

A Model of Complexity for the Legal Domain [†]

Cornelis N.J. de Vey Mestdagh *

Centre for Law & ICT, University of Groningen, P.O. Box 716, 9700AS, the Netherlands

* Correspondence: c.n.j.de.vey.mestdagh@rug.nl; Tel.: +31-6-51-18-08-51

† Extended abstract. Presented at the IS4SI 2017 Summit DIGITALISATION FOR A SUSTAINABLE SOCIETY, Section Theoretical Information Studies, Gothenburg, Sweden, 12-16 June 2017.

Published data: 9 June 2017

The complexity of the universe can only be defined in terms of the complexity of the perceptual apparatus. The simpler the perceptual apparatus the simpler the universe. The most complex perceptual apparatus must conclude that it is alone in its universe.

Abstract: The concept of complexity has been neglected in the legal domain. Both as a qualitative concept that could be used to legally and politically analyze and criticize legal proceedings and as a quantitative concept that could be used to compare, rank, plan and optimize these proceedings. In science the opposite is true. Especially in the field of Algorithmic Information Theory (AIT) the concept of complexity has been scrutinized. In this paper we introduce a model of problem complexity in the legal domain. We use a formal model of legal knowledge to describe the parameters of the problem complexity of legal cases represented in this model.

Keywords: Complexity; Model of Complexity; Legal; Legal Knowledge; Model of Legal Knowledge; Inconsistency; Logic; Logical Variety; Labeled Logical Variety; Algorithmic Information Theory

1. Complexity in the legal domain

The concept of complexity is hardly developed in the legal domain. Most of the descriptions of concepts related to complexity in legal literature refer to vagueness (of the intension of concepts), open texture (of the extension of concepts), sophistication (the number of elements and relations) and multiplicity of norms (competing opinions) - in most cases even without explicit reference to the concept of complexity. Complexity arises in all these cases from the existence and competition of alternative perspectives on legal concepts and legal norms. A complex concept or norm from a scientific point of view is not necessarily a complex concept or norm from a legal point of view. If all parties involved agree, i.e. have or choose the same perspective/opinion - there is no legal complexity, i.e. there is no case/the case is solved. In science more exact definitions of complexity are common and applied. Complexity is associated with i.a. uncertainty, improbability and quantified information content. Despite this discrepancy between the legal domain and the domain of science, in the legal domain quantifying complexity serves the same interests as in other knowledge domains.

2. How to develop a model of complexity in the legal domain (methodology)

In this paper we will try to bridge the gap between the intuitive definitions of complexity in the legal domain and the more exact way of defining complexity in science. We will do that on the basis of a formal model of legal knowledge - the Logic of Reasonable Inferences (LRI) and its extensions - that we introduced before [4], that was implemented as the algorithm of the computer program Argumentator and that was empirically validated against a multitude of real life legal cases [5]. The 'complexities' of these legal cases proved to be adequately represented in the formal model. In earlier research we actually tested the formal model against 430 cases of which 45 were deemed more complex and 385 less complex by lawyers. A first result was that the algorithm (Argumentator) - when provided with the same case facts and legal knowledge - was able to solve 42 of the more complex cases and 383 of the 385 less complex cases in exactly the same way as the legal experts did (including the systematic mistakes made by these experts). A second result was that the algorithm - when provided with complete data and knowledge - improved the decisions in 30 (66%) of the 45 more complex cases (of which 20% full revisions) and in 104 (27%) of the 385 less complex cases (of which only 2% full revisions). This result confirms the relative complexity of the first 45 cases. The selection of these 45 cases thus provides us with the material from which criteria for the definition of complexity in this paper could be derived. These criteria are translated to quantitative statements about the formal representation of the cases. Further research will focus on the fine tuning of this

quantitative model by comparing its results with new empirical data (new cases and opinions of lawyers about the (subjective) complexity of cases). Finally, the ability of the fine-tuned model to predict complexity in new cases will be tested.

3. Models of complexity in science

The aim of this research is to develop a measure of complexity for formal representations of legal knowledge and their algorithmic implementations. We therefore studied Algorithmic Information Theory (AIT) to get acquainted with the theoretical and practical models of complexity developed in this domain of science. Our conclusion was that to be able to apply concepts such as Algorithmic Probability (esp. Solomonoff), Algorithmic Complexity (esp. Kolmogorov), Dual Complexity Measures [1], Axiomatic (esp. Blum) and Inductive Complexity, etc. we first had to develop a model of *problem complexity* for the legal domain. In further research we can address the measures of *solution complexity* developed in AIT.

4. A formal model of legal knowledge (reasonable inferences)

The first step in developing a model of complexity in the legal domain is to describe the formal characteristics of legal knowledge that are related to the essence of complexity in this domain, i.e. the competition of opinions. To formalize this characteristic of (legal) knowledge we developed the Logic of Reasonable Inferences [4]. The LRI is a logical variety that handles inconsistency by preserving inconsistent positions and their antecedents using as many independent predicate calculi as there are inconsistent positions [2,3]. In order to be able to make inferences about the relations between different positions (e.g. make local and temporal decisions), labels were added to the LRI. In [6] formulas and sets of formulas are named and characterized by labeling them in the form (A_i, H_i, P_i, C_i) . These labels are used to define and restrict different possible inference relations (Axioms A_i and Hypotheses H_i , i.e. labeled signed formulas and control labels) and to define and restrict the composition of consistent sets of formulas (Positions P_i and Contexts C_i). A set of formulas labeled P_i represents a position, i.e. a consistent set of formulas including all Axioms (e.g., a perspective on a world, without inferences about that world). A set of formulas labeled C_i represents a context (a maximal set of consistent formulas within the (sub)domain and their justifications, c.f. the world under consideration). Certain metacharacteristics of formulas and pairs of formulas were finally described by labels (e.g., metapredicates like Valid, Excludes, Prefer) describing some of their legal source characteristics and their legal relations which could be used to rank the different positions externally [6]. In [7] we showed that labels can be used formally to describe the ranking process of positions and contexts. In the next paragraph we will use the extended LRI to identify the quantitative parameters of complexity in the legal domain.

5. A formal model of the complexity of legal knowledge (parameters for a reasonable calculation of complexity)

The processing of legal knowledge takes place in five successive phases. Each phase is characterized by its own perspectives and associated parameters of complexity in terms of the formal model introduced above:

1. Constructing a number of sets n (the number of parties involved) of labeled formula $H_{i,l}$ representing the initial positions of each of the parties in a legal discourse, i.e. hypotheses $_i$ of parties $_l$ about the (alleged) facts and applicable norms in a legal case;
2. Determining the intersection between these sets $H_{i,l}$ which defines A_i representing the agreed case facts and norms and determining the union of all complements which defines H_i . (A_i, H_i) represents the initial case description;
3. Calculating all possible minimal consistent positions P_i that can be inferred from (A_i, H_i) applying a logic, e.g. the LRI, a logical variety that allows each position to be established by its own calculus;
4. Calculating all maximal consistent contexts (cf. possible consistent worlds) C_i on the basis of (A_i, H_i, P_i) ;
5. Making a ranking of these contexts on the basis of the application of the metanorms (decision criteria) included in them. A formal description and an example of this process are comprised in [7].

Each step in this process is characterized by its own *parameters of complexity*. In legal practice different procedures are used to determine and handle complexity in these different phases:

1. In the first phase a direct, static measure of complexity is commonly applied. The number of parties and the number of Hypotheses. This is a rough estimate of the number of different positions (interpretations, perspectives, interests);

2. In the second phase a direct, relative measure of complexity is commonly applied. The number of A_i and its relative size to H_i . The larger the relative size of A_i the less complex a case is considered to be, because there is supposed to be more consensus;
3. &4. In the third and fourth phases all positions P_i and contexts C_i are derived. Given the resulting set of labeled formula (A_i, H_i, P_i, C_i) representing the legal knowledge presented in a certain case, the problem complexity of this set can be defined as follows:
 - a. The subset A_i (agreed case facts and norms) is by definition included in each P_i and C_i so its inclusion as such is not a measure for complexity as it reflects absolute consent;
 - b. The elements of the subset H_i are by definition not included in each P_i and C_i so the relative size of the inclusion of its elements is a measure of complexity as it reflects relative consent. If there is more conformity there is less complexity;
 - c. The relative size of the fraction of subset A_i in (A_i, H_i) - relative to the fraction of A_i in other cases - is a measure of complexity as it reflects the size of shared (consented) knowledge;
 - d. The relative size of the fraction of subset H_i in (A_i, H_i) - relative to the fraction of H_i in other cases - is a measure of complexity as it reflects the size of disputed knowledge;
 - e. The relative size of the subset P_i (relative to P_i in other cases) is a measure of complexity as it reflects the number of different minimal positions that can be taken logically;
 - f. The relative size of the subset C_i (relative to C_i in other cases) is a measure of complexity as it reflects the number of different consistent contexts (possible decisions) that can be distinguished;
5. In the fifth phase ranking of the contexts takes place. The number of rankings depends on the inclusion of metanorms in the respective contexts. Metanorms that are agreed upon are part of A_i , metanorms that are not agreed upon are part of H_i . The process of applying the metanorms is fully recursive, since the objects of the metanorms are other (meta)norms, which are themselves also part of (A_i, H_i) . This means that the determination of the complexity of the application of the metanorms is included in the previous phases. In this phase only the resulting number of rankings is established and can be considered to be an independent measure of complexity.

6. Conclusion and further research

In this paper we have given a description of a formal representation of legal knowledge - the extended Logic of Reasonable Inferences (LRI) - and we have described the quantitative parameters of complexity for this model. The result of this we would like to call *Reasonable Complexity*, because it is based on the LRI and because it inherits its relative, perspective bound character. Complexity is specifically relative to the number of perspectives combined in the knowledge under consideration. Further research will focus on extending the model of complexity to solution complexity, using - amongst others - available algorithms (i.a. Argumentator, a computer program we developed to implement the LRI). It will also use the available dataset of 430 environmental law cases that have been described and analysed before and that have already been represented in Argumentator.

References

1. Burgin, M. *Super-Recursive Algorithms*; Springer Science & Business Media, 2005.
2. Burgin, M.; de Vey Mestdagh, C.N.J. The Representation of Inconsistent Knowledge in Advanced Knowledge Based Systems. In *Knowledge-Based and Intelligent Information and Engineering Systems*; Koenig, A. Dengel, A., Hinkelmann, K., Kise, K., Howlett, R.J., Jain, L.C., Eds.; Springer Verlag, 2011; Volume 2, pp. 524-537.
3. Burgin, M.; de Vey Mestdagh, C.N.J. Consistent structuring of inconsistent knowledge. *J. of Intelligent Information Systems*, Springer US, 2013, pp 1-24, ISSN: 0925-9902.
4. Vey Mestdagh, C.N.J. de; Verwaard, W.; Hoepman, J.H. The Logic of Reasonable Inferences. In *Legal Knowledge Based Systems, Model-based legal reasoning*; Breuker, J.A., Mulder, R.V. de, Hage, J.C., Eds.; Vermande, Lelystad, 1991; pp. 60-76.
5. Vey Mestdagh, C.N.J. de Legal Expert Systems. Experts or Expedients? In *The Law in the Information Society*, Proc. of the Fifth International Conference of the Italian National Research Council, Florence, Italy, 2-5 December 1998; Ciampi, C., E. Marinai, Eds.; Istituto per la documentazione giuridica del Centro Nazionale delle Ricerche, Firenze, 1998, 8 pp.
6. Vey Mestdagh, C.N.J. de; Hoepman, J.H. Inconsistent Knowledge as a Natural Phenomenon: The Ranking of Reasonable Inferences as a Computational Approach to Naturally Inconsistent (Legal) Theories. In *Information and Computation*; Dodig-Crnkovic, G., Burgin, M., Eds.; World Scientific, New Jersey, 2011; pp. 439-476.
7. Vey Mestdagh, C.N.J. de; Burgin, M. Reasoning and Decision Making in an Inconsistent World: Labeled Logical Varieties as a Tool for Inconsistency Robustness. In *Intelligent Decision Technologies*; Neves-Silva, R., Jain, L. C., Howlett, R. J., Eds.; Springer, 2015; Smart Innovation, Systems and Technologies, Volume 39, pp. 411-438.