Extracting Normative Relationships from Business Contracts

Xibin Gao
Microsoft Corporation
One Microsoft Way
Redmond, WA 98052-7329, USA
xibin.gao@gmail.com

Munindar P. Singh
Department of Computer Science
North Carolina State University
Raleigh, NC 27695-8206, USA
singh@ncsu.edu

ABSTRACT

The normative concepts offer a principled basis for engineering flexible multiagent systems for business and other cross-organizational settings. However, producing suitable specifications is nontrivial: the difficulty is an obstacle to the adoption of multiagent systems in industry. This paper considers *normative* relationships of six main types, namely, commitments (both practical and dialectical), authorizations, powers, prohibitions, and sanctions. It applies natural language processing and machine learning to extract these relationships from business contracts, establishing that they are realistic and their encoding can assist modelers, thereby lowering a barrier to adoption. A ten-fold cross-validation over more than 800 sentences randomly drawn from a corpus of real-life contracts (and manually labeled) yields promising results for the viability of this approach.

Categories and Subject Descriptors

I.2.11 [**Artificial Intelligence**]: Distributed Artificial Intelligence—*multiagent systems*; D.2.1 [**Software Engineering**]: Requirements/Specifications—*Tools*

General Terms

Algorithm, Experimentation

Keywords

Norms; Agent-oriented software requirements

1. INTRODUCTION

Normative concepts provide a natural way to engineer multiagent systems for settings involving two or more autonomous parties, such as cross-organizational business service engagements. However, coming up with appropriate normative specifications for such systems is nontrivial. This paper is premised on the idea that some of the essential normative specifications might be encoded in human-produced artifacts, specifically, business contracts. Although it may not be viable to automatically execute them (because they are not specified in formal notation), contracts can provide a clear account of the norms in play, which a multiagent systems engineer would be able to suitably formalize.

For example, a contract[1] may state "Generico will issue an invoice to New Alpha for each Purchase Order, based upon and reflecting the applicable price" to specify that Generico will issue invoices in a certain manner. The above specification regulates Generico and New Alpha's interactions: it yields a requirement for Generico's agent expressed as a standard for correctness. The agent would map the above requirement to a functional requirement on its IT systems to make sure that the invoices are properly generated. Or, the agent may exercise its autonomy and violate this part of the contract, thereby risking sanctions, which the contract itself may specify or the surrounding society may impose.

A typical contract could contain hundreds of normative statements describing each party's expectations of the others. This complexity hints at the potential benefits and challenges of dealing with contracts. Contracts embody significant domain knowledge and past experience with various business situations: attempting to achieve such detail from scratch would be a daunting task, but though contracts carry such knowledge, extracting it can be nontrivial as well.

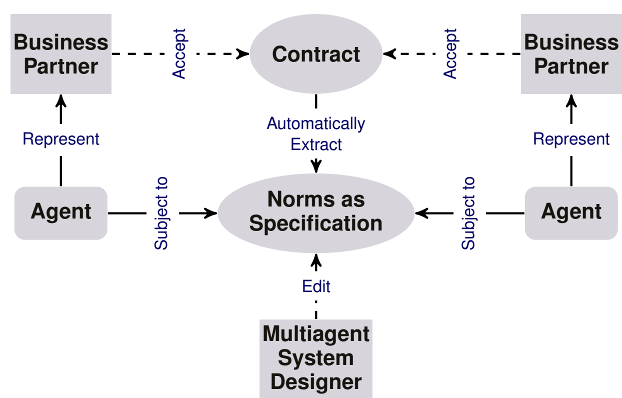

Figure 1: Envisioned use of contracts in design.

Figure 1 illustrates how our envisioned approach can apply in practice. A contract would exist between principals (business partners). Our tool would extract norms from the contract. A designer would use the extracted norms and produce a specification for a multiagent system. Each principal

[1]http://contracts.onecle.com/app-pharmaceuticals/abraxis-manufacturing-2007-11-13.shtml

Appears in: *Proceedings of the 13th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2014), Lomuscio, Scerri, Bazzan, Huhns (eds.), May, 5–9, 2014, Paris, France.*

would implement (or obtain) an agent who would represent that principal and engage in the multiagent system to realize the contract. The main intended benefit of our approach is to reduce the manual effort required of the multiagent systems designer.

Both in engineering multiagent systems and in executing contracts, it is important to know what each party expects from the others and a way to judge their potential compliance. For this reason, we adopt Singh's [26] formulation of directed *normative relationships* between the concerned parties. (Below, *norm* is a shorthand for normative relationship, and deviates from the vernacular use of "norm" as a generic, undirected ideal.) In this formulation, a norm has a fixed structure: a type, a subject (on whom the norm applies), an object (with respect to whom the norm applies), an antecedent, and a consequent. The directionality is crucial: it clarifies who is accountable to whom.

The first contribution of this paper is in developing an approach for extracting norms of the above structure from naturally occurring contract text. Its results are promising and though they are not perfect, which would be difficult when dealing with human language, they show that assisting a designer in formalizing a multiagent system based on a contract is a viable endeavor. This work is realized in a prototype tool suite for extracting norms and allied concepts from contracts. Although we focus on norms in this paper, its contribution can be readily combined with emerging approaches for extracting exceptions, business events, and temporal constraints from contracts, e.g., [8, 9].

The second contribution of the work is evaluating the realism of normative models in multiagent systems (specifically, Singh's [26] formulation) by determining how well those concepts can be automatically identified in contracts. Doing so may lead to deeper understanding of the theory of norms, reflecting how they are employed in practice.

Organization. Section 2 motivates and defines the normative relationship extraction task. Section 3 describes our approach. Section 4 presents the results of our evaluation. Section 5 surveys the relevant literature. Section 6 discusses some challenges and directions for future research.

2. NORMS IN CONTRACTS

Autonomous parties enter into business contracts that express in high-level terms their mutual expectations regarding their interactions. A party's compliance with a contract presumes its being able to model the norms as described in a contract and to treat those norms as requirements for its IT artifacts that would enable its operations to be compliant with the contract. For example, a vendor may enter into a contract to deliver products correctly configured according to customer specifications. In that case, it may impose a requirement on its IT system to read in customer orders and convey the specifications to its manufacturing subcontractor. Capturing the norms explicitly can thus greatly facilitate a business in complying with its contracts and monitoring its counterparties for their compliance.

2.1 Background: Norm Types

We adopt Singh's [25, 26] model of norms with one extension [21]. A norm is directed from a subject to an object and is constructed as a conditional relationship involving an antecedent (which brings the norm in force) and a consequent (which brings the norm to satisfaction). The advantage of this representation is that it yields clarity on who is accountable to whom. Specifically, a norm has four core elements—*subject*, *object*, *antecedent*, and *consequent*. Norms in our approach are of the following main types.

A commitment means that its subject commits to its object to ensure the consequent if the antecedent holds. (Singh [25] refers to the subject and object of a commitment as debtor and creditor, respectively. We adopt the more neutral terminology to support a unified model of all six norm types [26].) In a purchasing contract, commitments are generally prominent in scenarios such as product delivery and payment. Commitments are of two subtypes [21].

- **A dialectical** commitment is a claim staked by its subject, i.e., that the consequent is true if the antecedent is. A party's representations and warranties (e.g., the seller owns what she is selling) are its dialectical commitments. Likewise, an agreement as to the facts is a dialectical commitment by each of the agreeing parties.
- **A practical** commitment is a promise to ensure that the consequent will be brought about if the antecedent becomes true. For example, a seller's offer to a prospective buyer to provide specified goods for a specified payment is a practical commitment.

An authorization means that its subject is authorized by its object for bringing about the consequent if the antecedent holds. For example, in a manufacturing contract, the manufacturing facility owner may authorize a client to visit a facility with restricted access.

A power means that its subject is empowered by its object to bring about the consequent if the antecedent holds. Following Jones and Sergot [14], a power refers to institutional actions involving change of normative relationships. For example, in a manufacturing contract, the purchaser may cancel an order with prior notice, that is, it can terminate a commitment at will, thereby changing the normative relationship between itself and the manufacturer.

A prohibition means that its subject is forbidden by its object from bringing about the consequent if the antecedent holds. For example, in an employment contract, the employee may be forbidden from revealing the employer's confidential information to outsiders.

A sanction specifies the consequences its subject faces from its object for violating another norm, such as a commitment, an authorization, or a prohibition. For example, in a software licensing contract, if one party reverse engineers a software module and infringes a patent or other intellectual property, it may be sanctioned by having to pay damages.

2.2 Norms as Evident in Contract Text

The text excerpts below show examples of norm types from a manufacturing and distribution agreement between Sharp Corporation and Danger Inc.[2]

Practical commitment: Sharp will be responsible for ensuring an adequate supply of spare parts to distributors or repair facilities so warranty returns can be repaired and sent back to the field without undue delay.

[2]http://contracts.onecle.com/danger/sharp-mfg-2004-04-28.shtml

Dialectical commitment: Each Party represents and warrants to the other Party that (a) as of the Effective Date it has the full right and authority to enter into this Agreement and grant the rights and licenses granted herein.

Authorization: Danger will have the right to use the test results internally for product management and planning purposes.

Power: In the event this Agreement terminates for a reason other than for Danger's material breach, Danger shall have the right to purchase any Sharp owned tooling and test equipment for a Product at a reasonable price.

Prohibition: Danger shall not issue a press release mentioning Sharp without Sharp's prior consent, which shall not be unreasonably withheld or delayed.

Sanction: This Agreement may be terminated in its entirety by either Party immediately upon the occurrence of any the following events: (a) if the other ceases to do business, or otherwise terminates its business operations; (b) if the other materially breaches any material provision of this Agreement and fails to cure such breach within forty-five (45) days after receiving written notice from the non-breaching Party describing such breach; [...]

2.3 Technical Problem

In simple terms, our technical problem is, given a contract expressed in natural language, extract each norm and assign a type to it. Norms are expressed in sentence-level clauses in contracts. A sentence is either not a norm or one of the six mutually exclusive norm types. (When a sentence may express more than one norm—in which case, for simplicity, we consider only the norm expressed by its main verb phrase: we return to this point in Section 6.1.) Hence our problem boils down to the following two tasks: (1) Given a sentence from a contract, assign one of the seven labels—{not a norm, dialectical commitment, practical commitment, authorization, power, prohibition, sanction} to it. The last six are norm types. (2) Given a sentence from a contract that expresses a norm, extract the elements (subject, object, antecedent, consequent) of the norm.

3. APPROACH

Approaches for information extraction from unstructured text fall into two broad families. Pattern-based methods are simple and effective whereas machine learning-based methods can yield greater flexibility and robustness.

Figure 2 summarizes our hybrid approac hh. First, we preprocess a contract by stripping off any HTML tags, headlines, signoffs, and addresses. Second, we segment a preprocessed contract into a collection of sentences and filter the sentences based on signal words to obtain the norm candidates. Third, we produce a grammar tree (parse tree with annotations) from each candidate sentence. Fourth, with automatically extracted features harvested from the grammar tree, we apply a classification method to decide if a candidate expresses a true normative relationship, and further identify the type of the norm. Fifth, we use heuristics to extract the elements of a norm—subject, object, antecedent, and consequent.

Algorithm 1 formalizes the steps described above.

3.1 Preprocess and Filter

Contract text includes information such as definitions of terms and the parties' addresses that is largely irrelevant to our purposes. We identify sentences that are candidates for expressing norms. In English, such norms are usually associated with modal verbs such as "can" and "must." We identify several *signal words*, including the modal verbs and selected contracting-specific verbs such as "warrant" and "agree." We treat a sentence that contains even one such signal as a candidate for expressing a norm and filter out those sentences that lack such signals.

To validate the above step of filtering sentences based on signal words, we examined a randomly selected contract.[3] We identified 66 sentences describing norms in this contract, of which 63 sentences included signal words. That is, signal words yield a coverage of 95% on this contract. An example sentence that lacks a signal word is "The Customer is then responsible for all disposing cost of unacceptable materials." Moreover, all the signal words indicate normative relations.

We parse each norm candidate sentence using the Stanford Parser (a leading parser [7] and freely available) and generate parts of speech (POS) tags, phrase chunks, and dependencies among the tokens in the sentence.

3.2 Compute Features

We train a classifier for norms using a set of *features* of sentences: potentially relevant to classification. These features are phrasal (from the modal phrase, e.g., main verb) or contextual (from the surrounding text, e.g., "if" signi-

[3]http://contracts.onecle.com/hansen/southeast.mfg.2004.07.24.shtml

Algorithm 1 Norm extraction.

Require: Contract collection C

Preprocess contracts
for all contract c in C **do**
    for all sentence s in c that contains a signal word **do**
        Parse sentence s to induce grammar tree t
        Build feature vector for s based on tree t
        Predict norm type for s with classification model
        Extract norm elements of s based on heuristics
    end for
end for

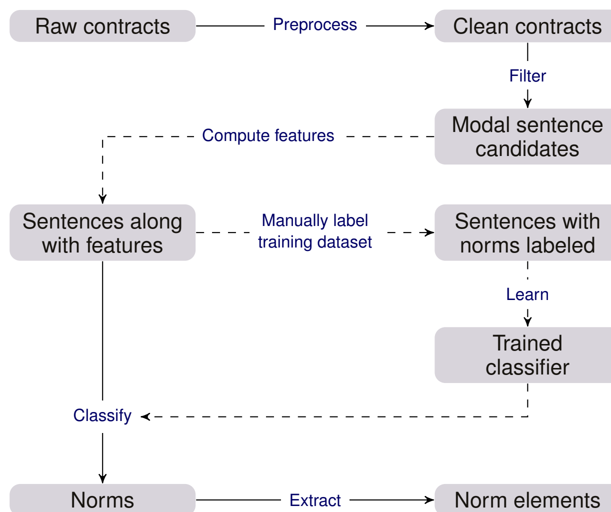

Figure 2: Approach (automatic except labeling).

fying a clause). The choice of features is crucial. In the spirit of grounded theory [10], applying our understanding of linguistics and contracts, we identify features that help us differentiate among the various norm types arising in a corpus of contracts. One feature (e.g., "may") may indicate more than one type (e.g., authorization or power). And, a sentence may include multiple features that correlate with distinct norm types. Table 1 shows the features we use. The paragraphs below describe the main features.

Table 1: Classification features.

Feature	Example
Subject contains organization	Motorola; Google
Clause signal	if; unless
Modal verb	may; should
Main verb expresses an event	deliver; perform
Main verb expresses a state	have; be
Main verb has physical impact	produce; pay
Main verb has social impact	terminate; approve
Negation present	not; neither
Only present	only
Dialectical commitment signal	warrants; understands
Practical commitment signal	agrees to
Authorization signal	the right to ⟨physical⟩
Power signal	the right to ⟨social⟩
Prohibition signal	must not
Sanction signal	responsible for breach

Grammatical subject and predicate, as the two main constituents of a sentence, carry norm-type specific information. A grammatical subject often serves as the argument of the predicate in a sentence. In contracts, participants such as companies, businesses, and organizations, often expressed as named entities, are the actors. Therefore, the appearance of a named entity in a sentence's grammatical subject helps reveal normative relationships. A practical commitment is generally directed from a grammatical subject that is a named entity. On the contrary, a dialectical commitment, which describes facts and makes general statements, often has a contractual term such as "article," "agreement," or "terms and conditions" as its grammatical subject.

Clause conjunctions express qualifications, conditioning, and modifiers. A clause led by "if" expresses a conditional dependency, whereas a clause led by "that" or "which" often expresses a modifier, and indicates the presence of a norm.

As part of the predicate, a verb identifies the action a subject performs. Vendler [27] presents four types of verbs: states and three event types, namely, activity, accomplishment, and achievement. Wu and Palmer [28] study the semantic representation of verbs. With inspiration from the above works, we divide verbs into semantic subtypes that indicate different norms. The distinction between practical and dialectical commitments, and between power and authorization often lies in the verb type. We divide verbs in two dimensions, as below.

- An *event* verb describes an action or event, e.g., "move," "perform," and "deliver." A *state* verb describes a state, e.g., "retain," "have," and "be." State verbs often indicate a dialectical commitment, whereas event verbs often indicate a practical commitment.
- A *physical* verb describes a change of physical characteristics such as location and appearance, e.g., "deliver," "produce," and "pay." A *social* verb involves changes in social or business relationships such as ownership and rights, e.g., "terminate," "change," and "approve." Physical verbs often indicate an authorization, whereas social verbs often indicate a power. For example, (from the above contract) "The Company may terminate the contract without cause and at no liability to the Company" is a power (its action verb, "terminate," indicates social consequence). And, "... [T]he Company may take possession of all Ingredients and Packaging Materials held at the Bottling Facility" is an authorization ("take possession" indicates physical consequence).

Thus, a custom lexicon that classifies verbs as "event," "state," "social," or "physical" is necessary.

Some expression patterns help predict the norm type. For instance, expressions of the form "X should be responsible for ..." and "X shall be liable for ..." indicate a practical commitment of X, whereas expressions of the form "X shall be authorized to ..." indicate an authorization. We encode such patterns as features.

A negation can transform a commitment or an authorization into a prohibition. Negations are thus relevant to classifying a norm.

3.3 Classify

Different classification approaches can be applied on the features we extracted. Naïve Bayes, Support Vector Machine (SVM), and logistic regression are state-of-art methods for text classification. Assuming the independence of tokens, Naïve Bayes estimates the posterior probability of membership in a class based on a prior probability and the available evidence. SVM solves a constrained quadratic optimization problem to find a hyperplane to separate instances of different class. Applications of SVM in text classification include news categorization and spam email detection, in which tasks SVM yields performance competitive with humans. Logistic regression predicts a categorical output via a linear function of a few predictor variables. It is widely applied in medical diagnosis in predicting disease type based on laboratory results and patient medical history. With the features as described above, we classify the normative relationship types using different classification methods, specifically as implemented in the Weka tool [12].

3.4 Extract

We extract the elements of a norm—subject, object, antecedent, and consequent—based on the following heuristics.

HEURISTIC 1. *The subject and object of a norm are usually contained in the grammatical subject and object, respectively, of a norm sentence.*

HEURISTIC 2. *If a norm sentence has a subordinate clause led by conjunction words such as "if" and "unless," the subordinate clause expresses the antecedent.*

HEURISTIC 3. *A consequent is usually conveyed in the main clause that contains a modal verb phrase.*

HEURISTIC 4. *If a sentence does not include a subordinate clause led by "if" and "unless," the antecedent is true.*

Using the above heuristics, the extraction process is automatic, largely relying on the above-mentioned grammar tree. Now we illustrate the element extraction results for a sentence expressing a practical commitment.

- *Norm sentence* (input): If future carrier customer requirements require payment of third party costs in addition to those identified above, Danger and Sharp shall discuss how to allocate such costs.
- *Subject*: Danger and Sharp (here, the norm subject is the grammatical subject). (Heuristic 1)
- *Object*: Danger and Sharp (here, "discuss" indicates a mutual relationship and, therefore, the grammatical subject is also the norm object). (Heuristic 1)
- *Antecedent*: If future carrier customer requirements require payment of third party costs in addition to those identified above (here, the antecedent is the subordinate clause led by "if"). (Heuristic 2)
- *Consequent*: Danger and Sharp shall discuss how to allocate such costs (here, the consequent is the main modal verb clause). (Heuristic 3)

4. EMPIRICAL EVALUATION

We now discuss our evaluation, including explanations of our results and the insights we gleaned from the exercise.

4.1 Experimental Results

We selected 1,000 sentences from manufacturing contracts[4] that met the modal filter. From these we removed 38 long sentences, each containing more than 80 words (usually with complex structure and not handled by the Stanford Parser) and 94 duplicate sentences (contract text is quite repetitious), leaving us with 868 sentences. The authors (assumed sufficiently competent in the study of norms and text mining) annotated the norm type for each sentence, resolving conflicting annotations via discussion. The 868 contract sentences, each labeled with a norm type, form our gold standard. The Frequency column of Table 2 shows the distribution of the norm types in the gold standard. As it shows, practical commitment, dialectical commitment, authorization, power, and prohibition account for a majority of the instances, whereas sanction and "not a norm" are rare. We conduct three experimental evaluations: (1) a ten-fold cross validation of the model, (2) a blind test set evaluation of the model, and (3) an evaluation for extracting norm elements. Additionally, to illustrate the effectiveness of our approach in extracting norms, we list some sample norms extracted by it.

Table 2: Results from ten-fold cross validation.

Type	Frequency	Precision	Recall	F-Measure
PC	223	0.87	0.80	0.83
DC	226	0.75	0.79	0.77
AU	146	0.67	0.69	0.68
PO	146	0.74	0.78	0.76
PR	85	0.64	0.68	0.66
SA	12	0.43	0.25	0.32
NN	30	0.58	0.47	0.52
Mean	–	*0.75*	*0.74*	*0.74*

[4] http://contracts.onecle.com/type/47.shtml

First, we perform a ten-fold cross-validation on our gold standard. For reasons of space, Table 2 shows the results only for LR, which are slightly better than those for SVM and NB: the latter two yield mean F-measures of 0.71 and 0.72, respectively. No classification method performs well on classes SA and NN, presumably because of the sparsity of such class instances in the dataset (there being only 12 SA and 30 NN instances out of a total of 868 sentences).

Second, because reviewing the above 868 sentences influenced our choice of features, we conducted an additional test. We gathered and manually annotated 99 norm candidate sentences from the manufacturing domain. In this evaluation, we trained our LR model on the entire set of 868 sentences and tested it on the new 99 sentences. Table 3 shows we obtained a weighted average F-measure of 84%. The model performs especially well on commitments, both dialectical and practical. The paucity of other norm types in the test dataset indicates that their results are not reliable.

Table 3: Evaluation over an independent dataset.

Type	Frequency	Precision	Recall	F-Measure
DC	41	0.90	0.88	0.89
PC	46	0.91	0.89	0.90
AU	7	0.75	0.43	0.55
PR	3	0.25	0.67	0.36
SA	1	0	0	0
PO	1	0	0	0
NN	0	0	0	0
Mean	—	*0.86*	*0.83*	*0.84*

Third, we extract norm elements. Subject and object depend on the contract: their extraction relies upon a named-entity recognizer. Extracting the antecedent and consequent is somewhat more subtle. Most times the antecedent is true: out of the 838 (= 868 − 30) norms in our gold standard, 788 norms have an antecedent of true and only 50 (6% of the total) have antecedents that are other than true. Among the 50 extracted antecedents, only one is erroneous (with respect to a manual evaluation), yielding a precision of 98%. Table 4 shows some automatically extracted antecedents. Consequents exist mostly in the context of modal phrases. Our extraction method yields high precision on a spot check. We arbitrarily selected 50 consequents extracted by our tool and manually verified each, obtaining a result of 100%, indicating the robustness of consequent extraction.

4.2 Sample Norms Extracted

With an LR model trained on the entire gold standard, we extract norms from over 1,800 contracts cutting across multiple domains. Below, we show some sample norms including their elements. An object missing from the sentence is shown as null; however, usually it is the counterparty of the subject, so it is easy to extract from the contract. Some norms show subjects that include both parties to the contract, e.g., Motorola and ASE; in such a case, the object too is the same two parties.

- Sentence: If Motorola rejects any Contract Products, Motorola and ASE shall confer to determine the reason for the rejection.
  - Norm type: practical commitment

Table 4: Sample antecedents extracted.

Antecedents
• unless the parties agree otherwise
• if Danger ceases operations but transfers or assigns all or substantially all of its assets or business to a third party successor
• if any provision of this agreement is held to be illegal or unenforceable
• if future carrier customer requirements require payment of third party costs in addition to those identified above
• unless the parties mutually agree in writing otherwise
• if the parties fail to agree on such terms and conditions
• if Intertrust's costs and expenses exceed such limits
• unless otherwise provided herein
• if the equipment is certified to be ATP ready for use prior to the day before the installation date
• if the dispute is not resolved by the mediation
• if such breach is not cured by reasonable measures by BeBig within 90 days from the date of notice

  - Subject: Motorola and ASE
  - Object: null
  - Antecedent: Motorola rejects any Contract Products
  - Consequent: Motorola and ASE shall confer to determine the reason for the rejection.

- Sentence: TieTech warrants and represents that it shall strictly adhere to the Product Specifications set forth in Appendix III attached hereto and by reference made a part hereof.
  - Norm type: dialectical commitment
  - Subject: TieTech
  - Object: null
  - Antecedent: true
  - Consequent: TieTech warrants and represents that it shall strictly adhere to the Product Specifications set forth in Appendix III attached hereto and by reference made a part hereof

- Sentence: Motorola shall have the right to approve such changes if customer approval is required.
  - Norm type: power
  - Subject: Motorola
  - Object: null
  - Antecedent: customer approval is required
  - Consequent: Motorola shall have the right to approve such changes

- Sentence: Plantronics shall have the right to audit such bill of materials and other information upon its request.
  - Norm type: authorization
  - Subject: Plantronics
  - Object: null
  - Antecedent: its request
  - Consequent: Plantronics shall have the right to audit such bill of materials and other information

- Sentence: GoerTek shall not take action to purchase materials or to manufacture Products based on any forecasts.
  - Norm type: prohibition
  - Subject: GoerTek
  - Object: null
  - Antecedent: true
  - Consequent: GoerTek shall not take action to purchase materials or to manufacture Products based on any forecasts

5. RELATED WORK

We partition the relevant literature into two main classes.

5.1 Norms and Contracts

Variations of the normative concepts have been studied in (agent-oriented as well as traditional) software engineering [2, 4, 19]. A key challenge that such approaches face is one of elicitation. Some approaches extract such concepts from regulatory text, but manually [3, 29]. Automatic extraction of norms has received only limited attention.

In MAS, studies on norms have focused on modeling [22], representation, monitoring [20], and reasoning [1]. Additionally, norms provide a powerful basis for business process modeling and requirement compliance [11]. Marengo et al. [17] propose a language REGULA for commitments in which the applicable temporal properties are encoded within commitments, which makes clear who the accountable party is. This representation is well-suited to contracts. For example, instead of "payment must occur before delivery" we would state that "the shipper commits to the seller to ensure that the customer has paid before delivering the goods to the customer." Our approach is compatible with the above intuition. Although we do not focus on temporal expressions here, our approach can be combined with that of Gao and Singh [9].

Meneguzzi et al. [18] present a normative model of contracts as applied to the aerospace domain. Their main topic is demonstrating their model and the reasoning they can perform to support the contract life cycle. Meneguzzi et al. describe the norms in contracts using natural language descriptions that they propose: these are not found in a natural contract. Lomuscio et al. [16] formalize contracts as a basis for verifying service compositions. However, their formalization does not involve norms as we have discussed them here. Identifying norms from natural language descriptions is not the focus of the above approaches, but it is clear how our approach could be used as a basis for extracting norms that could be incorporated into a formal tool chain such as those proposed by Meneguzzi et al. and Lomuscio et al.

Siena et al. [24] propose a conceptual metamodel for requirements engineering that incorporates legal prescriptions. With normative propositions, their metamodel captures goals, actions, and strategic requirements from regulations. By applying normative concepts, Siena et al. split legal statements into atomic elements to tackle the complexity of legal requirements. Their approach relies upon coming up with a normative model (an instance of their metamodel) by hand, which is tedious. Our approach, in contrast, shows how to produce norms that comply with our chosen metamodel [26]. Thus, our approach can greatly facilitate the above kinds of analyses.

5.2 Extracting Norms, Regulations, Policies

Norms as provisions in legal text have also been studied. De Maat and Winkels [6] automatically detect and classify norm types from Dutch law. They regard a sentence as a unit of a single norm, and use 88 textual patterns from about twenty Dutch laws to classify the sentences. De Maat and Winkels obtain 91% precision on 592 sentences from 15 Dutch laws. However, their approach is limited in its patterns, which sometimes introduce errors and have limited flexibility and portability across domains. Importantly, norms in the above approach, as in most previous approaches, are loosely defined and lack the focus of contract-specific features. In contrast, norm in contracts are well structured and our approach identifies and extracts their elements.

Indukuri and Krishna [13] use SVM to classify contract clauses as either related to payment or not. They propose an n-gram model (best results for four-grams) but lack linguistic modeling. Curtotti and McCreath [5] study the segmentation of contracts with a combination of rules and machine learning. They use 40 features including structural and statistical information to classify a sentence into one of 32 classes in regard to contract structural properties—definitions, addresses. However, our focus is not on the structure of a contract but the business dealings the contract specifies through normative relationships. Gao et al. [8] extract exceptions from contracts via a handful of manually crafted textual patterns, but without machine learning.

Savarimuthu et al. [23] propose to mine conventions and norms from communications among developers as recorded in open source software repositories. They propose extracting emerging norms mostly in terms of obligations and prohibitions. Our approach could potentially be leveraged to mining software documentation as well as developer communications in their setting.

Kalia et al. [15] present an approach for extracting commitments from email and chat conversations among people. In contrast to contracts in the present study, their approach deals with ad hoc communications. Their approach tackles commitments but not the other norms. Also, the commitments in their setting exhibit significantly simpler structure than those in contracts. It would be an interesting challenge to understand how norms established through contracts progress in conversations among the participants.

6. DISCUSSION

We introduce an automatic approach to extract norms from business contracts and further classify norm types. We use a hybrid of textual patterns, heuristics, and machine learning methods and obtain promising results in evaluation. In ten-fold cross validation, we obtain F-measures of 83% and 77% for practical and dialectical commitments and somewhat lower for the other normative types. This result indicates the potential viability of an automated approach in extracting contracts for requirements.

6.1 Threats to Validity

Our approach seeks to identify normative relationships expressed in contract sentences and extract their key elements (subject, object, antecedent, and consequent). Our results face the following main threats to their validity.

First, our approach has limitations due to the complexity of legal language. Important features in our approach are based on the grammar tree and, therefore, sufficiently accurate parsing is important. Long sentences often challenge state-of-the-art parsers, including the Stanford Parser, which we use. Parsing errors cascade down the pipeline and affect end-to-end precision and recall. Therefore, we restrict ourselves to sentences of up to 80 words. From the same repository that we study here, Gao et al. [8] examined a sample dataset consisting of 2,647 contracts from domains of licensing, consulting, outsourcing, supply, manufacturing, purchase, and stock options. They found that the median length of a sentence across those contracts is 58 words. Improved parsing tools would help address this threat.

Second, due to the effort required to build a large training dataset, we have annotated fewer than one thousand sentences through the efforts of two annotators. As reported above, this dataset includes only a few instances of some norm types, as a result of which our approach does not yield adequate F-measure for them. Although we sampled the sentences in our dataset from contracts of multiple domains, it is likely that we have not considered important styles of drafting contracts. A more diverse training dataset annotated by multiple annotators (ideally, without formal training in norms) could help address this threat.

Third, we focus on sentences that express one norm per sentence and tha express a norm via a distinct predicate. There are cases where one long sentence expresses multiple norms and some norms may be expressed without predicates. Such cases are not common in our experience and we excluded them because of their high complexity. In the randomly selected manufacturing contract we examined above, out of the 66 sentences that bear normative relations, four sentences bear more than one norm, constituting 6% of the total. The following example sentence shows a dialectical commitment about a power and a prohibition: "The Customer agrees that the Company, in its sole discretion may change the Bottling Schedule upon five (5) business days written notice to the Customer, however, should not delay the production more than fifteen (15) days."

An associated threat to validity is that if we encounter contract drafting styles that involve such formulations in larger frequencies, our approach would prove less effective than it does at present.

6.2 Significance and Future Directions

Traditional software engineering (and its subarea, requirements engineering) generally presume a closed system where the IT artifacts necessarily serve the interests of the stakeholders. In cross-organizational settings, however, the parties involved are autonomous and heterogeneous. In such settings, norms are essential because no party controls the implementation or operation of another. The agents research community has long argued for the applicability and value of multiagent systems approaches to practical problems in open, cross-organizational settings. However, despite the strong arguments in favor of multiagent systems, their application has always been resisted because of the perceived cost of eliciting normative models.

This paper is about exploiting an opportunity in helping build normative models based on legacy documents, namely, contracts. Not only does a contract describe the services and value exchanges between the participants, it is also expressed in a familiar form. Further, a contract is presumably accepted (or being considered for acceptance) both by management and by technical staff. Thus using a contract as a basis for formalization would not be controversial.

Several directions can be explored based on our extraction results. First, a major practical challenge is the development of suitable datasets to support extending our method to multiple norm types, and to conduct a deeper investigation of the features arising in sentences of different norm types.

Second, our approach would enable building a repository of norms from different domains to provide examples and templates for authoring specifications. The consequent of one norm may be the antecedent of another: a chain of dependencies often exists in contracts whereby the success or failure of one business action triggers another. The study of such interdependencies can help discover and abstractly describe the business processes associated with contracts.

Third, it would help to capture criteria on requirements in normative terms and validate requirements gleaned from a contract with respect to such criteria. For instance, a poorly authored contract may contain no prohibitions or sanctions and thus lack the teeth to ensure compliance. Our norm extraction approach can help determine whether the requirements as expressed in contracts satisfy important criteria, such as being enforceable or providing coverage against each possible normative violation.

Independently of the specific technical approach one may adopt, the study of norms as a foundation for requirements for interactions among autonomous parties can have far-reaching consequences on software engineering. Analyzing contracts is a natural entry point for this program of research to expand the practical reach of multiagent systems.

Acknowledgments

Thanks to Dennis Bahler, Amit Chopra, Robert Handfield, Chung-Wei Hang, Chris Healey, Anup Kalia, Pankaj Telang, and the anonymous reviewers for helpful comments on previous versions. This work was partially supported by the Army Research Laboratory in its Network Sciences Collaborative Technology Alliance (NS-CTA) under Cooperative Agreement Number W911NF-09-2-0053 and was partially supported by U.S. Department of Defense under the Science of Security Lablet grant.

7. REFERENCES

[1] H. Aldewereld, S. Álvarez-Napagao, F. Dignum, J. Vázquez-Salceda. Making norms concrete. *Proc. AAMAS*, pp. 807–814, 2010.

[2] A. Artikis, M. Sergot, J. Pitt. Specifying norm governed computational societies. *ACM TOCL*, 10(1):1:1–1:42 2009.

[3] T. Breaux, A. Antón. Analyzing regulatory rules for privacy and security requirements. *IEEE TSE*, 34(1):5–20, 2008.

[4] A. K. Chopra, E. Paja, P. Giorgini. Sociotechnical trust: An architectural approach. *Proc. Conf. Conceptual Modeling*, LNCS 6998, pp. 104–117, 2011.

[5] M. Curtotti, E. McCreath. Corpus based classification of text in Australian contracts. *Proc. Aust. Lang. Tech. Assoc. Wkshp*, pp. 18–26, 2010.

[6] E. de Maat, R. Winkels. Automated classification of norms in sources of law. *Proc. Semantic Processing of Legal Texts*, LNCS 6036, pp. 170–191, 2010.

[7] M. De Marneffe, B. MacCartney, C. Manning. Generating typed dependency parses from phrase structure parses. *Proc. LREC*, pp. 449–454, 2006.

[8] X. Gao, M. Singh, P. Mehra. Mining business contracts for service exceptions. *IEEE TSC*, 5(3):333–344 2012.

[9] X. Gao, M. Singh. Mining contracts for business events and temporal constraints in service engagements. *IEEE TSC*, 2013. In press. http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6482126

[10] B. Glaser, A. Strauss. *The Discovery of Grounded Theory*. Aldine de Gruyter, 1967.

[11] G. Governatori, A. Rotolo. Norm compliance in business process modeling. *Proc. Conf. Semantic Web Rules*, pp. 194–209, 2010. Springer-Verlag.

[12] M. Hall, E. Frank, G Holmes, B. Porringer, P. Reutemann, I. H. Witten The WEKA data mining software: An update. *ACM SIGKDD Explorations Newsletter*, 11(1):10–18, 2009.

[13] K. V. Indukuri, P. Radha Krishna. Mining e-contract documents to classify clauses. *Proc. Bangalore Compute Conf.*, pp. 1–5, 2010.

[14] A. Jones, M. Sergot. A formal characterisation of institutionalised power. *Log. J. IGPL*, 4(3), 1996.

[15] A. Kalia, H. Motahari Nezhad, C. Bartolini, M. Singh. Monitoring commitments in people-driven service engagements. *Proc. SCC*, pp. 160–167 2013.

[16] A. Lomuscio, H, Qu, M. Solanki. Towards verifying contract regulated service composition. *JAAMAS*, 24(3):345–373, 2012.

[17] E. Marengo, M. Baldoni, A. K. Chopra, C. Baroglio, V. Patti, M. P. Singh. Commitments with regulations. *Proc. AAMAS*, pp. 467–474 2011.

[18] F. Meneguzzi, S. Modgil, N. Oren, S. Miles, M. Luck, N. Faci. Applying electronic contracting to the aerospace aftercare domain. *Eng. Appl. AI*, 25(7):1471–1487, 2012.

[19] N. H. Minsky, V. Ungureanu. Law-governed interaction. *ACM TOSEM*, 9(3):273–305, 2000.

[20] S. Modgil, N. Faci, F. Meneguzzi, N. Oren, S. Miles, M. Luck. A framework for monitoring agent-based normative systems. *Proc. AAMAS*, pp. 153–160, 2009.

[21] T. J. Norman, D. V. Carbogim, E. C. W. Krabbe, D. N. Walton. Argument and multi-agent systems. In C. Reed, T. J. Norman, eds., *Argumentation Machines*, ch. 2, Kluwer, 2004.

[22] O. Pacheco, J. Carmo. A role based model for the normative specification of organized collective agency and agents interaction. *JAAMAS*, 6(2):145–184, 2003.

[23] B. Savarimuthu, S. Cranefield, M. Purvis, M. Purvis. A data mining approach to identify obligation norms in agent societies. *Proc. ADMI*, pp. 54–69, 2010.

[24] A. Siena, A. Perini, A. Susi, J. Mylopoulos. A meta-model for modelling law-compliant requirements. *Proc. Wkshp. Req. Engg. & Law*, pp. 45–51, 2009.

[25] M. P. Singh. An ontology for commitments in multiagent systems. *AI & Law*, 7(1):97–113, 1999.

[26] M. P. Singh. Norms as a basis for governing sociotechnical systems. *ACM TIST*, 5(1)21:1–21:23, 2013.

[27] Z. Vendler. Verbs and times. *Phil. Rev.*, 66(2):143–160, 1957.

[28] Z. Wu, M. Palmer. Verbs semantics and lexical selection. *Proc. ACL*, pp. 133–138, 1994.

[29] J. Young. Commitment analysis to operationalize software requirements from privacy policies. *Requirements Engineering J*, 16(1):33–46, 2011.