Frame Semantics of Continuous Processes

Clifton McFate and Kenneth Forbus
Qualitative Reasoning Group, Northwestern University
2133 Sheridan Road, Evanston, IL, 60208, USA
{c-mcfate, forbus}@ northwestern.edu

Abstract

Qualitative Process theory provides a formal representation for human-like models of continuous processes. However, qualitative reasoning systems frequently rely on hand-made models which inhibits expansion into new domains. A representation that unifies QP theory with natural language expressions would allow these systems to expand their models by reading and interaction, thus greatly increasing their flexibility. Prior research mapped qualitative process elements onto English language constructions, but did not connect the representations to existing frame semantic resources. Here we identify and classify QP language constituents through their instantiation in FrameNet frames to provide a unified semantics for linguistic and non-linguistic representations of processes. We demonstrate that all core QP relations can map to FN, though larger QP evoking phrasal constructions do exist outside of this mapping. We conclude with a corpus analysis showing that these frames occur in natural text involving a variety of continuous processes.

Introduction & Background

Much of daily experience involves interacting with, and reasoning about, continuous processes. They can be common, such as coffee flowing into your mug, or they can be abstract, like economic growth. Despite the mathematical complexity of continuous processes, people rapidly generate predictions based on their mental models of these situations. Forbus' (1984) qualitative process theory (QP) provides a formal language for representing mental models of these continuous systems. The theory benefits from being domain general and has significant predictive power.

An important issue is bridging the gap between purely linguistic models and QP mental models, to provide a unified semantics. Doing so not only sheds light on the semantics of continuous phenomena, but also lays the groundwork for developing systems that can learn from and reason with natural language (McFate, Forbus, & Hinrichs, 2014).

Kuehne (2004) developed QP frames, a frame semantic representation based on Fillmore *et al's* (2001) FrameNet. This approach was revised and expanded by McFate *et al* (2014). While useful, both approaches suffered from limited coverage and did not connect QP frames to frame semantic resources more broadly. This paper further bridges that gap by providing a QP mapping of specific process types in FrameNet as well as their constraints, including limit points, which mark the boundaries of qualitative states. This in turn provides a broader coverage analysis of QP elements in English that also grounds them in process specific linguistic constructions.

Qualitative Process Theory

In QP theory, changes within a continuous system are always the result of processes. Causality starts with direct influences, which express the relationship between the rate of a process and the constrained quantity. A direct influence provides partial knowledge of a differential equation, where the set of direct influences must be combined to determine the derivative. Indirect influences propagate the direct effects of a process through the rest of a system by providing partial information about instantaneous (e.g. algebraic) causal relationships.

Processes and their influences are represented by *model fragments* which describe the entities that participate in a process, the conditions under which instances of it are active, and what consequences hold when active. The conditions typically include ordinal relationships, which involve a quantity and one of its limit points

QP theory provides a framework for representing mental models for many naturally occurring phenomena across a range of expertise. However, the incremental nature of language makes extracting complete QP models difficult. We turn to frame semantics to help provide the needed flexibility.

Frame Semantics

Semantic frames are conceptual schemas that relate lexical items in a sentence to their role in a semantic description

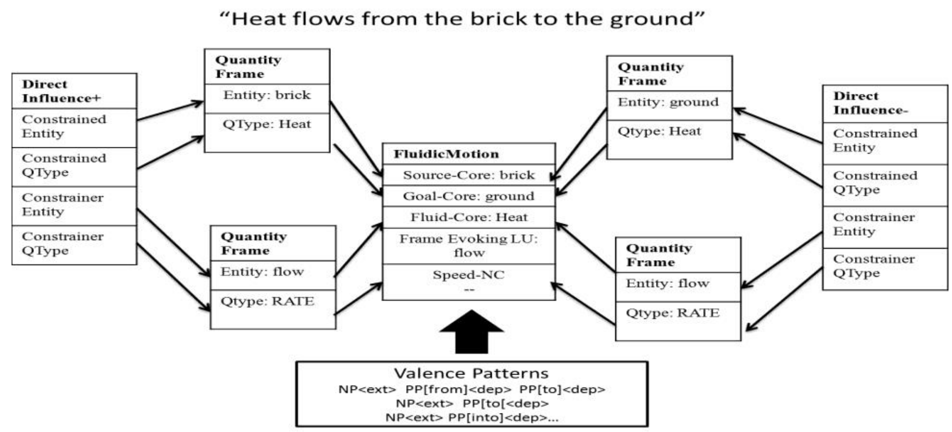

Figure 1: QP frames for a sentence

(Fillmore, Wooters, & Baker, 2001). Fillmore *et* al’s (2001) FrameNet is a frame semantics for English. FrameNet frames are evoked by a frame-bearing lexical unit. The dependent structures in the sentence form arguments to that frame’s frame elements (FEs). For example, the `Motion` frame includes frame elements for the `Source`, `Goal`, and `Theme`. It is instantiated in a specific construction (a pairing of syntactic form and semantic meaning) by a frame evoking lexical unit (LU) such as the word *fly* in “The bird flew to Florida”. Here, the noun phrase (NP) subject fills the role of `Theme` and the prepositional phrase (PP), *‘to Florida’*, fills the role of `Goal`. The specific grammatical instantiation of these roles is called a *valence pattern* for that lexical unit. Thus a simple transitive would have a valence pattern of an NP subject and an NP object.

Two core inter-frame relationships in FrameNet are inheritance and sub-frame. A frame that inherits from a parent must have a corresponding frame element for each element of the parent and can introduce others. Inheritance can be partial or multiple (Baker, Fillmore & Cronin, 2003). Sub-frames allow FrameNet to represent complex scenarios as a sequence of smaller parts related through precedence.

QP Language

Kuehne (2004) recast QP theory in a frame-semantic representation to better handle compositionality in language. Quantity frames fill the argument slots of influence frames. The influences participate in quantity transfer descriptions and process frames which express the results and activation conditions of the process. This representation differs from FrameNet in that sub-frames in FrameNet are related through their role relations and frame-frame relations, instead of filling a frame element of the superordinate frame. Our representation benefits from compactness and allows us to define entire frames as FE arguments, though it is straightforward to transform them into FN conventions (e.g. Figure 1).

Kuehne (2004) identified several syntactic patterns that instantiated QP frames, and used them to automatically extract QP frames from text, using Davidsonian lexical representations from Cyc KB[1] contents. This approach, while successful, was limited by the syntactic patterns in the controlled grammar and did not integrate into the language system more broadly. McFate *et al* (2014) extended this approach and introduced narrative functions to guide disambiguation (Tomai, 2009). This system’s coverage was limited by the coverage of Cyc’s semantic templates, and it also became evident that a finer-grained set of distinctions would be useful. Integrating QP frames with FrameNet helps solve both problems by providing valence patterns by frame type. It benefits frame semantics by providing rich representations for mental models.

Unifying FrameNet with QP Frames

Continuous Processes

Continuous processes are process verbs and nominalized verbs in English. Since direct influences are only allowed within processes, we start with them. The direct influence (DI) frame has four required elements: `constrained`, `constrainer`, `entity`, and `sign`. The constrained and constrainer are both quantity frames. The `entity` is the process, and the `sign` indicates the direction of contribution for the rate.

FrameNet has many frames that instantiate continuous processes and thus DIs. A straightforward mapping for these frames aligns QP elements to potential FrameNet elements. Figure 1, above, is an example of `Fluidic_motion`. It’s difficult to evaluate precisely how many frames in FrameNet instantiate a DI, but we can get a

[1] http://www.cyc.com/platform/researchcyc

rough estimate by examining the FrameNet frame lattice which broadly categorizes sets of interrelated frames. An upper bound involves all members of the `Event` category which has 300 frames with 695 total members (counting inheritors). However, the `Event` category includes non-durative events such as `name_conferral`. A better starting place would be several of FN's smaller groupings. The vast majority of direct influence relevant frames are captured in the Motion_Scenario cluster (45 members) and Change_Of_Phase_Scenario (3 members). A few outliers exist in other clusters such as Catching_Fire (4 members) and Transfer_Scenario (27 members). Furthermore, many of these frames reuse the same frame elements which greatly simplifies mapping additional frames.

Next we provide several process frame examples for two high level frames, Motion and State Change/Conversion. For each process type we match Frame Elements (FEs) to QP Elements and analyze several unique instantiations for that process class. They are exemplary of the productivity of mapping these two resources.

Motion

Motion in FrameNet is captured with the `Motion` frame, which has several inheritors. Kuehne (2004) focused specifically on processes that align with `Fluidic_Motion` which is where we start. Table 1 summarizes relevant patterns for `Fluidic_Motion` and its alignment to DI Frames. The QP elements are `constrainedEntity` and `constrainedQuantityType`. These map to FrameNet elements `Goal/Source` and `Fluid`.

The left-hand column in the following tables illustrates FrameNet valence patterns that instantiate the QP elements. The patterns are presented with FrameNet grammatical functions. The most common are `ext`, `dep`, and `obj` which indicate that the phrase is an external argument (subject), verb dependent, or object of the sentence.

As an example, the top left cell in Table 1 says that a positive direct influence can appear as a transitive-pp sentence with *to*, *into*, and *in*. The `constrainedQuantityType` is indicated by the NP subject, and the `constrainedEntity` is instantiated by the prepositional phrase (PP) dependent. These valence patterns are found by mapping the QP representation to the FrameNet frame elements (FEs) and then finding the patterns that instantiate the target FEs in individual frames.

Table 1

Valence Pattern		Example	FrameNet Frames
DI+ Transitive-PP		Water **flows** to the basin.	Fluidic Motion: *Core:* Goal, Source, Path, Fluid, Area *Non-core:* Speed
Cnd-Qtype	Cnd-Ent		
NP<ext>	PP[to]<dep>		
	PP[into]<dep>		
	PP[in]<dep>		
DI- : Transitive-PP		Water **flows** from the basin	
Cnd-QType	Cnd-Ent		
NP<ext>	PP[from]<dep>		
	PP[out]<dep>		

Positive direct influences correspond to frame instantiations that include the `Goal` while negative direct influences correspond to the `Source` FE. The constrained quantity corresponds to the `Theme` in the `Motion` frame and its corresponding element in the inheritors (ie. `Fluid`).

FrameNet further separates the `Motion` and `Cause_motion` frame, a distinction which holds for the inheritors as well and requires an `Agent` or `Cause FE`. `Cause_motion` verbs can appear in additional valence patterns, as shown in Table 2.

Table 2

Valence Pattern		Example	FrameNet Frames
DI +/- Transitive		New York pumps water from the subway.	Cause_fluidic_motion: *Core:* Goal, Source, Path, Fluid, Area, Agent/Cause
Cnd-QType	Cnd-Ent		
NP<obj>	PP[from]<dep>		
	PP[to]<dep>		
	…		
DI+: Passive		Water is spilled from the bucket.	
Cnd-QType	Cnd-Ent		
NP<obj>	PP[from]<dep>		

While the verb examples above rely on dependent prepositional phrases to indicate the sign of the DI, some verbs encapsulate the sign as well, which allows a unique FE assignment with the simple transitive (Table 3).

Table 3

Valence Pattern		Example	FrameNet Frames
DI- : SimTrans		The bucket **leaks** water.	**Fluidic_motion**: *Core:* Goal, Source, Path, Fluid, Area
Cnd-QType	Cnd-Ent		
NP<obj>	NP<ext>		
DI- : SimTrans		The sun **emits** heat.	**Emitting**: *Core:* Emission, source_Emitter
Cnd-QType	Cnd-Ent		
NP<obj>	NP<ext>		

State Changes & Conversion

State change is a limited class and includes processes such as boiling. A state change process can be represented as a pair of direct influences representing the increase in substance at one phase and decrease in substance at another (e.g. steam and water)

This is an important representational decision because it separates state-changes from their preconditions. This holds up in the valence patterns of state-change verbs which differ from motion (Table 4). FrameNet captures change with the `Change_of_phase` frame and its causer. The core element, `Undergoer`, is the changing entity. This differs from prior analyses of flow and motion, because they are referring to an event without referring to the details of what happened during it (e.g. "The water boils away"). Note that the quantity type of the direct influence frequently relies on the semantics of the verb (e.g. the resultative in the top two cells).

Table 4

Valence Pattern		Example	FrameNet Frames
DI+ : Resultative-State		The water **froze** to ice.	Change_of_phase: *Core*: Undergoer *Non-core*: Result, Speed, Place Cause_change_of_phase: Core: Agent, Cause, Undergoer
Cnd-QType	Cnd-Ent		
PP[to]<ext>	NP<ext>		
DI- : Resultative-State			
Cnd-QType	Cnd-Ent		
NP<ext>	NP<ext>		
DI- : Intransitive		The water **froze**.	
Cnd-QType	Cnd-Ent		
DEN	NP<ext>		

This is clearest in the intransitive where the knowledge that *ice* is the resulting state is indicated only by the verb, freeze. The conversion interpretation is supported by multiple PP attachments as in "The water froze from liquid to ice".

Thus, in mapping to state changes we extend the `Change_of_phase` frame to include required initial and final states. This puts our interpretation closer to the `Undergo_Change` and `Cause_Change` frames. These include verbs such as *change*, *convert*, and *turn* which have core `Final` and `Initial_category` elements.

As with motion, an `Agent` frame-element (causative) results in unique transitive and by-PP constructions. One difference between conversion and state-change is that conversion agents tend to also be the entity that possesses the quantity while in a state-change it is usually an outside performer. This isn't always the case and can be over-ruled by explicit possession, but it is an interesting side effect of conversion having less regular participants.

Constrainers

So far we have ignored rates. Kuehne (2004) found that English frequently left rates implicit despite being a crucial part of direct influences. Non-causative frames above contain the non-core `Speed` frame element which includes rate constructions (e.g. at a rate of X per Y)

For process frames lacking the `Speed` element, a QP analysis necessitates including a rate FE. Since the rate is a necessary constraining quantity in a direct influence, we assume it is implicitly evoked by the frame evoking lexical unit when not otherwise stated.

Indirect Influences

The indirect influence frame, also called a qualitative proportionality (Qprop), has a `consequent` frame element, `antecedent` frame element, and `sign`. Qprop patterns occur in the `Contingency` and `Objective_Influence` frames which include words such as *influence* and *depend***.** Both FN frames have a frame element corresponding to a causal and dependent entity which corresponds to the antecedent and consequent of the influence. These lexical units can be modified with an adverb such as *negatively* to reverse directionality.

Table 5

Qprop+ : Transitive-PP		The rate **depends** on the surface area.	**Contingency** *Core:* Determinant, Outcome
Conseq-qtype	Antecedent-qtype		
NP<ext>	PP[on]<dep>		
	PP[upon]<dep>		
Qprop+ : SimTrans		The amount of water **influences** the water level.	**Objective_Influence** *Core*: Dependent_entity Influencing_entity or situation
Conseq-qtype	Antecedent-qtype		
NP<ext>	NP<obj>		
Qprop+ : Passive		The water level is **influenced** by the amount of water.	
Conseq-qtype	Antecedent-qtype		
NP<obj>	PP[by]<dep>		

A similar mapping holds to the `Actor` and `Affected` of the `Causation` frame. For example, when indirect influences hold between rates and quantities governed by another process we see underspecified statements such as:

"Deforestation causes less carbon to be taken out of the atmosphere."

This sentence indicates a constraint on carbon absorption tied to the quantity governed by deforestation. Similar patterns tie the rates of two processes together:

"Heating water causes it to boil."

Disambiguating whether the rate or governed quantity is the antecedent frequently requires domain knowledge. Additionally, several phrasal constructions can indicate covariance and thus evoke an indirect influence (Table 6).

Table 6

1. Comparative Correlative	The higher the water level in the bucket, the greater the water pressure in the bucket.
2. Correlative Conjunction	Both temperature and pressure increase.
3. As X, Y	As the temperature in the boiler increases, the pressure in the boiler increases.
4. Changes with Y	The pressure of the boiler changes with the temperature.

Construction 1 is the comparative correlative (Culicover & Jackendoff, 1999). Here the first and second phrases map to the antecedent and consequent roles of the indirect influence. The sign is given by the directionality of the adjectives. A similar mapping exists for some correlative conjunction constructions. For an indirect influence, the conjunct phrases must both involve change verbs (e.g. increase/decrease). Kuehne (2004) noted that the construction in 3 was one of the most common.

Quantity Frames

In QP frames, influences operate over lexical units or phrases that evoke quantity frames. They have the frame elements `quantityType, entity, quantityValue,` quantityUnit, and signOfDerivative. The role quantityType relates the frame to the Cyc collection denoting the quantity of the sentence. An example would be *Heat* or *Pressure*. The entity is the object that the quantity attribute pertains to. Value and unit are optional and relate the quantity to numerical data. signOfDerivative is an optional relation that indicates direction of change if any.

In many cases, identifying a quantity is very direct. Quantity evoking lexical units include "*heat*", "*pressure*", and "*volume*". These units fit most cleanly into the Measurable_attribute frame, though it has thus far only been applied to gradable adjectives (e.g. The hot brick). Furthermore, in FrameNet the frame explicitly evokes the notion of deviation from a norm.

While adjectives modify an entity directly, quantity frames often rely on possession to indicate the entity. These instantiations map to the Possession frame and include possessive verbs (e.g. *have*) and genitive constructions, though only a subset of Possession verbs is suitable. QP possessives seem restricted based on the nature of the thing possessed. A counterexample is "The brick owns mass."

Containment also links quantities to entities. One can describe the "energy contained in the boiler" or separate out quantities with "the air and water pressure in the container". We represent this using the ContainedStuffFn function which defines a compound entity of the sub-part contained at a specific phase. These constructions fit into the Containers and Containing frame.

The QType can be compositional with the unit. This occurs in measurement phrase constructions (Table 7).

Table 7

QValue	Qunit	Qtype	5 **liters** of water	**Measures** *Core*: Count, Entity, Unit
Num.Quant	DEN	PP[of]<dep>		

QValue	Qunit	Qtype	Entity	The wall is 6 feet **tall**	**Dimension** *Core*: Dimension, Measurement, Object
Num.Quant	N<dep>	DEN	NP<ext>		
QValue	Qunit	Qtype	Entity	The 6 foot **tall** wall	
Num.Quant	N<dep>	DEN	NP<obj>		

Qunit	Qtype	**Amount** of water	**Quantity** *Core*: Entity, Quantity
DEN	PP[of]<dep>		

The first valence pattern consists of a measurement expression modifying an "of" PP. The next two feature constructions that take a measurement expression and adjective phrase and return either a noun-modifier or a predicate (Fillmore, Lee-Goldman, & Rhodes, 2012). Additionally, *'amount'* constructions can be used to explicitly define a substance as a quantity.

Finally, QP direct influences are constrained by rates which are represented by a quantity frame where the entity is the process itself. In FrameNet, rates are usually an optional role in the event, evoked through modifiers such as *'quickly'*. In our mapping, rate is a required role, instantiating a quantity frame even if not explicitly mentioned (see the rate frame in Figure 1). FN supports null-instantiated element, but in our approach, evoking a direct influence results in the verb also evoking a rate quantity frame. This is not a mechanism in FN, though a simple extension would be to add rate quantity frames to FN and have them be evoked by any process verb. McFate *et al*, (2014) simply create rate frames as a side-effect of evoking a direct influence.

Rates can also be explicitly referenced using nominalized verbs (e.g. 'the rate of flow'). These phrases are often anaphoric and are supported by FrameNet with the Rate_quantification frame. Finally, Fillmore *et al* (2012) identify a rate phrase construction which consists of a numerator, the definite NP, and a denominator, the indefinite np (e.g. "twice a day" "miles per hour"). As FN is a lexical resource, it does not have frames for these constructions. Representing these and a few others (like those in Table 6) could be resolved by adding construction frames to FrameNet as in Fillmore *et al*, (2012).

Ordinals

Inequalities frequently drive processes. An ordinal frame has two frame elements, QF1 and QF2 which take quantity frames. There is a relation FE that defines the direction of the ordinal (>, <, =, negligible) (Dauge, 1993). In FrameNet style, these could be expanded to Entity1, Entity2, sign, and Qtype.

The most direct way these appear in language is through gradable adjectives in a comparative construction (e.g. 'cooler than' or 'more cool than'). In FrameNet, this fits most neatly into the Evaluative_comparison frame, though it does not include comparative adjectives. Individual Measurable_attribute adjectives such as *big* and *small* do have valence patterns where a than-PP (e.g. bigger than X) is mapped to the Degree frame element, though this same element applies to modifiers such as *very*. There is no comparative construction based on *more* or *less*. This could be addressed with a construction frame (e.g. less X than or X-er than Y).

Ordinal frames are also evoked by non-comparative Measurable_attribute expressions such as *hot* or *cold* since they explicitly evoke deviation from a norm. Through inheritance, FrameNet associates many of these adjectives with frames of their QuantityType such as Temperature and Size. What it doesn't provide is the relative positions of the adjectives on a scale: that *hot* is greater in temperature than *cold*. Thus while the appearance of two measurable attributes in the same paragraph can evoke an ordinal, we rely on assertions from ResearchCyc to know the sign of the ordinal relationship.

There are also several ways to indicate a difference without specifying directionality. These are often referential and fall under the `Similarity` frame (Table 8). A difference can be introduced and referred to with an NP as in the first valence pattern below, or with similarity verbs (Intransitive/ Transitive-PP).

Table 8

Ent1	Ent2	Qtype	The temperature **difference** between the objects	**Similarity** *Core*: Differentiating_fact, Dimension, Entities, Ent1, Ent
PP[between] <dep>	PP[between] <dep>	N<dep>		
PP[between] <dep>	PP[between] <dep>	PP[in] <dep>		
Intransitive			The temperatures of A and B **differ**.	
Ent1	Ent2	Qtype		
PP[of]<dep>	PP[of]<dep>	NP<ext>		
Transitive-PP			The temperatures **differ** between the bricks.	
Ent1	Ent2	Qtype		
PP[across] <dep>	PP[across] <dep>	NP<ext>		
PP[between] <dep>	PP[between] <dep>			
PP[from] <dep>	PP[from] <dep>			

One interesting feature of similarity verbs is that the entity can be left out. Usually this occurs either in a generic statement or when the quantity is referential. In these cases, the `entity` or `entities` frame element maps to the Qtype of the two participating quantity frames.

One significant difference between our representation and FrameNet's is that FrameNet provides the `Entities` frame element which groups multiple individuals to fill one role. We separate them since they indicate different quantity frames.

Limit Points, Transitions, and Constraints

Limit points are quantities that define a value where a model fragment changes status (e.g. *boiling point*). Limit points are vital in understanding when and for what reasons a state change occurs, but prior work has not mapped them to linguistic frames. We have found that several FrameNet frames can instantiate a limit point. Frequently they occur as a compound nouns consisting of a constrained process and a barrier. They can also include numerical values, and can participate in possession and containment. Limit points are also evoked with `Extreme_value` adjectives such as maximum or minimum. Verbs that signify arrival at a point as in the `Arriving` frame can also evoke them (e.g. "The oven reached 400 degrees.").

Many limit points are left implicit or referred to only as a deviation from the norm as in: "The water gets cold which causes condensation" The fact that these are limit points can be made explicit with a modifier such as *enough* (`Sufficiency`)

FrameNet's `Process` frame has sub-frames indicating different states. Starting conditions are captured with patterns from the `Process_start` frame which includes verbs such as 'begin'. Similar frames exist for stopping, continuing, pausing, and resuming. When used in conditionals, lexical units evoking sub-frames and limit points define the constraints of a model fragment (Table 9).

Table 9

Condition	Process	Once the submarine **reaches** crush depth, compression **begins**.	**Process_Start** *Core* Unxp: Event *Non-Core*: Time **Arriving** *Core:* Goal, Theme
PP[after]<dep>	NP<ext>		
PP[when]<dep>			
PP[if]<dep>			
PP[once]<dep>			
Condition	Process	After **reaching** 2,070 degrees, the steel **begins** melting.	
PP[after]<dep>	Vping<dep> VP[to]<dep>		
PP[when]<dep>			
PP[if]<dep>			
PP[once]<dep>			

These patterns involve multiple frames. The conditional PP (e.g. when, after…) indicates a condition which is provided by the limit-point evoking verb (e.g. reach). The process, as an NP or VP, then becomes the argument of a process state change verb such as *begin*.

Many frames in FrameNet can indicate their own activation conditions. Consider the `Cause_Motion` valence patterns in Table 2. The required `Agent` can be replaced with a non-animate `Cause` and viewed as preconditions:

"A temperature difference drives heat to the brick."

These constructions can also be used to elaborate on previous instantiations of frames that they are causative of.

"A temperature difference drives heat flow."

Explicit `Causation` verbs and modifiers can also indicate a process constraint (e.g. 'because the temperatures differ…').

Finally, quantities within a process can be constrained at certain values using correspondence statements such as:

"The force of the spring is zero when the block is at zero."

This temporal correspondence is captured by the frame `Temporal_collocation` which relates a trajectory and landmark. This is a vast frame and includes indexical terms such as "today". We constrain the QP mapping to valence patterns that relate two events as with the adverb '*when*'.

Corpus Analysis

We have demonstrated that the core elements of QP theory all have corresponding FrameNet frames. This unified semantics both sheds light on the semantics of continuous processes, and provides a basis for future work in computational models of semantics. Next we ask, how frequently do these patterns occur in natural language? We answer this question with a corpus analysis.

Our corpus consisted of grade school science topics from the Simple English Wikipedia: full articles on the water cycle and Bernoulli's principle as well as the first 6 sections about the sun and the introduction of the global warming article.

There were a total of 77 sentences. Each sentence was annotated for QP frames. For each process evoking LU (e.g. *flow*) we evaluated its FrameNet correspondence. We counted FN as having the valence pattern if the specific LU in the correct frame had the complete pattern annotated either alone or as a part of a larger pattern with the same core elements. Not having a specific valence pattern does not preclude it from occurring in other frame evoking units, and future work could rely on similar lexical units as a scaffold for missing annotations. The results are summarized by article in Table 10.

Table 10

Corpus Analysis Results				
	QP Evoking	Process Evoking	Valence in FN	Motion or Conversion
water	9 (.64)	8	5	4
sun	15 (.56)	20	7	13
Bernoulli	8 (.62)	4	2	2
Global-warming	12 (.60)	7	3	1

Out of all 77 sentences, 44 (57.1%) had QP material. We identified 39 process evoking lexical units (*e.g. flow*) across 28 of those 44 sentences. Other QP sentences described entity attributes (e.g. 'the sun is big') or introduced indirect influences between other quantities.

These results suggest a substantial number of sentences in science texts convey QP information. Furthermore, in our verb analysis, we found that 43.5% of process evoking units already had their specific valence pattern annotated in FrameNet, and only one of the units lacked a FrameNet entry ('build up'). Thus, mapping QP theory, and possibly other non-linguistic representations, to FrameNet is a good method for illuminating how these models are expressed linguistically.

20 of the 39 process evoking units either directly evoked or evoked inheritors of the set of frames analyzed above. The remaining verbs evoked `Giving`, `Receiving` `Creating`, `Gathering_up`, `Arriving`, `Departing`, `Removing`, `Soaking_up`, `Using_resource`, and `Fire_burning`. Of these, all but the last three are members of the Event category. Only `Soaking_up` and `Using_resource` are uncategorized in the frame lattice.

Like in Kuehne's (2004) analysis, we found that reference to an explicit rate was rare, only occurring in four sentences of the Bernoulli article. Furthermore, we found no compound-noun limit points (e.g. boiling point) but did find constraints based on deviation from the norm (e.g. when it gets cold…). In part this was due to choice of articles, e.g. articles on boiling or phase changes per se do mention them.

Related Work

While not connected to a specific cognitive theory, Ovchinnikova *et al* (2010) used a data-driven analysis to cluster and enrich FN frames about medical treatment. A similar approach could be used to generalize QP evoking frames and identify new relationships between them. Furthermore, our analysis provides a mapping to a non-linguistic ontological mental model that could provide an interface between linguistic and non-linguistic representations in cognitive architectures.

Recently, other researchers have developed general frame semantic parsing systems such as Ovchinnikova's (2012) statistical abduction system and Das *et al*'s (2014) SEMAFOR. Both systems could use a QP mapping to turn parsed frames into QP representations. Furthermore, knowledge of the constraints on qualitative models could help resolve ambiguities in language processing. McFate *et al* (2014) has suggested that QP specific narrative functions could provide such a mechanism.

Discussion and Future Work

Qualitative Process Theory provides a powerful formalism for representing mental models about continuous processes. Linking the QP formalism to frame semantic resources both enriches the linguistic representations with a higher-order model and provides concrete linguistic instantiations of QP concepts.

In this paper, we have demonstrated how each of the core elements of QP theory appears in select FrameNet frames. In illustrating the productivity of this mapping, we have focused on two broad types of processes, *flow* and *state change*. Just these two processes were able to account for 20 out of 39 direct influences in our corpus analysis, and all but three of the missing process types would be covered by completing the mapping for the rest of FrameNet's event lattice. Furthermore, the FrameNet lattice has clustered events into several smaller sub-collections which provides leverage for selectively expanding the mapping in future work.

While FrameNet frames and valence patterns provide some method for expressing each QP element, FrameNet does not support all methods of expressing the information. This is most clear with phrasal constructions such as the comparative correlative or if-then constructions. These multi-word and phrasal constructions are beyond the initial goal of FrameNet, but could be addressed with construction frames such as those suggested by Fillmore *et al* (2012).

These higher-order constructions are especially important for ordinal frames, since a very common method of expressing ordinal constraints involve comparative constructions. Furthermore, while FrameNet does provide the knowledge that terms like *big* and *small* are measurements along the same scale, the ordinal relationship between the two terms requires outside knowledge. Evoking an ordinal frame in this way is pragmatic rather than syntactic. Future work would involve placing the FrameNet lexical-units for each measurable attribute onto a scale.

One candidate representation for expanding this work with constructions is Sign Based Construction Grammar (Boas & Saag, 2012) which combines functional and construction grammar approaches. SBCG is amenable to frame semantics and could provide hierarchies for further classifying these phenomena. One suggestion for learning linguistic generalizations and has been through a process of structural alignment (Taylor *et al*, 2011). Given the structure of frame semantics, this seems a likely candidate for future research.

Finally, future work will involve reasoning with QP models extracted from text. This presents several challenges. Perhaps the greatest is that models in text are often only partially complete. Frame semantic representations are very well suited for accumulating fragmentary knowledge. Future work will focus on treating frames extracted from text as process examples for analogical generalization. Over the course of several examples, these generalizations could be used to specify a complete process model.

Acknowledgments

This research was supported by the Intelligent and Autonomous Systems Program of the Office of Naval Research.

References

Baker, C. F., Fillmore, C. J., & Cronin, B. 2003. The structure of the FrameNet database. *International Journal of Lexicography*, 16(3), 281-296.

Boas, H. & Saag, I, (Eds.) 2012. Sign Based Construction Grammar. Stanford, CA: CSLI Publications.

Das, D., Chen, D., Martins, A. F. T., Schneider, N., & Smith, N.A. 2014. Frame-Semantic Parsing. Computational Linguistics, 40(1), 9-56.

Dauge, P. 1993. Symbolic Reasoning with Relative Orders of Magnitude. In Proceedings of the 13th IJCAI.

Fillmore, C. J., Wooters, C., & Baker, C. F. 2001. Building a Large Lexical Databank Which Provides Deep Semantics. *Proceedings of PACLIC-2001*, Hong Kong, China.

Fillmore, C. J., Lee-Goldman, R., & Rhodes, R. 2012. The FrameNet Constructicon. *Sign-based Construction Grammar. CSLI, Stanford, CA*.

Forbus, K. D. 1984. Qualitative process theory. *Artificial Intelligence*, *24*, 85–168.

Hinrichs, T. and Forbus, K. 2012. Learning Qualitative Models by Demonstration. *Proceedings of QR2012*.

Kuehne, S. E. 2004. Understanding natural language descriptions of physical phenomena. Doctoral Dissertation, Northwestern University, Evanston, Illinois.

McFate, C.J., Forbus, K. and Hinrichs, T. 2014. Using Narrative Function to Extract Qualitative Information from Natural Language Texts. *Proceedings of AAAI-2014* Québec City, Canada.

Ovchinnikova, E., Vieu, L., Oltramari, A., Borgo, S., & Alexandrov, T. 2010. Data-Driven and Ontological Analysis of FrameNet for Natural Language Reasoning. In LREC-2010.

Ovchinnikova, E. 2012. Integration of World Knowledge for Natural Language Understanding. Atlantis Press, Springer.

Taylor, J. L. M., Friedman, S. E., Forbus, K. D., Goldwater, M. and Gentner, D. 2011. Modeling structural priming in sentence production via analogical processes. *Proceedings of CogSci-2011*. Boston, MA.

Tomai, E. 2009. *A Pragmatic Approach to Computational Narrative Understanding.* Doctoral Dissertation, Northwestern University, Evanston, Illinois.