UIUCDCS-R-88-1469; UILU-ENG-88-1775, University of Illinois at Urbana-Champaign, Department of Computer Science (1988)
17. Forbus, K.D.: Qualitative process theory. *Artificial intelligence* 24, 85–168 (1984)
18. Friedman, S.E., Forbus, K.D.: Repairing incorrect knowledge with model formulation and metareasoning. In: *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence*, pp. 887–893 (2011)
19. Klenk, M., Forbus, K.: Analogical model formulation for transfer learning in AP Physics. *Artificial intelligence* 173, 1615–1638 (2009)
20. Weld, D.S.: Comparative analysis. *Artificial intelligence* 36, 333–373 (1988)
21. Chang, M.D., Wetzel, J., Forbus, K.D.: Qualitative and Quantitative Reasoning over Physics Textbook Diagrams. In: 25th International Workshop on Qualitative Reasoning (2011)
22. Klenk, M., Forbus, K.D., Tomai, E., Kim, H., Kyckelhahn, B.: Solving everyday physical reasoning problems by analogy using sketches. In: *AAAI*, p. 209 (2005)
23. Novak, G.S., Bulko, W.C.: Understanding Natural Language with Diagrams. In: *AAAI*, pp. 465–470 (1990)
24. Rajagopalan, R.: Picture semantics for integrating text and diagram input. *Artificial intelligence Review* 10, 321–344 (1996)
25. Lockwood, K., Forbus, K.: Multimodal knowledge capture from text and diagrams. In: *Fifth International Conference on Knowledge Capture*, pp. 65–72. ACM (2009)
26. Graesser, A.C., Chipman, P., Haynes, B.C., Olney, A.: AutoTutor: An intelligent tutoring system with mixed-initiative dialogue. *IEEE Transactions on Education* 48, 612–618 (2005)
27. VanLehn, K., Jordan, P.W., Penstein Rosé, C., Bhembe, D., Böttner, M., Gaydos, A., Makatchev, M., Pappuswamy, U., Ringenberg, M.A., Roque, A.C., Siler, S., Srivastava, R.: The architecture of why2-atlas: A coach for qualitative physics essay writing. In: Cerri, S.A., Gouardères, G., Paraguaçu, F. (eds.) *ITS 2002. LNCS*, vol. 2363, pp. 158–167. Springer, Heidelberg (2002)
28. VanLehn, K., Lynch, C., Schulze, K., Shapiro, J.A., Shelby, R., Taylor, L., Treacy, D., Weinstein, A., Wintersgill, M.: The Andes physics tutoring system: Five years of evaluations. In: *12th International Conference on Artificial Intelligence in Education*, pp. 678–685. IOS Press (2005)
29. Klenk, M., Forbus, K.: Exploiting persistent mappings in cross-domain analogical learning of physical domains. *Artificial Intelligence* 195, 398–417 (2013)
30. Pisan, Y.: An Integrated Architecture for Engineering Problem Solving. Doctoral Dissertation, Northwestern University, Evanston, IL. UMI No. 733042431 (1998)