First Steps Towards a Formal Analysis of Law

Tom Maarten van Engers
Leibniz Center for Law
University of Amsterdam
Amsterdam, the Netherlands
e-mail: vanengers@uva.nl

Robert van Doesburg
Immigration and Naturalization Service
Rijswijk, the Netherlands
e-mail: r.v.doesburg@ind.minvenj.nl

Abstract—In this paper, the authors present some work recently done within the Dutch Immigration and Naturalization Service (IND). Being responsible for the implementation and execution of complex and ever changing regulations, for which the use of IT systems is a necessity, this organization has become aware of their dependence on trustworthy methods to assure the correct implementation of law into their operations and services. While many attempts to automate law, even in the domain of migration law, have been made before, hardly any attention has been paid to the ‘translation process’ from legal rules expressed in natural language to specifications in computer executable form. In this paper, we will explain the method we have developed and illustrate its application with some concrete examples. The work is part of a larger innovation programme initiative that we collaboratively conduct within a virtual collaboration, called the ‘Blue Chamber’.

Keywords—knowledge acquisition; legal engineering; legal analysis; Hohfeldian analysis.

I. INTRODUCTION

Making a formal analysis of law is problematic because sources of law contain huge amounts of implicit information. As a result, law can be difficult to understand and to interpret. In countries that follow common law tradition understanding the law requires the knowledge of a growing body of potentially relevant preceding cases, next to the knowledge of ‘black letter’ law, including bills such as tax law, or immigration law. The authors, living in a civil law country, conduct their research in governmental organizations. In civil law cultures, the most important legal sources are the laws that are produced by the parliamentary system and the sources of law that are produced by legal bodies based on delegated powers. Case law is relevant as well, especially in explaining the ‘correct’ interpretation and application of legal sources onto actual cases. But rather then taking case decisions as a primary source, institutions like the IND describe the legal consequences of court decisions in changed regulation.

Translating sources of law into formal specifications for IT systems is a necessary step for (partial) automation of public services. The translation process of sources of law should be transparent to make sure that legal and policy experts are able to validate the correctness and completeness of the ‘legal rules’ the IT system. The authors have also pointed at the importance of inter-coder independency and scalability with respect to human resources [1]. The method to be developed should allow us to work with many knowledge analysts and knowledge engineers that have normal professional skills. Like us, other people have identified the need for a systematic approach that would allow to cope with law, and enable its translation into formal models that consequently could be executed by IT systems. In fact the very idea of translating law into computational models that we could use for solving cases using a computer, goes back to Gottlieb Wilhelm Leibniz who even build a mechanical reckoner that was supposed to solve complex (legal) problems once we would have the magical numbers of the legal concepts. Leibniz before Wilkins, both deserving recognition as founding fathers of modern computer science, stressed the importance of formal conceptualizations, these days usually referred to as ontologies, centuries before computer science came into existence [9][17].

Shortly after the invention of the modern computer, scientist recognized its power as a symbol processor allowing it to be used for reasoning processes. This includes reasoning about legal cases. Consequently, people have been working on using IT systems within the domain of law, particularly within public administrations. In literature, one can find quite some research papers on systems aimed at deciding on legal cases. The authors of this paper also have been working on such legal decision-support systems [2][3].

The early work of Sergot et al., which is also within the domain of immigration law, like the work we present in this paper, uses the expert as the main source of legal knowledge, despite of its title: ‘the British Nationality Act as a Logic Program’ [13][14]. In the POWER-programme [4], conducted in the late nineteen-nineties by the authors, written sources of law were the main source of (legal) knowledge. Legal experts did, and do still, play an important role, but only as interpreters and validators of knowledge that can be traced back to sources of law.

In a series of projects since the POWER-programme, the authors have been systemizing the translation process using (semi) automated norm extraction [1][11][12], particularly by looking at invariant language patterns typically used in written sources of law. Using computational linguistics we were able to identify the most important patterns and showed that we could use parsing to ‘translate’ written sources of law.
written in natural language into model sentences in a formal language.

Also, we have been working on the representation of norms in ways that enable multiple task contexts and multiple agents perspectives. The typical single task orientation used in Sergot et al. could be avoided. In order to achieve this objective, we used formal models [5][6] that are based upon an extended version of Hohfeld’s model [8]. In previous research projects we have showed that Hohfeld’s initial model is already a big improvement compared to traditional interpretation of rights, duties and allowances in (modal) logic. Our extended and completely formalized version of Hohfeld’s model enables us to express all typical jural relations in a formal way, but it is also expressed in a relational model that can be implemented in a straight forward way. Furthermore, we worked on the development of an agent-role based model, allowing us to reason about the consequences of norms in a social context [15][16]. We certainly do not claim that we have solved all issues, but by working on these related topics the depth and complexity of understanding law became much clearer to us and has inspired us to continue our quest.

One of the issues that we did not address thus far was the scoping problem. While legal experts are perfectly able to list the regulations that are relevant to solving a legal problem of some kind (or at least claim to be), we have experienced that one of the problems for our knowledge analysts was where to start analyzing, what rules of law to include, and where to stop looking for additional sources of law. Obviously, a top down approach, analyzing all sources of law in a country, could not possibly work since there are simply too many sources of law available to allow for an analysis that finishes within a reasonable time. Furthermore, the concept ‘top down’ would be problematic. Although the constitution normally is considered to be the highest source of law (in the Netherlands just after ‘the grace of God’ that gives the formal power to the King), we are bound by even higher forces, such as international treaties. In the POWER-programme we have experienced that a serial approach, where analysts worked through a source of law from the first to the last article, was both time consuming and required an integration step to ‘glue’ the different partial models together. At that time we did not have an explicit method for doing that and much was left to the insights of the analysts.

So the challenge we took up was to make explicit the issue of scoping relevant sources of law. If we would have a method that solves our scoping problem in such way that it is coder independent, and results in a model that could be mapped to the original sources (like we aimed for in the POWER-programme), we would be a step closer to our final aims, i.e. a method for the formal analysis of law.

In this paper, we describe the approach we have developed and explain it by illustrating its application on Dutch Alien Law. We will explain the issues that were raised during the analysis and explain their relevance.

In Section 2, we give a short description of recent IT developments within the IND. In Section 3, we give a brief overview of legal sources relevant to the domain. In Section 4, we present a scoping procedure for sources of law. Section 5, briefly describes the conceptual-semantic analyses. Section 6, contains conclusion and future work.

II. INNOVATION OF THE IND

In 2005 the IND has chosen to fundamentally rebuild its organization. One of the steps toward a newborn IND was the redesign of all processes, including all supporting IT systems. This new system, called INDiGO, would be based on rule governance principles, separating procedural knowledge (workflow) from legal case content related knowledge. This principle became known as separating the know from the flow [7]. In order to create the actual IT solution, the IND selected several middle ware components, following the architectural principle of having one type of functionality within one middle ware component (one thing in a box). In order to support its knowledge management the IND has chosen to work with an inference engine that works on an explicit knowledge model derived from sources of law. This set up was also intended to enhance adaptivity and reduce maintenance efforts as a result of changing law and policies. Frequent changes of the knowledge model should not compromise the systems stability.

The INDiGO system is now operational for more than five years, and while the maintenance efforts are significantly less compared to the previous IT-systems, the effort required to implement changes is still substantially higher than expected when the solution was chosen.

The biggest issue is that legal experts and policy advisors lack the skills or will to read and validate the knowledge representation used by the inference engine in order to adapt changes when necessary. This is not a new problem. The Common Business Oriented Language (COBOL) was also once expected to be used by business people, rather than the very specialized computer programmers that actually created the systems code.

If systems are relatively stable the absence of direct insight of legal experts and policy advisors in the knowledge models would not hurt too much. But in volatile environments with constantly changing laws and policies, organizations will lose control of their IT systems unless they find ways to support the ‘translation’ of legal knowledge from the sources of law into a formal representation in their IT systems.

As a result of not using the knowledge models in the inference engine the user organization of the IND specifies changes without direct insight into the existing set of rules. Integrating the amended rules in the existing knowledge models is left to knowledge modelers. Although knowledge modelers within the IND have a history in the primary process of the organization, this leads to an unnecessary burden on the IT change process. Because requests for change in the inference engine are not fully specified. Interpretation errors made by knowledge workers, because of the lack of details in specifications, will not be detected until the start of the test phase. This leads to relatively expensive modifications.

The absence of unambiguous specifications also has consequences for the acceptance of changes by users. Having formal specifications is a condition for making a complete
test set. Test employees lacking full understanding of the required specifications will by definition carry out incomplete acceptance tests. Recognizing errors in the last stages of a release, or even in the initial period after delivery, leads to even higher repair cost than corrections made during the design and testing phase of a release.

The IND is seeking a formal method for coder independent and traceable specifications for the implementations of policy changes and changes in sources of law in information systems. These specifications should be available at the start of the process of changing knowledge models.

The authors have not found any existing solution for solving this issue.

III. LAWS AND REGULATIONS IN THE NETHERLANDS AND EUROPE

Why is scoping difficult? Laws are referring, explicitly and implicitly, to other laws and subordinate legislation. This creates a network of relationships. Determining relevant relationships can only be done on the basis of a given context. The scoping process should lead to a set of rules that can be traced back to sources of law and to a context description.

On October 15, 2014 the following rules were effective in the Netherlands, and thus potentially relevant for the IND:

1. 1.100 Dutch Laws;
2. 1.748 Dutch Orders in Counsel
3. 5.273 Dutch ministerial regulations
4. 679 Dutch international treaties
5. 2.796 European Laws
6. 3.164 European regulations
7. Dutch and European case law

The IND is responsible for the implementation of the Aliens Act and the Netherlands Nationality Act. This article focuses on the implementation of the Aliens Act. The Aliens Act consists of 173 articles, 288 lines, 28.766 words. Subordinate legislation includes:

• Order in Council: the Aliens Decree (322 articles, 5,642 lines, 56,331 words)
• Ministerial regulation: Aliens Regulations (167 articles, 3,983 lines, 27,213 words)
• Aliens Act Implementation Guidelines, part A to D (7,491 paragraphs, 13,935 lines, 153,068 words).

The challenge of the scoping process is to distil a relevant set of rules out of the entire collection of sources of law.

IV. A PROCEDURE FOR SCOPING

To select a workable set of rules that provides an adequate basis for developing specifications for a specified context, we used the following procedure:

1. Select an acting person (this can also be an organization). Describe the context of the acts of the agent
2. Choose a starting point for the analysis: a legal statement that contains a condition and is relevant in the chosen context.
3. Perform a linguistic analysis on a selected article.
4. Transform the text to the active voice, thus insuring there the subject of the sentence is an acting person.
5. Identify explicit references and terms that need a definition.
6. Select words or constituents that contain or might contain an implicit reference. Make these references explicit, or make an explicit decision that further analysis is not relevant in the chosen context.
7. Analyze all the selected words and constituents, starting with point 2 of this procedure.
8. The procedure ends when all relevant references analyzed. The decision to end the analysis is being made by a multidisciplinary team in which legal experts, policy advisors, and practitioners are represented.

The reader must be aware that the linguistic analyses (step 3 of the procedure) for the example cases described below have been conducted manually. The authors aim at supporting this step by automated devices in the near future and have worked on automated tools for the analysis of legal sources in the various previous projects [10][11]. The examples described below show that even a quite limited linguistic analyses is sufficient for our purpose.

A. Case: foreign students in the Netherlands

The procedure for scoping has been tested for the context of services provided by the IND (acting person) to foreign students studying or wanting to study in the Netherlands at a university.

B. Starting the analysis

Choosing a starting point for the analysis is step 2 of the scoping procedure. Two logical starting points for the analysis are:

1. A foreign student wants to come to the Netherlands. Conditions for admission to the Netherlands are stated in article 3, Aliens Act (Chapter 3: Entry).
2. A foreign student wants to reside in the Netherlands. Conditions for residence in the Netherlands are stated in article 8, Aliens Act (Chapter 3: Residence).

Both starting points lead to the same results within five iterations, because of implicit cross-references between article 3 and article 8, Aliens Act.

C. The analysis of an article

Step 3 of the procedure, the linguistic analysis of a selected article, is illustrated on the basis of Aliens Act article 16, paragraph 1, preamble and under b and article 4: 1 of the General Administrative Law.

Aliens Act article 16, paragraph 1, preamble and under b reads: “An application for a temporary residence permit can be rejected, if the alien does not possess a travel document.” This is a sentence in the passive form and it contains the following components, see Figure 1:

- can be rejected (verb)
- an application for the granting of a residence permit for a fixed period as referred to in article 14 (subject)
- if (conditional conjunction)
• the alien does not possess a travel document  
  (adverbial of condition; subordinate clause).

The subordinate clause of the sentence can be decomposed to:
• possess (verb)
• not (adverb of denial)
• the alien (subject)
• a valid travel document (direct object).

Because the sentence is in the passive voice there is no acting subject present.

In step 4 of the procedure the sentence is put in the active voice. Resulting in the question: who is the acting person that can reject an application for a residence permit.

The question is answered in step 5 of the procedure. The explicit reference to Aliens Act article 14, paragraph 1 reveals that it is Our Minister (the minister of Security and Justice) who is authorized to grant, reject or disregard an application for a residence permit.

After annotating the explicit references in the article, step 6 of the procedure commences: the search for implicit references. There is an implicit relationship between ‘an application for a temporary residence permit’ and ‘the alien’ that has to possess a travel document. Also the existing of ‘an application for a temporary residence permit’ implies that such an application can be submitted. To be able to make this implicit reference explicit, relevant rules are searched for in sources of law.

To make explicit what ‘the alien’ should do to proof he possesses a valid travel document, the requirements that enable ‘Our Minister’ to assess whether ‘the alien’ has a valid travel document are added, including references to the relevant sources of law, see Figure 2.

Step 7 of the procedure is the analyses of the articles of sources of law to which implicit or explicit references have been found.

The procedure continues with the analysis of article 4:1 of the General Administrative Law is shown, see Figure 3.

Article 4:1 of the General Administrative Law reads: “The application for taking a decision is submitted in writing to the administrative authority authorized to decide on the application, unless otherwise prescribed by law.” The sentence is in the passive form and contains the following components:
• is submitted (verb)
• the application for taking a decision (subject)
• unless (conditional conjunction)
• otherwise prescribed by law (adverbial of condition)
• to the administrative authority authorized to decide on the application (adverbial of place)
• in writing (adverbial of condition)

Repeating the scoping procedure leads to new implicit references, see Figure 4.

D. The size of the set of rules

The scoping procedure leaves 38 articles relevant for the entry and residence of foreign students in the Netherlands.
out of the original 173 articles of the Aliens Act. The set of rules also contains related sources of law, including rules from other legal domains.

The overall set of rules thus covers the following components (articles and paragraphs) of the regulations:

- 38 of the 173 articles of the Aliens Act
- 26 of the 322 in the Aliens Decree articles
- 12 of the 167 articles in the Aliens Regulations
- 280 of the 7,491 paragraphs of the Aliens Act Implementation Guidelines
- 1 article from a European law
- 13 articles of the General Administrative Law Act
- 2 articles from the Public Health Act
- 1 article of the Criminal Code
- 1 article from the Student Finance Act.

E. The value of a linguistic analysis

New in the proposed method is the explicit annotation of every legal rule that is needed for capturing the formal specifications of a service. Reasons for adding a legal rule to the relevant set of rules are recorded. Annotations for not adding a rule because of lack of relevance is only recorded when the analysts involved decide that this information might be relevant in the future. Every rule and every annotation in the model is explicitly related to a source of law.

The set of legal rules and the annotations made in establishing the relevant set of rules provides a basis for discussion and further analysis of the accuracy of the determined scope.

Adjusting the set of rules is possible at any time. Reapplying the procedure is necessary when changes are being made onto sources of law referred to in the relevant set of rules, or when any expert involved asks for adjustments.

The obtained legal rules must be converted to a formal (executable) specification. We use a specific application of CogNIAM (see [5][6]). This results in a formal model expressed in some relational algebra. It is therefore quite different from other specification languages that have a rule-like syntax, such as Prolog, most production rule syntaxes, business rules, etc. CogNIAM specifications can be used as formal specifications for IT systems and also provides a full set of test scenarios. With respect to the analytical approach used so far for the making of CogNIAM representations, the scoping procedure presented in this paper, yields the great advantage that the linguistic analysis is easy to learn. The constituents resulting from the linguistic analysis are much more structured than the original legal text, leading to a decrease of coder independency. Additionally the full traceability to legal text enables standardization of the translation of legal text in natural language to formal language. In the near future unambiguous automated methods for linguistic analysis on legal text will be available (see also [1][11][12]). For multiple languages including Dutch and English parsers that can perform this analysis automatically are available.

The intermediate step of linguistic analysis simplifies the often complex sentence structures in legal texts and thus provides for legal experts better understanding of the intended meaning of legislation, and makes anomalies that would otherwise remain hidden, visible.

V. Conceptual-Semantic Analysis

In [5][6] Van Engers and Nijssen describe the conceptual analysis of law. In order to allow the reader to understand how the models that are produced follow the steps described in the previous sections, we will briefly give the Hohfeldian analysis of one of the partial models (Aliens Act article 16, paragraph 1, preamble and under b). This model (see Figure 2) describes potential legal relations between ‘Our Minister’ and ‘the alien’. In Hohfeldian terms this model is read as follows: ‘Our Minister’ has a POWER to reject a request for a temporary residence permit, when ‘the alien’ cannot produce evidence of being in possession of a travel document.

When ‘Our Minister’ executes that POWER ‘the alien’ has a LIABILITY towards the rejection and a new jural relation between ‘Our Minister’ and ‘the alien’ comes into existence, namely the DUTY of ‘Our Minister’ to produce a decision (rejection) and a (CLAIM)RIGHT of ‘the alien’ on that decision. Once ‘Our Minister’ has fulfilled his DUTY the DUTY-(CLAIM)RIGHT relation terminates.

However, if ‘the alien’ can produce evidence of being in possession of a travel document, then ‘Our Minister’ has a DISABILITY and ‘the alien’ an IMMUNITY with respect to the rejection of the request for a temporary residence permit.

The model does not answer the question what happens then? We may not infer that by not having the POWER to reject the request will then be granted. So we have to look

![Figure 3](image3.png) Linguistic analysis of article 4:1 General Administrative Law.

![Figure 4](image4.png) Translation of article 4:1 General Administrative Law.
for other partial models that will provide us an answer to this question.

Also one may wonder how ‘Our Minister’ would know about the alien’s request unless it is submitted (by ‘the alien’)? Looking for an answer to this question we might find that this is arranged for in AWB 4.1 (see the corresponding model in Figure 3). According to this model ‘the alien’ got a POWER to submit a request for a temporary residence permit. If ‘the alien’ would execute this POWER ‘Our Minister’ would have the LIABILITY to decide upon the request thus creating a new jural relation between ‘Our Minister’ and ‘the alien’. This DUTY-(CLAIM)RIGHT relation would give a DUTY to ‘Our Minister’ to make a decision and a (CLAIM)RIGHT on that decision for ‘the alien’. This jural relation terminates once ‘Our Minister’ makes the decision.

If we would combine all these partial model we would find that the intended scenario would be as follows: When ‘the alien’ executes his POWER to submit a request for a temporary residence permit, ‘Our Minister’ has the LIABILITY to decide upon the request. This is the first POWER-LIABILITY relation. This creates a new legal relation between ‘the alien’ and ‘Our Minister’. Following from this relation ‘Our Minister’ got the DUTY to decide and ‘the alien’ the (CLAIM)RIGHT on the decision. When ‘Our Minister’ fulfills his DUTY the relation between ‘the alien’ and ‘Our Minister’ is terminated. The second POWER-LIABILITY relation is relevant for content of the decision. According to the combined model ‘Our Minister’ has POWER and ‘the alien’ has a LIABILITY towards a rejection if ‘the alien’ cannot produce evidence of being in possession of a travel document. ‘Our Minister’ then has the DUTY to produce a decision (rejection) and ‘the alien’ has a (CLAIM)RIGHT on that decision. Once ‘Our Minister’ has fulfilled his DUTY the DUTY-(CLAIM)RIGHT relation terminates. If ‘the alien’ produces evidence of being in possession of a travel document, then ‘Our Minister’ has a DISABILITY and ‘the alien’ has an IMMUNITY with respect to the rejection of the request for a temporary residence permit. What should happen in this case remains unclear.

The Hohfeldian analysis given above shows that the deeper analysis of the partial models further clarifies the meaning of sources of law in terms of (potential) legal consequences of agents’ acts. Furthermore the analysis helps to identify issues that need further clarification to fully cover the potential situations that may occur within a given context.

VI. CONCLUSION AND FUTURE WORK

The title of this paper was deliberately chosen. One could argue that the first steps towards formal analysis of law have been set long ago. For centuries researchers have been trying to formalize law, and certainly the last decades a lot of progress has been achieved. Legal information systems have become essential components in our modern society and without them efficient and effective application of law would be impossible. Despite all of this a complete method that translates law written in natural language into formal specifications is still lacking, although parts of such a method are already available.

The work we report on in this paper, fits within a series of studies, experiments and trials that are all aimed at developing a method for the formal analysis of law. In this paper, we addressed an issue - scoping - that has been neglected so far, but is an essential part of the method we are working on. In this sense we address the first steps in the translation towards a formal model of law.

By developing our approach and testing it on concrete cases we have learned a lot about how humans, including knowledge engineers and legal experts read and interpret sources of law and attribute meaning to them.

Following the scoping procedure explained in this paper, we have not only been able to identify relevant pieces of law, and thus helped to find a solution to our scoping issue, but while building the partial models and transforming them, we have been able to identify and clarify implicit relations between constituting parts of sources of law, terminological and conceptual unclarities and vagueness. This makes that our approach offers much more than merely filling the scoping gap in our method, as it contributes to improving the quality of the analysis of law and consequently helps public administrations such as the IND to improve the quality of their operation.

ACKNOWLEDGMENT

The research reported upon in this paper, wouldn’t have been possible without the support of the Dutch Immigration and Naturalization Service. Also the cooperation in the Blue Chamber has inspired us to continue to work on this intriguing topic and discuss our ideas even in late hours and at the cost of not paying attention to our spouses.

REFERENCES


