Checking compliance of a system with regulations : towards a formalisation

Laurence Cholvy and Claire Saurel

ONERA Centre de Toulouse
2 avenue Edouard Belin
31055 Toulouse, France
{cholvy,saurel}@cert.fr

Abstract. This paper addresses the problem of checking if an updated system is in compliance with the current regulations which apply on the domain.
We first present the applicative context in which this problem has been met. We sketch a formalisation of the problem of compliance and we show that is can be split in several sub-problems of different types, the solutions of which are discussed.

Keywords Compliance, regulations, formalisation.

1 The applicative context

This study is part of the ONERA program named IESTA[1], which aims at developping models, methods and a platform of simulation for analysing innovative concepts of Air Transportation Systems (ATS). This platform will allow its customers (air companies, aircraft manufacturers, official regulation providers, research laboratories and relevant ATS stakeholders) to virtually modify some parameters of the ATS (for instance modify landing procedure, modify types of airplanes, equipe aircrafts with new kinds of fuel...) and analyse the impacts of these scenarii on some environmental metrics.

In particular, this platform will allow its users to study the impacts of expected ATS modifications, on the noise level and chemical emission in the vicinity of a given airport. Modifications of interest will thus be the ones which, according to the simulation, lead to a reduction of noise or/and pollution levels.

In this context, we focus on the particular problem of the compliance of these modifications with the current regulations.

For instance, assume that the simulation shows that, for such a given aiport, modifying the landing procedure in such a given way leads to noise reduction. Before discussing the adoption of this interesting modification, it would be helpful to help users to check if it complies with the current regulations, or if it is incompatible with them or if it is not even ruled.

This is the problem we address.

More generally, the problem we are interested in can be described as follows:

[1] with financial support of DGA, FNADT, FEDER and Région Midi Pyrénées

- given a system composed of several components,
- given the set of regulations which rule this system,
- given a modification which is proposed for the system (modification concerning its structure or the way of performing its function),
- we first want to be able to check if the modified system is in compliance with regulations. If it is not, we want to help users to understand where are the causes of non compliance. Users will then have to revise the given regulations or to revise the proposed modification.

This paper is organized as follows. Section 2 tends to formalise the problem of checking compliance of a system with regulations. Section 3 analyses more deeply this problem and shows that it can be split in several sub-problems of different types. Section 4 focuses on the problem of providing the users an assistance to revise violated regulations. Section 5 mentions some relevant works, the scientific domains they belong to (Information Retrieval and Normative Reasoning) being candidate to provide us with solutions. Section 6 concludes this paper.

2 Towards a formalisation of the problem of checking compliance

The variables of the problem we address are the following.

Definition 1. *new denotes the updated system the compliance of which has to be checked.*

Definition 2. *KB denotes the background knowledge, i.e the knowledge about the considered environment.*

Example 1. In the ATS case, if the problem is to check the compliance of the ATS when aircrafts are given a new fuel, then KB may include characteristics and properties of this new fuel (density, volumic mass, inflammation temperature...), but also characteristics of airport environment (atmospheric pressure, physical models...). Any information describing the modified ATS when aircrafts are given a new fuel is in *new*.

¿From a more formal point of view, *new* and KB should be modelled in a common model. For instance, KB could be modelled by ontologies, hierarchy of concepts, dependance graphs between concepts, or more generally, and this will be supposed in the rest of the paper, by logical formulas. In the same way, *new* could be modelled by sets of nodes in the ontologies, sets of concepts, or more generally, and this will be supposed in the rest of the paper, by logical formulas.

Assumption In the following, we will suppose that *new* is compatible with KB i.e, $KB \cup new$ will be supposed to be a consistent set of formulas.

Consistency is a prerequisite to the problem of compliance since, if $KB \cup new$ is inconsistant, this means that *new* contradicts the domain knowldege, thus, building *new* is impossible. Consequently, the question of checking his compliance is not posed.

Definition 3. *A regulation r is a triplet:* $r = < st_r, ref_r, norms_r, def_r >$*, where*

- st_r *is an ordonned list of the levels which structure the text of r in an arborescent way.*
- ref_r *denotes the set of other regulations r refers to.*
- $norms_r$ *denotes the normative contents of r. If n is in* $norms_r$*, n is assigned a position label related to* st_r *which denotes its position in the arborescent structure of r.*
- def_r *is a set of the definitions of the concepts which are used in the regulation text.*

Note that several members of $norms_r$ may have the same position label. It means that they are in the same text unit in r.

What we want here to capture is that a regulation contains information of very different natures :

- conceptual definition information (def_r) : that is an optional part of the regulation. Concepts which are ruled by r are defined in def_r.
- rules ($norms_r$) : that is the core of the regulation. The rules apply on the real world by stating what is obligatory, permitted or forbidden under which conditions. Formal modelling rules requires the use of a logical formalism dedicated with normative reasoning i.e a deontic logic (see section 5). In the following, we will thus suppose that these rules are modelled by formulas of such a logic.
- In a regulation r, conceptual definition and rules are expressed according a given structure (st_r).
- information about other regulations ref_r : regulations which inspire the regulation, regulations which are abrogated by the regulation...One generally mainly find it in the head of the text of r.

Example 2. For r being [2], $st_r = [article, alinea]$. That means that r is composed of several articles, each of them being eventually composed of alineas.

Example 3. Consider now r_0 a regulation such that $st_{r_0} = [part, subpart, article]$. If the formula n is in the 2d article of the 3rd subpart of the 1st part of r_0, then the position label assigned to n is $[(part, 1), (subpart, 3), (article, 2)]$.

Example 4. Finally, for r being [2], ref_r includes regulations [16] and [3].

Definition 4. $R = \{r_1, ...r_n\}$ *denotes the set of all the regulations which apply on the domain.*

Example 5. In the ATS case, the set R of regulations which rule the aeronautic domain is composed of CEE regulations, national regulations such as environment code, civil aviation code, and lots of orders and procedures.

Several relations exist in R such as:

- $r_2 \geq_S r_1$ is true if regulation r_1 is a specialization of regulation r_2. $\geq_S$ is a partial pre-order defined upon R.
- $r_2 \geq_A r_1$ is true if regulation r_2 abrogates regulation r_1 : it means that r_1 doesn't apply anymore since r_2 applies. $\geq_A$ is a partial pre-order defined upon R.
- a binary relation *replace* so that $replace(a_{i_j}, a_{k_l})$ is true if the i^{th} article of regulation r_j replaces the k^{th} article of regulation r_l.

Some of relations of this type has been presented in [4]. Our model represents a slightly simplified form of real regulations, since as for instance, $\geq_A$ could be defined between text units of regulations, and not only between regulations (as for the relation *replace*) . We here suppose that instances of such relations concerning a regulation r are explicited in ref_r.

Example 6. Let r_1 and r_2 respectively being regulations [2] and [16]. $r_2 \geq_S r_1$, because the regulation dealing with rules concerning the Blagnac airport specializes the regulation about french public air transport.

Definition 5. *If $R = \{r_1, ..., r_n\}$ is the set of regulations which apply on the domain, we define $norms_R$, as the set of all the rules of all the regulations of R, i.e, $norms_R = \cup_{i=1}^{n} norms_{r_i}$*

$norms_R$ is thus a set of formulas of a particular deontic logic.

More formally, we assume that a formal model (formal language and formal inference denoted $\models$ in the following) has been chosen for modelling and reason with rules of $norm_R$, background knowledge KB and modification new[2].

In the rest of the paper, we will suppose that $norm_R$ is consistent i.e is a consistent set of rules.[3]

[2] Ideally, this formal model is a logic which allows to express and reason with any type of deontic notions which appear in regulations, any type of knowledge, causal or temporal, which appear in KB. Such a logic should then be a deontic logic ([5], [8]) allowing to reason with causality and time as well. Defining such a general logic remains to be done.

[3] Notice that consistency of sets of rules has been defined in [6] so that, $norm_R$ is a consistent set of rules if there is no situation (or state of the world) s, consistent with KB (i.e possible) such that : $s \cup norms_R \models false$. This general definition is not taken here for simplicity, but notice that if $norm_R$ is a consistent set of rules according to this definition, then $norm_R$ is a consistent set of rules

Checking compliance is then defined by checking one of the following assertions:

1. "case of permitted modification"

   $$\forall\phi \quad KB \cup new \models \phi \quad \implies \quad \models norms_R \rightarrow permitted(\phi)$$

   This expresses that all the consequences the system modification *new* (under context KB) are explicitely permitted by the rules in the regulations. In the first case, *new* could be accepted without any other modification since it is compliant with the regulations.

2. "case of forbidden modification"

   $$\exists\phi \quad \exists r \in R \quad KB \cup new \models \phi \quad and \quad \models norms_r \rightarrow forbidden(\phi)$$

   This expresses that the system modification, *new*, has some consequences (under context KB) which are explicitely forbidden by one regulation In this case, *new* cannot be taken into account since it explicitely leads to violate regulations, unless modifying regulations themselves. Localizing the very rules which are violated by *new* is addressed in section 4.

   Notice that in the general case, the prohibition is not caused by only one regulation. So the case of forbidden modification should be described by:

   $$\exists\phi \quad KB \cup new \models \phi \quad and \quad \models norms_R \rightarrow forbidden(\phi)$$

   However, in this paper, we assume that the prohibition is caused by a single regulation because it simplifies the presentation of localizing violated rules (see section 4).

3. "case of a non ruled modification"

   $$\exists\phi \quad KB \cup new \models \phi \quad and$$

   $$\not\models norms_R \rightarrow permitted(\phi) \quad and \quad \not\models norms_R \rightarrow forbidden(\phi)$$

   This expresses that the system modification, *new*, has some consequences (under context KB) which are neither explicitely permitted nor explicitely forbidden by the regulations. In this case, it will be possible to accept *new* only after an analyse and modifications of regulations so that consequences of *new* are permitted.

By definition, these three assertions are exhaustive. Furthermore, they are exclusive only if $norm_R$ is consistent. This is the reason why assuming consistency of rules is a prerequisite to the definition of compliance.

3 Decomposing the problem of checking compliance

Checking compliance can be decomposed into several sub-problems. The idea is to check compliance only on a subset of R, made of regulations which "apply at present" and "concerned by *new*". At this step, the property "being a regulation concerned by *new*" remains to be formally defined. This property could be defined so that the test of checking compliance is more efficient in time. It could also be defined so that we can help the user (in the second case) to find the precise articles of the regulations that are violated.

- **Problem pb 1 : find the "regulations which could be violated"**
  This problem can be divised into two sub-problems as follows:

  - **Problem pb 1.1 : find the "regulations which apply at present"**
    This problem consists in selecting the regulations which are not abrogated nor replaced by other regulations. In other words, the problem is to focus only on the regulations that apply at the moment.
    This problem may be defined by : find $max_{\geq_A}(R)$ [4]
    Information in $\cup_{i=1}^{n} ref_{r_i}$ will be hepful to solve this sub-problem .

  - **Problem pb 1.2 : find the "regulations concerned by *new*"**
    This problem is a problem of Information Retrieval, the information to be retrieved being regulations.
    In order to solve it, considering information in $\cup_{i=1}^{n} def_{r_i}$ (i.e definitions of the concepts used in the regulation text) will be necessary.

  The two above sub-problems may be solved in any sequence order : each of them contributes towards reducing the set of regulations to be considered in checking compliance.

  Let us denote R_{new} the set of the regulations of R which apply at present and which are concerned by *new*.

- **Problem pb 2 : checking compliance of *new* with R_{new}**

  This comes to check the three assertions:

$$\forall \phi \quad KB \cup new \models \phi \quad \Longrightarrow \quad \models norms_{R_{new}} \rightarrow permitted(\phi)$$

$$\exists \phi \quad \exists r \in R_{new} \quad KB \cup new \models \phi \quad and \quad \models norms_r \rightarrow forbidden(\phi)$$

$$\exists \phi \quad KB \cup new \models \phi \;\; and \;\; \not\models norms_{R_{new}} \rightarrow permitted(\phi) \;\; and \;\; \not\models norms_{R_{new}} \rightarrow forbidden(\phi)$$

[4] If $\geq$ is a partial pre-order defined on R, then $max_{\geq}(R)$ is defined by: $max_{\geq}(R) = \{r \in R : \forall r' \in R, r' \geq r \Rightarrow r \geq r'\}$

4 Towards assisting localization of violated rules

In this section, we suppose the case when *new* doesn't comply with R_{new}. I.e, we assume that the second assertion of problem pb 2 is true.

Definition 6.

$$R_{Forbidden} = \{r \in R_{new} \quad , \exists \phi \quad KB \cup new \models \phi \quad and \quad \models norms_r \rightarrow forbidden(\phi)\}$$

$R_{Forbidden}$ denotes the set of regulations which are involved in the cause of non compliance of *new* with R_{new} under KB.

In order to assist users in revising such regulations, several kinds of aids may be proposed to him. Below we sketch some induced problems.

- **Problem pb 3 : localize a cause of non compliance in a regulation** Let $r \in R_{Forbidden}$. This problem consists in:

  1. exhibiting the elements in $norms_r$ which are involved in a demonstration of "forbidden modification".
  2. finding their position label in r (according to st_r, as defined in definition 3).

- **Problem pb 4 : localize the least specialized regulations involved in non compliance**
  The problem is to identify the uppest regulations involved in non compliance, towards the specialization relation defined on R : in other words, it is to identify sources (in the specialization or hierarchical sense) of non compliance.

  Formally : find $max_{\geq_S}(R_{Forbidden})$

- **Problem pb 5 : propagate a cause of non compliance in a set of regulations**
  The problem is, given a regulation involved in non compliance, to identify all the regulations which specialize it. These regulations, because they take their inspiration from the violated regulations, are are also involved as causes of non compliance.

  Formally : let $r \in R_{Forbidden}$, find $\{r' \in R_{Forbidden},\ r \geq_S r'\}$.

- **Problem pb 6 : explanation for non compliance** The problem is to give an informative explanation based upon non compliance demonstrations. This comes to a problem of Explanation Generation, which has been studied for many years [20], [1].

5 Relevant works

Among the different sub-problems we have raised in the previous sections, two of them are of particular interest. More specifically, these are: a problem of information retrieval, the information to be retrieved are regulations (cf problem 1.2) and a problem of normative reasoning (cf problem 2). These two very different questions have been addressed by many works we mention some of them below.

5.1 Information retrieval, regulation retrieval

Information Retrieval is a vast domain of research whose works aim to define models and methods or algorithms, to retrieve information among a large set of information, like the web space. See [22] for an interesting overview. The user's demand is formalised by a query Q (of the form "retrieve documents which contains terms $t_1 ... t_n$").

The three most used models in Information Retrieval are the vector space model, the probabilistic model and the inference network model.

According to the vector space model, the user query as well as the documents the query is addressed to, are represented by vectors of terms (words of a given vocabulary for instance). The score of a document is defined as a similarity degree between its vector and the query vector. Several similarity degrees are usually used, among which the dot product defined by: if $\boldsymbol{D}$ denotes the document vector and $\boldsymbol{Q}$ denotes the query vector, then the score of D for Q is:

$$sim(D, Q) = \sum_{i=1}^{n} d_i . q_i$$

where d_i is the value of the i^{th} component of (D) and q_i is the value of the i^{th} component of (Q). The value d_i (resp q_i) is called the weight of the di^{th} term in the document (resp, query).

Various methods for weighting terms have been defined. All of them are based on different parameters which are : term frequency (words that repeat several times in a text are considered salient), document frequency (words that appear in many documents are considered common and are not very indicative of document content), the number of documents that contain a given term, the document lenght (in bytes), the average document length (in bytes)...

As for Probabilistic models, they assume that documents in a collection should be ranked by decreasing probability of their relevance to a query (*probabilistic ranking principle*). Since knowing its true value is impossible, the probability of relevance of a document to a query has to be estimated. In this family, the models differ from the way they estimate that probability of relevance.

Last models are Inference network models. In these models, document retrieval is modeled as an inference process in an inference network.

Since we consider Regulation Retrieval as a particular case of Information Retrieval, solving pb 1.2 could be done by adapting a model of Information Retrieval. For doing so, the definitions of concepts used in a regulation (i.e the

def_r part) has obviously an important role to play in the process. Furthermore, the very structure of regulations (i.e the st_r part) is also something particular which must be taken into account by Regulation Retrieval models to be defined ([4], [9]).

Let us also mention [10], in which the authors define a tool to analyse a regulation and extract rights and duties expressed in the regulation. Legal texts are annotated in order to identify the agents, their rights (actions that the agents have the permission to perform under come conditions) and duties (the actions they have to perform under some conditions)... A semantic model in then built from these annotations.

Let us finally cite [13], in which the author defines a legal ontology of the french Law. Such an ontology could be used as a common concept description language and links with the def_r part of regulations should be establised.

5.2 Reasoning with regulations, normative reasoning

Problem pb 2 raises the question of checking if a given formula, (here $permitted(\phi)$ or $forbidden(\phi)$), is implied by some rules (here $norms_{R_{new}}$). This is a particular case of what is called "normative reasoning" i.e, reasoning with norms.

Reasoning with norms requires at least modelling deontic notions (permission, prohibition, obligation...). But it also requires modelling individuals (agents on which obligations, permission and prohibition apply) and properties on individuals. It sometimes also require modelling several dimensions of time (time of validity of norms, deadlines...) and different types of norms (defeasible norms, Contrary-to-duties...). To our knowledge, there is no general logical formalism which allow to reason with so many different notions. However, there are several kinds of formalisms which allow to model some aspects of the norms. These are deontic logics [8].

Most of deontic logics are modal ones, [5], since deontic operators are not very-functional operators (for instance, it may be the case that smoking is forbidden, even if somebody is smoking). Some of them are based on dynamic logics [15], or based on temporal logics, or both [7]. They also may be non monotonic [11], [23].

However, First Order Logic (FOL) can also been used to reason with deontic notions ([21], [14], [17], [19], [12]...) and is a compromise between expressivity and simplicity. In this case, normative reasoning comes to a problem of theorem proving in FOL which is solved (at least from a theoretic pioint of view) by different means: provers based on Resolution Rule, tableaux methods, or any method defined for the SAT problem.

Let us finally mention a very theoretical but interesting work, [18], in which the authors define a dynamic deontic logic for reasoning with consequences on permissions and prohibitions, that the modification of a policy generates. This aims at helping the user who wants to modify a regulation, by allowing him/her to derive the permissions which were valid before the modification and which are no more valid after; or the permissions which become valid after the modification etc.

6 Conclusion

In this paper we have addressed the problem of checking compliance of an updated system with regulations. The main contributions are a formalisation of this problem and its decomposition in several sub-problems of different types. We also sketched some functionalities which could help a user to revise regulations in case of non compliance. We finally quickly presented relevant litterature, more precisely Information retrieval and Normative Reasoning, which could offer solutions.

Notice however that this work is very preliminary and thus raises many open questions, the most important one being the definitions of the solutions of the different sub-problems and their applicability as well. The case named "case of a non ruled modification" has also to be studied. And the model of regulations used in this work, has to be refined in order to take into account a finer granularity of representation. Indeed, for instance, this model does not allow to represent relations between text units .

However, even preliminary, this work enlights the complexity of the problem of checking compliance and the varieties of questions to be solved.

Acknowledgments. We would like to thank the anonymous reviewers for their comments which helped us to improve the paper.

References

1. F. Benamara et P. Saint Dizier. WEBCOOP: A Cooperative Question-Answering System on the Web. 10th European Chapter of the Association of Computational Linguistics (EACL), Budapest, Hongrie, avril 2003.
2. Arrêté du 21 mars 2003 portant restriction d'exploitation de l'aérodrome de Toulouse-Blagnac (Haute-Garonne), NOR: EQUA0300268A
3. Réglement (CEE) n 2408/92 du Conseil, du 23 juillet 1992, concernant l'accès des transporteurs aériens communautaires aux liaisons aériennes intracommunautaires
4. B. Chabbat. Modélisation multiparadigme de textes réglementaires. Thèse soutenue le 8 décembre 1997, à l'INSA de Lyon.
5. B. F. Chellas Modal logic, an introduction. Cambridge University Press, 1980.
6. L. Cholvy Checking regulation consistency by using SOL-deduction Proc. of AI and Law conference, Oslo, Norway, june 1999.
7. F. Dignum and R. Kuiper. Combining dynamic deontic logic and temporal logic for the specification of deadlines. In Jr. R. Sprague, editor, Proc. of thirtieth HICSS, Wailea, Hawaii, 1997.
8. Proceedings of the DEON workshops. http://deon2008.uni.lu/publications.html
9. D. Jouve Modélisation sémantique de la réglementation Thèse de l4INSA de Lyon, novembre 2003
10. N. Kiyavitskaya et al. Extracting Rights and Obligations from regulations: Toward a Tool-supported Process 22^{nd} IEE/ACM Int. Conf. on Automated Software Engineering (ASE'07), Atlanta, november 2007.
11. J. Horthy. Deontic Logic as Founded on Nonmonotonic Logic Annals of Mathematics and Artificial Intelligence, 1993.

12. J.Y. Halpern, V. Weissman. Using First-Order Logic to Reason about Policies Proc. of the 6th IEEE Computer Security Foundations Workshop CSFW-03, 2003.
13. Guiraude Lame. Construction d'ontologies partir de textes.
Une ontologie du droit dédiée à la rechecrhe d'information sur le web.
14. R. Lee. Bureaucracies as deontic systems *ACM Transactions on Information Systems (TOIS)* 6(2), pp 87 - 108 , April 1988.
15. J.-J.Ch. Meyer. A different approach to deontic logic: Deontic logic viewed as a variant of dynamic logic. In Notre Dame Journal of Formal Logic, vol.29, 1988.
16. Arrêté du 12 mai 1997 relatif aux conditions techniques d'exploitation d'avions par une entreprise de transport aérien public (OPS1)- NOR: EQUA9700893A
17. Ong, R. Lee Detecting deontic dilemmas in bureaucratic rules: a first-order implementation using abduction Proc of *Proceedings of the second Workshop on Deontic Logic and Computer Science (DEON'94)*, Oslo, 1994.
18. R. Pucella, V. Weissman. Reasoning about Dynamic Policies. In proc. of the 6th Int. Conf. on Foundations of Software Science and Computation Structure (FOSSACS 04), 2004.
19. U. Ryu, R.M. Lee. Defeasible Deontic Reasoning and Its Applications to Normative Systems. *Decision Support Systems*, 14(1), pp. 59-73, 1995.
20. Cl. Saurel. Méthode de génération d'explications négatives à partir d'une base de connaissances en logique des prédicats 8èmes journées internationales sur les systèmes experts et leurs applications, Avignon, 1988.
21. M.J. Sergot. Prospects for representing the law as logic programs. . In *K.L. Clark and S.. Tarnlund, editors, Logic Programming* , pp 33–42. .Academic Press, London, 1982.
22. A. Singhal. Modern Information retrieval: a brief overview IEEE Data Engineering Bulletin 24(4), 35-43, 2001.
23. L. van der Torre. Violated obligations in a defeasible deontic logic. Proc of ECAI 1994.