On the Representation of Physical Quantities in Natural Language

Sven E. Kuehne
Qualitative Reasoning Group
Northwestern University
1890 Maple Avenue, Ste. 300
Evanston, IL 60201, USA
skuehne@cs.northwestern.edu

Abstract

In this paper we investigate the forms in which continuous properties can appear in written natural language. Our focus is on physical quantities found in descriptions of physical processes, such as expansion, movement, or transfer. Using Qualitative Process Theory as our underlying formalism, we show how information extracted from natural language text corresponds to the five constituents of physical quantities, i.e. entities, quantity types, values, unit, and the sign of derivative. Furthermore, based on the analysis of a corpus of descriptions of physical processes we develop two hypotheses that allow us to distinguish between two forms in which adjectives and adverbs can be used as references to physical quantities.

1 Introduction

Ordinary people know a lot about the physical world around them. They know that water will eventually boil if you heat it on a stove, that a ball placed at the top of a steep ramp will roll down, and that a cup will overflow if you continue to pour coffee in it. When people talk and write about such phenomena in everyday language, references to continuous properties are usually part of these descriptions. From simple utterances like *"The coffee is hot"* to a more complicated comparison like *"The velocity of gas molecules is higher than the velocity of molecules in a liquid."* being able to identify and extract the information about physical quantities is essential to understand these sentences. Using Qualitative Process Theory (Forbus, 1984) as the underlying formalism, we investigate the forms in which continuous properties can appear in written natural language. Our focus is on *physical* quantities found in descriptions of *physical* processes, such as expansion, movement, or transfer.[1]

The way in which continuous parameters and processes are described in natural language is not accidental, even though the Qualitative Reasoning community has often assumed these correspondences in their representations of knowledge. Since Qualitative Process Theory is a formalism of how people reason about the physical world, the basic ideas of the theory should be reflected in the language that people use to communicate their understanding of physical phenomena. This paper shows that the natural language descriptions of physical processes contain abundant information about the constituents of physical quantities. Moreover, the results of this study can be used in a variety of applications, such as grammatical rules of a parser or in the design of information extraction algorithms.[2]

2 Physical quantities

In Qualitative Process Theory, all physical changes in continuous properties are caused by *physical processes*. The identification of continuous parameters is therefore an essential step in the extraction of information about physical processes from natural language text. In (Kuehne & Forbus, 2002) we presented a scheme for the extraction process that uses FrameNet-compatible representations (Baker, Fillmore, & Lowe, 1998; Fillmore, Wooters, & Baker, 2001) to capture information about physical processes. The examples presented draw from the same corpus material used in our previous analysis (Buckley, 1979; Maton et al., 1994; Moran & Morgan, 1994), although some of the sentences have been shortened and simplified.

This paper presents an overview of how information about continuous properties can appear in natural language, and the ways in which this information corresponds to the following five constituents of physical quantities:

- The *Entity* is a uniquely named object or an instance of a process associated with the quantity. For example, the word 'brick' in the noun phrase 'the temperature of the brick' is an entity.[3]
- The *Quantity Type* specifies the kind of parameter. The word 'temperature' in the noun phrase 'the temperature of the brick' is a reference to a quantity type.
- The *Value* specifies the numerical or symbolic value of the property. The number '3' in '3 liters of water'

[1] The findings of this analysis are applicable to other types of quantities as well. Abstract and conceptual quantities are often referred to metaphorically by words with a physical basis and require a different semantic interpretation. 'The price is hot.' is does not have anything to do with temperature, unlike 'The water is hot.' However, the techniques for the extraction of information about such quantities are essentially the same.

[2] Although the results of the analysis are used by us in a particular way, for the purpose of generality, we do not limit the findings presented here to any particular type of grammar or pattern language.

[3] The noun 'brick' actually refers a particular individual, maybe 'brick32', not the collection of all bricks.

or the adjective 'hot' in 'the hot ground' are values associated with a quantity.

- The *Unit* specifies the physical units of the property. Example: The word 'kilograms' in '3 kilograms of lead.' Units usually appear in combination with a numerical value or with a quantifier.
- The *Sign of the Derivative* specifies how the parameter is changing. In the sentence "The temperature is increasing." the sign of the derivative is expressed by the word 'increasing', which indicates that the parameter is changing in a positive direction.

Only the first two of these five constituents are required to identify a physical quantity. The quantity type together with the entity are sufficient to talk about quantities like 'the temperature of a brick' or the 'the flow rate of heat'. Values, units, and information about changes are optional and often not explicitly stated.

Entities and quantity types can be named by unique labels. These labels are usually introduced together with the noun, e.g. 'the brick B1' or 'the pressure P32'. After it has been introduced, the label can than be used on its own, acting as a discourse variable that refers to the entity or quantity type. The patterns presented in this analysis do not use such labels, but are applicable to named entities and quantity types as well.

3 Physical quantities in NL text

Descriptions of physical phenomena necessarily need to mention physical quantities. Talking about a physical process without references to continuous parameters is difficult and in most cases impossible. In this section we will show some forms which are commonly used in natural language descriptions to express information about physical quantities. Most of the examples presented in this analysis come from the same corpus that we used in (Kuehne & Forbus, 2002), the analysis is mainly concerned about the different parts of speech in which information about physical quantities appears, not about the syntactic constructs represented by particular sentences.[4]

3.1 Explicitly referenced quantities

Natural language text can refer to physical quantities either directly or indirectly, depending on whether the type of the quantity is explicitly mentioned in the sentence. *Explicit references* to quantities can be found in nouns, verbs, and adjectives that are morphologically related to quantity types.

3.1.1 Nouns

The quantity type can be explicitly mentioned as a noun, together with one or more entities that it is associated with.

(1) VOLUME flows from the *can* to the *ground*.
(2) The TEMPERATURE of the *brick* is rising.

Sentence 1 contains information about two physical quantities, the volume of some substance in the can and on the ground. The quantity type 'volume' is associated with both locations, i.e. the 'can' and the 'ground'. In (2) the quantity type 'temperature' is associated with a single entity.

3.1.2 Compound nouns

The quantity type can also appear as the head of a compound noun. The remaining constituents of the compound noun can be treated as information about a specialization of the quantity type. For example, in (3) the quantity type 'radiation heat' is a specialization of 'heat'; in (4) 'heat energy' is a type of 'energy'.

(3) RADIATION HEAT originates from the *heater.*
(4) The HEAT ENERGY of the *water* increases.

3.1.3 Verbs

Verbs can refer to events as well as to quantity types associated with these events.[5] The verb in sentence 5 appears as a direct reference to the quantity type 'length'. Sentence 6 is slightly more complicated, because it allows two different interpretations. The obvious interpretation is to treat the verb as an explicit reference to a quantity, as it is in (5). In this case, the quantity type 'heat' is tied to both entities, the stove as the source of the heat flow and the kettle as the destination of the heat flow.

(5) The press LENGTHENS the *iron beam*.
(6) The stove HEATS the *kettle*.

Alternatively, sentence 6 can be interpreted as an increase in temperature of the kettle caused by the stove. Even though the quantity type 'temperature' is not mentioned in (6), we might infer that heating the kettle also increases the temperature of the kettle. This is an inference that most readers of such a descriptions draw, and it coincides with the kind of conclusions that are supported by QP Theory.

3.1.4 Adjectives

Certain adjectives can refer to quantity types directly, if the adjective is morphologically related to a quantity type. For example, in (7) the adjective 'denser' refers to the quantity type 'density'. The quantity type in this sentence is associated with both entities, the subject 'iron' and the object 'wood'. The quantity type referenced in (8) by the adjective 'deep' is 'depth' and associated with the noun 'pit'.

[4] The goal of this analysis is to highlight the forms in which information about physical quantities can appear in natural language, instead of defining a specific grammar for descriptions of physical quantities.

[5] Events such the increase or decrease of a parameter, e.g. the temperature of a brick, can be involved in an instance of a physical process. For one linguistic perspective on actions, processes, and events, see (Parsons, 1990).

(7) *Iron* is DENSER than *wood*.
(8) The DEEP *pit* is covered with dirt.

3.2 Implicitly referenced quantities

While the quantity types in explicitly referenced quantities are usually easy to determine, *implicit references* to quantities are more difficult to figure out. Implicitly referenced quantities do not mention a quantity type. Instead, the reader has to use the contextual information provided by the sentence as well as available background knowledge. The following section shows how nouns, verbs, adjectives, and adverbs can determine a quantity that is not explicitly mentioned in a sentence.

3.2.1 Verbs

A quantity type can be implicitly referenced by a verb that describes a physical process, e.g. movement, expansion, or transfer. The sentence in which the verb occurs usually provides additional contextual information for the interpretation of the implicitly referenced quantity.

(9) As the temperature rises, the *liquid* EXPANDS.
(10) The *ball* ROLLS against the wall.

The verb 'expand' in (9) indicates that something is changing in one or more physical dimensions, i.e. in length, area, or volume. For the three-dimensional entity 'liquid' the appropriate quantity type is therefore 'volume'. The verb also includes implicit information about a positive change in the quantity, i.e. an increase in volume of the liquid, which we will address in section 6.

3.2.2 Adjectives

The quantity type can be implicitly referenced by certain adjectives. For example, the quantity type described by the adjective 'hot' in (11) is 'temperature'. The comparative also encodes the ordinal relationship between the quantities associated with the two entities, i.e. the fact that the temperature of the stone is greater than the temperature of the water. Similarly, the quantity type expressed by 'lighter' in (12) is 'weight'.

(11) The *stone* is HOTTER than the *water*.
(12) The *upper air masses* are LIGHTER than the *lower air masses*.

For a correct interpretation the relationship between the adjective and the associated quantity type has to be known. The fact that the adjective 'hot' is associated with 'temperature' is a fact learned by a human reader and is usually provided as background knowledge in an NLP system.

3.2.3 Verb/Adverb combination

Quantity types can also be determined by combining verbs and adverbs. The quantity type referenced in (13) is the rate of movement, or 'velocity'. The adverb alone is not sufficient to determine the quantity type. Although 'faster' is generally associated with velocity, it just qualifies the rate of change, i.e. that something is happening in less time. There are cases in which the quantity type referenced by 'faster' is not velocity. For example, 'expanding faster' in (14) refers to the rate of expansion.

(13) The *gas molecules* are MOVING FASTER than *molecules in a solid*.
(14) *Liquid A* is EXPANDING FASTER than *liquid B*.

All these cases have one thing in common: the referenced quantity is a rate, most likely associated with a process referenced by the verb ('movement', 'expansion', 'decay').

3.2.4 Noun/Verb combination

This type of implicitly referenced quantity uses a noun/verb combination to refer to the rate of change of a quantity.

(15) [The greater the thermal resistance,] the HARDER the *heat* FLOWS.
(16) [The less heat is supplied,] the SLOWER the *temperature* RISES.

The quantity type in (15) is not 'heat' but the flowrate of heat. The combination of 'flows' and 'heat' determines the quantity type, while the combination of 'flows' and 'harder' gives the direction of change. Sentence 16 looks similar to (15) but differs in an important domain-specific way: temperature is not an extensive property, i.e. temperature cannot be added directly to an object. The quantity type referenced in (16) is the rate of change in temperature, resulting from a change in the amount of heat.

3.2.5 Noun/Adjective combination

The quantity type is only implicitly referenced by a combination of a noun and an adjective.

(17) The BIGGER the *surface*, [the more heat is absorbed.]

The quantity type in (17) is the size of the surface (not the surface itself) associated with an unnamed participant or the size of a participant 'surface'. The adjective 'bigger' refers to the quantity type 'size' (or 'area'). It also encodes a change of the quantity, i.e. an increase in surface area.

As in Verb/Adverb combinations, the adjective determines the referenced quantity type. For example, replacing 'bigger' with 'shinier' will change the resulting quantity type from 'area' to 'reflectiveness'. In the following section we will investigate the roles of adjectives and adverbs in determining implicitly referenced quantities in more detail.

4 A closer look at adjectives and adverbs

Changing the adverb in a Verb/Adverb combination or the adjective in a Noun/Adjective combination for an implicitly referenced quantities can change the interpretation of the underlying quantity type. Similarly, comparisons between two quantities can be presented by explicitly mentioning the quantity type or by using an adjective or adverb as an indirect reference to the quantity type. Another important distinction can be made about how adjectives and adverbs encode information about references to physical quantities, i.e. whether they are tied to a specific type of quantity or are neutral in regard of a quantity reference.

Research on the lexical semantics of adjectives has tried to establish taxonomies for the different semantic categories of adjectives (see Raskin & Nirenburg (1995) for an overview). Several of these taxonomies focus on the class of adjectives that we are most interested in for extracting information about physical quantities, i.e. qualitative (scalar, gradable) adjectives (Dixon, 1991; Frawley, 1992). However, none of these taxonomies has been used in any practical application until now (Raskin & Nirenburg, 1995).

From our perspective, using the semantics of Qualitative Process Theory, the taxonomies suggested by Dixon and Frawley are flawed and inconsistent. The breakup of types and subtypes appears to be arbitrary, because several of the types of quantities can be collapsed into a single type. In Dixon's taxonomy the adjectives of the 'speed' and 'physical property' types are separated from those classified as 'dimension'. Similarly, 'age' and 'value' are listed as separate types, while they could actually be treated as a single kind of quantity. Furthermore, labeling one of types a 'dimension' and another 'physical property' is misleading and incorrect. All dimensions are quantity types, but not all quantity types are dimensions.

We are not going to redefine any of the existing taxonomies. Instead, we are dividing the class of qualitative adjectives into two distinct subclasses: quantity-specific and quantity-neutral adjectives.

4.1 Quantity-specific adjectives and adverbs

Quantity-specific adjectives and adverbs encode information about an implicitly referenced quantity, i.e. the adjective or adverb determines (sometimes in combination with a noun or verb) the quantity type. Sentences that use quantity-specific adjectives and adverb do not contain explicitly referenced quantities. The information about the quantity type is encoded in the adjective or adverb itself and needs to be retrieved from there.

(18) The stone is HOTTER than the water.

The comparative 'hotter' in (18) refers to the quantity type 'temperature' that is associated with the two entities 'stone' and 'water'. The reader has to know that the adjective 'hot' is associated with 'temperature' or otherwise the interpretation of the sentence would fail.[6] Furthermore, the adjectives used in these types of comparison have to be gradable. Explicit quantity references in comparisons are usually not combined with quantity-specific adjectives, as it is illustrated by sentence 19.

(19)* The *temperature* of the stone is HOTTER than the *temperature* of the water.

The additional information about the quantity type originating from the explicit reference would be considered redundant, when the quantity-specific adjective and the noun refer to the same quantity type.[7]

The fact that sentences with quantity-specific adjectives and adverb do not contain directly referenced quantity types for physical quantities of the same type can be summarized in the following hypothesis:

Exclusion Hypothesis: *The use of explicitly referenced quantities and quantity-specific adjectives and adverbs (for the same entity and referring to the same quantity type) is mutually exclusive.*

An important aspect of the use of quantity-specific adjectives and adverbs is their interaction with nouns and verbs, and the quantity type that is referenced by the combination of them. Using different quantity-specific adjectives and adverbs with the same noun or verb changes the implicitly referenced quantity type.[8]

(20) *Gas molecules* are MOVING FASTER than *molecules of a liquid*.

In (20), the adverb 'faster' together with the main verb determines the type of the indirectly referenced quantity. The adverb 'fast' is quantity-specific, while the main verb of the sentence is not. The combination of 'to move' and 'fast' refers to the quantity type 'velocity'. However, replacing the quantity-specific adverb with another adverb of the same category will result in references to different quantity type, as demonstrated in the following variations of (20).

(21) *Gas molecules* are MOVING HIGHER than *molecules of a liquid*.

(22) *Gas molecules* are MOVING FARTHER APART than *molecules of a liquid*.

The indirectly referenced quantities in sentence 21 are the 'height' (or position) of the gas and liquid

[6] The use of the comparative also imposes an ordering on the two physical quantities, i.e. the temperature of the stone is higher than the temperature of the water.

[7] If they refer to different quantity types, the sentence would be considered quite problematic, e.g. if we try to replace 'temperature' in sentence (19) with 'weight'.

[8] With respect to Exclusion Hypothesis it is assumed that the verb or noun in the combination does not include an explicit quantity reference, i.e. it has to be quantity-neutral.

molecules, and the quantity type referred to in (22) is the 'distance' between the molecules.

4.2 Quantity-neutral adjectives and adverbs

Comparisons between two quantities do not always have to use quantity-specific adjectives and adverbs in their comparative form. Another class of adjectives and adverbs does not carry any quantity-determining information and is therefore labeled as *quantity-neutral*. Sentences with adverbs and adjectives of this class need to reference the quantity directly, because the quantity-neutral adjective or adverb does not contribute any information to determine the quantity type. In sentence 23 the quantity type ('temperature') is explicitly mentioned for both entities (the 'food' and the 'plate'). The comparison is done by a quantity-neutral adjective 'high'.[9]

(23) The *temperature of the food* is HIGHER than the *temperature of the plate*.

The direct reference to the quantity type used in the comparison does not need to be included in a noun phrase with the entities. A common form of comparison references the quantity type as a part of a noun phrase that includes a quantity-neutral adjective.

(24) The *tub* has a GREATER *volume* than the *can*.

The explicitly referenced quantities in sentence 24 are the 'volume' of the 'tub' and the 'can'. The direct reference to 'volume' applies to both entities in this pattern. The quantity-neutral comparison does not contribute any information to determine the quantity type; it just determines the ordering between the two quantities.

4.3 Transformation

The sentences in the previous two sections referred to physical quantities by using either a combination of explicit references to quantities and quantity-neutral adjectives and adverbs, or a combination of implicit quantity references and quantity-specific adjectives and adverbs. According to the Exclusion Hypothesis, these combinations are mutually exclusive for the same quantity type and the same associated entities. The following example illustrates how sentences containing quantity-specific adjectives and adverbs (25, 26) can be rewritten in quantity-neutral forms (27, 28).

(25) The *stone* is HEAVIER than the *brick*.
(26) The *food* is HOTTER than the *plate*.

(27) The *temperature of the stone* is GREATER than the *temperature of the water*.
(28) The *stone* has a GREATER *weight* than the *brick*.

These examples suggest that sentences with quantity-specific adjectives and adverbs can be changed into their quantity-neutral counterparts, and that sentences with quantity-neutral adjectives and adverbs with explicit references to quantities in nouns have an alternative quantity-specific form. The following hypothesis captures this relationship between the two different uses of quantity-specific and quantity-neutral adjectives and adverbs.

Transformation Hypothesis: *Sentences with quantity-specific constructs can be transformed into semantically equivalent quantity-neutral constructs, and vice versa.*

Being able to transform or rewrite sentences with implicit references to a quantity type into a semantically equivalent form that makes the quantity type explicit and uses only generic, quantity-neutral adjectives and adverbs, is an important step towards the creation of a simplified grammar and makes the semantic interpretation for physical quantities more transparent.

The Exclusion and Transformation Hypotheses are powerful tools in the construction of a controlled language for the constituents of Qualitative Process Theory. While the Exclusion Hypothesis eliminates (or, at least, limits) ambiguity, the Transformation Hypothesis allows the conversion of sentences into an intermediary canonical form.[10]

5 Representation of values in physical quantities

The previous sections were primarily concerned about the information about the entity and quantity type, the two mandatory constituents of physical quantities. Knowing the type of a quantity and the entity it is associated with enables us to talk and reason about it. A simple noun phrase such as 'the depth of the water' contains enough information recognize it as a physical quantity, even without having any information about a particular value the quantity might have, the unit of that value, or the direction in which the quantity is changing. On the other hand, sentences like (29) could provide information for all five constituents.

(29) The *temperature* of the *oil* is *rising* to *250 degrees Fahrenheit*.

The identification of the quantity type and the entity is just half the story when we are dealing with representations of physical quantities. The following two sections examine how values and units of quantities appear in natural language text, and how changes in quantities can be identified.

There are three common types of references to values and units that can be found in natural language text: in the context of comparisons, as symbolic labels, and as

[9] The adjective 'high' can be used either in a quantity-specific sense (referring to 'height' or 'depth' as a quantity type), or in a quantity-neutral way (in the sense of 'more' or 'greater').

[10] Although a cross-linguistic study has not been done yet, the two hypotheses appear to be true at least for Germanic languages such as English and German.

quantitative information. We will discuss values and units together because units usually appear in combination with values.[11]

5.1 Comparison

Values in the context of a *comparison* appear in sentences like "The brick is warmer than the plate." The comparison orders the quantities, i.e. the temperature of the brick is greater than the temperature of the plate. However, it does not contain exact information about the possible values of the quantities. Even though the base form ('warm') of the comparative might refer to a specific range of temperature, the exact values cannot be known or even guessed from the information provided by the sentence. The brick might be red hot, while the plate is frosted with ice. This fact becomes more explicit if the quantity-neutral form of the sentence is used, "The temperature of the brick is higher than the temperature of the plate." Replacing the comparative 'warmer' with 'hotter' will not change the ordering between the quantities or contribute any additional information for identifying an exact value.

It is impossible to determine how far the values associated with the two compared quantities are apart from each other. The only information that can be extracted from this sentence about the values of the two compared quantities is the fact, that the value of one quantity is greater than the other. With several of these comparisons along the same dimension, it is possible to identify the potential ranges of the values for particular quantities. For example, the temperature of a coffee is greater than the temperature of an ice cube, and it is lower than the temperature at the tip of a lit cigarette.

5.2 Symbolic labels

Values can also take the form of a *symbolic label* associated with an entity, e.g. "The brick is hot." Even though the exact temperature of the brick is unknown, the adjective 'hot' suggests a certain temperature range. The range might be different depending on the context of the sentence. In refrigeration 'hot' might be in a very different range of temperatures than in the context of metallurgy.

Nouns that are associated with the adjective can impose restrictions on the range of the value in certain cases. For example, (Bierwisch, 1967) compares two simple sentences, "*The room is tall*." and "*The space is tall*"' In the first sentence the noun 'room' might restrict the range of values for the height of a room to those for a typical room, e.g. between 8 and 10 feet. Without further information, this kind of assumption is more difficult to make for second sentence. Is the space a small compartment or a crawl space? Or is it the inside of a cathedral? The range of typical values would be quite different for these two cases.

The concepts of quantity-specific and quantity-neutral forms are applicable to these symbolic labels for values. Adjectives that represent a value are generally quantity-specific, as in the sentence "The brick is hot." Alternatively (and according to the Exclusion Hypothesis), a quantity-neutral form could be used to express the same fact, e.g. "The temperature of the brick is high."[12]

While adjectives and adverbs generally refer to a range of values along a dimension, natural language also uses symbolic labels to refer to concrete values, i.e. particular points along a dimension. The noun phrase 'boiling point of water' usually refers to the point where liquid water turns into steam and the value of approximately 212 degrees Fahrenheit. The noun phrase provides a label for this particular point. Note that the compound noun 'boiling point' would be an underspecified symbolic label because different substances have different boiling points. Other labels such as 'sound barrier' may not need the additional complement.

The structure for labels that describe limit points is not arbitrary. Usually the head of a noun phrase refers to a point on a scale (e.g. 'point', 'barrier'), while the noun modifier is associated with a process, a dimension, or a quantity type (i.e. 'boiling', 'sound'). These two parts are mandatory components of the label. Determining the quantity type and the dimension is difficult in many cases, e.g. we have to know that 'boiling point' is associated with 'temperature' and that 'sound barrier' actually refers to the speed of sound or velocity. Additionally, the label can take an optional complement phrase that restricts the compound noun. For example, the complement phrase 'of water' restricts the interpretation of boiling point to a particular substance. The key idea here is that the underlying mechanisms for handling limit points are essentially the same as those for symbolic references to intervals on a particular dimension.

5.3 Concrete numeric values and units

The most explicit form in which values can appear is as *quantitative information*, i.e. by using concrete numeric values and units. For example, in (30) the quantity type ('temperature') is explicitly stated, together with exact information about the numeric value ('120') and the unit ('degrees Fahrenheit').

(30) The temperature of the brick is *120 degrees Fahrenheit*.

Sentences that contain concrete numeric values and units might use quantity-specific adjectives or adverbs instead of explicit references to the quantity type.

(31) The water is 150 degrees Celsius hot.

[11] Units can appear separately from values in definitional statements, like "Length is measured in Meters." or, even more explicit, "The unit of power is the Watt."

[12] The Cyc knowledge base (Lenat & Guha, 1989) handles values in a similar way. For example, the value `#$Hot` is the result of `#$HighAmountFn` of `#$Temperature`.

The adjective 'hot' in (31) contains both an indirect reference to the quantity type 'temperature' and also provides information on which end of a scale the value is considered in the context of the sentence.

6 Representations of changes in physical quantities

The values of physical quantities cannot always be treated as static information; they will change as physical processes are active. The sign of the derivative indicates whether a quantity is changing and in which direction. This section takes a look at how changes in physical quantities are reflected in natural language.

When we talk about changes in the physical world, the most obvious choice to express them would probably by using a verb. For example, if water is flowing from one container into another, there are several ways of expressing the change of the amount of water in each container. It could be explicitly stated that the amount of water in one container is decreasing while the amount of water in the other is increasing. Alternatively, one could simply say that water is flowing from one container to another, without ever mentioning the increase and decrease in the two involved quantities.

6.1 Verbs with direct references to a quantity change

Verbs can directly refer to a change in a quantity and its direction, i.e. whether the quantity is increasing or decreasing, when the verb alone contains all the information about the change and the direction and we can therefore distinguish between verbs for positive and negative changes in quantities. For example, *gain, increase,* and *add* are verbs for positive changes, while *lose, decrease,* and *leak* are associated with negative changes.[13] Some verbs belonging to this class also allow prepositional phrase as a complement, which is restricted to the particular direction of change indicated by the verb itself (e.g. 'add to' vs. *'add from').

(32) The *brick* LOSES *heat* to the *room*.
(33) The *temperature of the water* is INCREASING.
(34) The *brick* GIVES OFF *heat*.

Some otherwise 'neutral' verbs can also fall into this class if they use specific particles to indicate a change in a quantity, as in (34).

6.2 Verbs with directional prepositional phrases

Verbs associated with Transfer and Motion event do not contain a direct reference to changes in quantity. For example, verbs like *flow* or *move* indicate a transfer of something between two physical or conceptual locations, but they do not contain information about the actual direction of the change. Instead, this information is provided by directional prepositional phrases attached to the verb. The description of the transfer can be complete when both the source and the destination are identified by prepositional phrases, as in (35), or partial when only one of the directional prepositional phrases is attached, as in (36) and (37).

(35) Heat is *transferred* FROM inside the house TO the outdoors.
(36) Energy is *moved* TO a new location.
(37) The *fan* moves heat away FROM the processor.

6.3 Verbs in combination with quantity-specific adverbs

Quantity-specific adverbs can determine the change in a quantity in conjunction with a verb. According to the Exclusion Hypothesis, the verb itself has to be quantity-neutral. Analogous to verbs with direct reference to a quantity change, the combination of verbs and quantity-specific adverbs can be associated with a decrease in a quantity, as in (38) or with an increase, as in (39).

Similar to the interpretation of the quantity type from verb/adverb combinations in section 3.2.3, there are cases in which the same adverb can refer to an increase or a decrease, depending on the verb with which it is used. For example, in the context of (39), the adverb 'faster' would indicate a positive change in the velocity of the molecules, while in (40) it will indicate an increase in the rate at which a substance dissolves.

(38) The *glass* is GETTING COLDER.
(39) The *molecules* are MOVING FASTER.
(40) The *substance* DISSOLVES FASTER.

6.4 Nouns with direct references to change

Nouns provide another way of describing changes in physical quantities. They can be divided into similar classes as verbs, i.e. nouns with direct references to a change in a quantity, and nouns that use directional prepositional phrases.

Nouns can directly refer to a change in a quantity, and analogous to section 6.1 the can be divided into nouns that refer to positive, as in (41), and negative changes, as in (42).

(41) The INCREASE in *temperature* is significant.
(42) The DECREASE in *pressure* caused a failure.

6.5 Nouns with directional prepositional phrases

Similar to verbs of the Transfer and Motion domain, the corresponding nouns will also need directional prepositional phrases to describe changes in a quantity. Again, the information about the transfer can be complete, as in (43) or partial as in (44).

[13] Another distinction could be made between verbs that can only used with extensive quantities. For example, heat can be *added*, while temperature cannot.

(43) The *flow* of oxygen FROM the tank TO the capsule is blocked.
(44) The *transfer* of heat TO the kettle has been completed.

7 Discussion

Other parts of our current research are concerned with the design of a controlled language for describing physical phenomena. One important aspect in the development of such a language is the goal to reduce possible syntactic and semantic ambiguity. The identification of patterns used for references to continuous parameters in natural language is an essential part of the semantic interpretation process, which must include the detection of directly referenced quantities as well as indirect references.

The Transformation Hypothesis provides us with a powerful mechanism for creating canonical references to physical quantities. It allows a transformation of sentences with quantity-specific adjectives and adverbs into an equivalent form that only uses generic, quantity-neutral adjectives and adverbs together with a direct reference to a quantity type. The ability to express implicit references to quantities as well as the use of quantity-specific adjectives and adverbs are important for the habitability of the controlled language.

Many quantity-specific adjectives and adverbs form opposing pairs for the same quantity type along a single dimension. For example, 'tall' is the opposite of 'short' for the quantity type 'height', and 'wide' the opposite of 'narrow' for the quantity type 'width' (see Bierwisch (1967, 1989) and Kennedy (2001) for a detailed analysis of polar adjectives). For certain quantity types we can identify not just a single opposing pair but a set of quantity-specific adjectives. For the quantity type 'temperature' we can find adjectives such as 'warm', 'cool', 'tepid', and variations such as 'lukewarm' as references besides just 'hot' and 'cold'. It is an interesting question to speculate why this variety of quantity-specific adjectives exists for some quantity types but not for others. Frequent use or familiarity with the concept 'temperature' cannot explain this fact alone. Quantity types of the length dimensions such as length, height, depth, width, or distance are also frequently used, yet they do not show the same variety of quantity-specific adjectives as the quantity type 'temperature'.

Understanding the connections between Qualitative Process Theory and natural language is important for understanding the general cognitive plausibility of qualitative models. It will also give us greater insight into how results from qualitative reasoning can be communicated back to human users in an intuitive way – by using natural language.

Acknowledgements

I would like to thank Ken Forbus and Dedre Gentner for insightful comments on this paper, as well as Praveen Paritosh and Chris Kennedy for interesting discussions on the topic. This research was supported by the Artificial Intelligence program of the Office of Naval Research.

References

Baker, C. F., Fillmore, C. J., & Lowe, J. B. (1998). *The Berkeley FrameNet Project.* In: Proceedings of the 17th International Conference on Computational Linguistics and 36th Annual Meeting of the Association for Computational Linguistics (COLING-ACL 98), Montreal, Canada.

Bierwisch, M. (1967). Some semantic universals of German adjectivals. *Foundations of Language, 3*, 1-36.

Bierwisch, M. (1989). The Semantics of Gradation. In M. Bierwisch & E. Lang (Eds.), *Dimensional Adjectives* (pp. 71-261). Berlin, Germany: Springer-Verlag.

Buckley, S. (1979). *From Sun Up to Sun Down*. New York: McGraw-Hill.

Dixon, R. M. W. (1991). *A New Approach to English Grammar, on Semantic Principles*. Oxford, England: Clarendon Press.

Fillmore, C. J., Wooters, C., & Baker, C. F. (2001). *Building a Large Lexical Databank Which Provides Deep Semantics.* In: Proceedings of the Pacific Asian Conference on Language, Information, and Computation, Hong Kong, China.

Forbus, K. D. (1984). Qualitative Process Theory. *Artificial Intelligence, 24*, 85-168.

Frawley, W. (1992). *Linguistic Semantics*. Hillsdale, NJ: Erlbaum.

Kennedy, C. (2001). Polar Opposition and the Ontology of 'Degrees'. *Linguistics and Philosophy, 24*(1), 33-70.

Kuehne, S. E., & Forbus, K. D. (2002). *Qualitative Physics as a component in natural language semantics: A progress report.* In: Proceedings of the Twenty-fourth Annual Conference of the Cognitive Science Society, George Mason University, Fairfax, VA.

Lenat, D. B., & Guha, R. V. (1989). *Building large knowledge-based systems : representation and inference in the Cyc project*. Reading, MA: Addison-Wesley.

Maton, A., Hopkins, J., Johnson, S., LaHart, D., McLaughlin, C. W., Warner, M. Q., & Wright, J. D. (1994). *Heat Energy* (annotated teacher's ed.). Englewood Cliffs, NJ: Prentice Hall.

Moran, J. M., & Morgan, M. D. (1994). *Meteorology - The Atmosphere and the Science of Weather* (4th ed.). New York, NY: Macmillan College Publishing.

Parsons, T. (1990). *Events in the Semantics of English*. Cambridge, MA: MIT Press.

Raskin, V., & Nirenburg, S. (1995). *Lexical Semantics of Adjectives: A Microtheory of Adjectival Meaning* (Technical Report MCCS-95-288). Las Cruces, NM: New Mexico State University.