

# Elliptical Lists in Legislative Texts

Réka Markovich  
Eötvös Loránd  
University/Budapest University  
of Technology and Economics  
markovich.reka@yahoo.com

Syi  
Budapest University of  
Technology and Economics  
i@syi.hu

Gábor Hamp  
Budapest University of  
Technology and Economics  
gabor.hamp@gmail.com

## ABSTRACT

Legal texts consist of hierarchically ordered and labeled (numbered) structural units (sections, subsections, paragraphs, etc.). Using the ordered layout and the labels the different parts of structural units can be easily localized and clearly referred. Nearly one-third of the structural units in the statutes we have examined are list items that can be considered as elliptical. In such cases the list items—each with unique identifying label (number)—are not complete propositions. We have trained the computer to recognize these lists and the different units and elements in them, and to create complete sentences from these. We will introduce some logical considerations that have to be reckoned with if we intend to use these complete sentences to create logical assignments to the legal regulation's content: we show how this technique influences the logical description of norms.

## Keywords

elliptical lists, text structure, logical analysis, deontic logic

## 1. INTRODUCTION

The decree regulating the legislative drafting in the Hungarian legal system (61/2009. (XII. 14.) decree of the Minister of Justice [1]) says: "Legal regulation drafts should be drawn according to the rules of Hungarian language, in a clear, comprehensive, consistent way" (§ 2). These requirements can be considered as fundamental: the legal system is expected to be clearly understandable by those who are subjects of the law. Also, a legal system containing contradictions makes law-abidingness impossible, therefore becomes senseless. Since consistency is a logical notion, the pursuit—which the authors are driven by—to provide a logical assignment of our legal rules seems to be legitimized by the legislation too. The issue of law's compliance with logic—as well as the efforts towards it—raises a great number of questions, details of which we would not go into at this point. The present paper aims to show that a proper interpretation of the results of natural language processing

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author.

Copyright is held by the owner/author(s).  
*ICAIL '15*, Jun 08-12, 2015, San Diego, CA, USA  
ACM 978-1-4503-3522-5/15/06.  
<http://dx.doi.org/10.1145/2746090.2746112>.

of legislative texts requires propositional and deontic logical tools.

## 2. STRUCTURE OF LEGAL NORMS

Texts of legal regulation consist of hierarchically ordered and labeled (numbered) structural units. Using the ordered layout and the labels, different parts of structural units can be easily localized and clearly referred to. The decree regulating legislative drafting specifies in details the structure of statutory instruments. It defines the following structural units: *Subparagraph, Paragraph, Subsection, Section, Subtitle, Chapter, Part, Book*. These are the basic categories of hierarchically ordered structural units.

The decree requires all components to be numbered using Arabic numerals or letters of the Latin alphabet, so the law does not contain unmarked sections (§ 37 (1)). Furthermore, it ordains that "if a section is divided into subsections every single sentence in the section entirely must belong to one of the subsections" (§ 44 (5)). From these it follows that subsections (or higher structural units) have to contain full sentences. This is not required from units on levels below subsections. In fact, in the text of laws it is very common that some ellipses occur in paragraphs and subparagraphs. In these cases the given unit (paragraph or subparagraph) in itself is not a complete sentence i.e. the sentences extend beyond the boundaries of the structural units. Such an elliptical item can only be interpreted as a complete proposition together with one of the preceding—and sometimes one of the following—units.

## 3. ELLIPSES

In linguistics, clauses from which one or more words are omitted are referred as 'ellipses'. Since in these cases the omission is governed by syntactic rules, the omitted words are clearly understood by the addressee. This—quite ordinary—technique is usually applied for avoiding word repetition or redundancy (e.g.: 'Anna goes to Amsterdam, Bob to Berlin.' 'I heard Anna's song, you heard Bob's.'). In theoretical syntax elliptical structures are important topics, and a number of its types is distinguished. The lists we have trained the computer to identify in legal texts have similarly elliptical structure. The elliptical quality comes from the written form.

## 4. LISTS

Jack Goody, a communication theorist of Toronto School, analyzing the nature of the written form of communication,

points out that writing allows or even forces us to use techniques that cannot be applied in oral communication [4]. Goody refers to these solutions based on knowledge representation of the written form as grapholinguistic techniques. Grapholinguistic techniques include lists, tables, matrices, and other text organizing methods. Although the primary content of a list or a table can be told, the written form as such can give additional information to the reader that a listener cannot receive through oral transmission. This additional content comes from the pictorial/visual character of the written text that can compose linearity into one (or more) other dimension(s).

What we find in subsections using paragraphs and subparagraphs are lists exploiting the expressive power of using an additional dimension. We will call these lists of elliptical items made up within subsections Goody-lists. Since Goody-lists can be considered as structured technique of ellipsis we can describe their structural properties. These lists always have a head unit introducing the list elements—at least two of them are needed in order for us to be able to speak about a list at all. Sometimes at the end of the list there can be found an additional unit that closes the list. To distinguish this latter from the head unit mentioned above we can call them 'closing head unit' and 'opening head unit', respectively. The opening head is always compulsory, the closing head is optional. Identifying these head units and the list elements, and concatenating them in the right order we get complete sentences. Subparagraphs in themselves are incomplete. The act of concatenation is a natural part of reading skills of both legislative and lay people; computers have to be trained to accomplish it.

## 5. WHAT DID WE DO?

To investigate Goody-lists we used six Hungarian acts to create our corpus. Processing the acts in a database in such a way that each structural unit goes separate row (separate record) we got a ten-thousand-record PostgreSQL database. Basically, we used a rule-based method (iterative refinement of regular expressions) to find and qualify all types of structural unit [3]. Sometimes SQL-queries based on regular expressions searching morphological features had to be complemented by structural information (in some cases the ranking order of the structural unit can be important). In the next step we identified the Goody-lists and their list elements—also with regular expressions. In the corpus we found 540 Goody-lists consisting of 3000 records (28%) in the corpus of 10,500 records in sum. That is, nearly one third of the texts of statutory instruments are settled in Goody-lists—i.e. elliptical texts. A quarter of the records consists of titles or marks of the structural units (25%). Concatenating the head units with the list elements we composed complete sentences (with SQL commands based on regular expressions), therefore we got 2200 semantically valid propositions—which we are going to call Goody-sentences—from the 3000 Goody-list related records. (The number of complete sentences that can be composed from a given list equals to the number of the elements of the list.)

## 6. LOGICAL CONSIDERATIONS

### 6.1 Part I—Propositional Logic

Since our work is driven by the pursuit to describe the legal regulation with logic, we must consider and explicate what

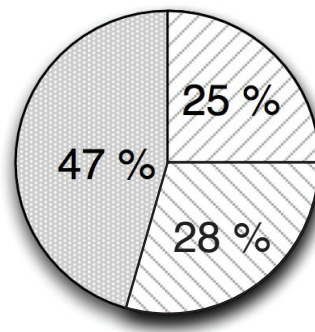


Figure 1: portion of Goody lists in laws

kind of phenomena we have to reckon with from the viewpoint of logic after carrying out the work described above. The decree regulating legislative drafting specifies four logical connections that can be between the list elements. It requires that the legislator make it clear in the case of a list what is necessary for setting the legal effect off: (1) all the (listed) conditions have to be fulfilled, (2) at least one of them has to be fulfilled, (3) exactly one of them has to be fulfilled, or (4) none of them can be fulfilled. The decree does not use their names from propositional logic but these are the following: conjunction ("and"), disjunction ("permissive or"), bisubstraction ("exclusive or") and conegation ("neither-nor"), respectively. The decree also mentions—as preferred solution—that the connection has to be indicated with a connective word at the end of the last but one list element. As it seems, according to this connective word we can easily identify the logical connection, therefore the logical connective too which we have to use between complete sentences we got as results of the process.

Let's see an example from Hungarian Civil Code [2]:

#### Section 3:5

*The instrument of constitution of a legal person shall expressly indicate the founders' intent to set up the legal person and shall contain:*

- a) the legal person's name;
- b) the legal person's registered office;
- c) the legal person's purpose or main activity;
- d) the names of the founders of the legal person, including their home address or registered office;
- e) the capital contributions prescribed, the value of such contributions, as well as how and when such assets are to be made available; and
- f) the legal person's chief executive officer.

At the end of paragraph e) we can see the 'and' so we know that with the Goody-sentences we have to reproduce the logical connection in the following way:

*The instrument of constitution of a legal person shall expressly indicate the founders' intent to set up the legal person and shall contain the legal person's name.*

AND

*The instrument of constitution of a legal person shall expressly indicate the founders' intent to set up the legal per-*

son and shall contain the legal person's registered office.

AND

The instrument of constitution of a legal person shall expressly indicate the founders' intent to set up the legal person and shall contain the legal person's purpose or main activity.

AND

The instrument of constitution of a legal person shall expressly indicate the founders' intent to set up the legal person and shall contain the names of the founders of the legal person, including their home address or registered office.

AND

The instrument of constitution of a legal person shall expressly indicate the founders' intent to set up the legal person and shall contain the capital contributions prescribed, the value of such contributions, as well as how and when such assets are to be made available.

AND

The instrument of constitution of a legal person shall expressly indicate the founders' intent to set up the legal person and shall contain the legal person's chief executive officer.

It seems if we intend to categorize our lists according to the logical connection settled between the list elements, it is enough to check the connective written at the end of the last but one list element. But in the case of connegation we have to check the head unit too to see what the real logical connection is. Let's see the following example:

Section 4:108

The presumption of paternity may not be contested if:

- a) descent results from a reproduction procedure, except if the mother's husband or partner has not given his prior consent for the procedure; or
- b) paternity was established by way of judicial process.

Here we can find an 'or' at the end of paragraph a) but we know from the expression 'may not...if' in the head unit that none of the listed conditions may be fulfilled for the presumption of paternity. Also, we have to take the following rule of propositional logic (one of De Morgan's law) into account when we would like to work with complete sentences instead of a list:

$$\neg(A \vee B) \Leftrightarrow \neg A \wedge \neg B$$

On the left side of the equivalence we can see what is described by the list, on the right side the conjunction of the two negations, i.e. what we have to get with the complete Goody-sentences. It means that the separate sentences containing the negation have to be connected by an 'AND'. But that is not all. Something else also changes. This is the "nature" of the conditions. To explicate this change, let's see some further considerations.

When the decree mentions the logical connections it regards the logical connective between the list elements only but one should also consider the logical relation between the head unit and the elements, especially—as we will see—that this relation changes if we divide the list in order to create complete sentences. It is useful to distinguish the two types of connection, so we refer to the logical relation between list elements as an internal connection, while the relation be-

tween the whole group of the list items and the head unit will be called external connection. Regarding the external connection—as the citation from the decree on legal drafting shows—these lists often used for giving the conditions for setting off the legal effect. If the legislator provides an exhaustive list of the conditions in one list, we can consider this list containing the necessary and sufficient conditions. How should we understand these features? This is shown by the formula below: if the legislator provides conditions a definition-like way, then (from a logical point of view) we should consider it as a biconditional which covers necessity and the sufficiency in the following way:

$$H \Leftrightarrow (A \vee B) \Leftrightarrow (A \rightarrow H) \wedge (B \rightarrow H) \wedge ((A \vee B) \leftarrow H)$$

where 'H' stands for head unit, 'A' and 'B' are representing list elements. The right arrow is a conditional that makes each list element to be sufficient condition, i.e. fulfilling any of them sets off the legal effect indicated in the head unit, while with the left arrow—which stands for the propositional logical connective called retroconditional—we show that the disjunction of the elements is a necessary condition since (at least) one of the list elements has to be fulfilled. This means that the isolated Goody-sentences will not be enough to express the real content, we also have to consider the logical connection described above.

## 6.2 Part II—Deontic Logic

Let's see a puzzling example that shows that besides propositional logical connectives something else is needed to properly reconstruct the logical connection between the Goody-sentences:

Section 2:28

(1) Placement under guardianship or conservatorship may be requested by:

- a) the spouse, domestic partner, relative in direct line, or sibling living with the person of legal age;
- b) the minor's legal representative;
- c) the guardian authority; and
- d) the public prosecutor.

Section 2:30

(2) Termination of guardianship or conservatorship may be requested by:

- a) the person in custody;
- b) the spouse, domestic partner, relative in direct line, or sibling living with the person in custody;
- c) the conservator;
- d) the guardian authority; or
- e) the public prosecutor.

It can be seen that the structure of these two sections is the same. But the connective word at the (1) c) is 'and', at (2) d) is 'or'. Should these lists be interpreted differently because of the difference of the connectives despite their structural similarity? Then some discrepancy arises drafting the two lists in this way though. This—presumably—unintended discrepancy comes from a deontic logical paradox called free choice permission. [6] [7] According to it, the permission of the disjunction of two things is equivalent with the conjunction of these permitted things. That is, the proposition

'Fred or Bill may come' is equivalent with the proposition 'Fred may come and Bill may come' [5]. With formula:

$$P(A \vee B) \Leftrightarrow PA \wedge PB$$

The inconsequence of the connective word use in the sections above of Hungarian Civil Code is probably the result of this weird feature of presence of 'and' and 'or' at the same time. This phenomenon highlights also that when we get complete Goody-sentences from the list elements we have to be careful with using a logical connective between the complete sentences: if the head unit contains a permission then the 'or' between the list elements becomes an 'and' between the complete made up sentences. For example from the Section 2:30 the adequate sentence list is the following.

*Termination of guardianship or conservatorship may be requested by the person in custody.*

AND

*Termination of guardianship or conservatorship may be requested by the spouse, domestic partner, relative in direct line, or sibling living with the person in custody.*

AND

*Termination of guardianship or conservatorship may be requested by the conservator.*

AND

*Termination of guardianship or conservatorship may be requested by the guardian authority.*

AND

*Termination of guardianship or conservatorship may be requested by the public prosecutor.*

## 7. SUMMARY

As it was showed above, we can automatize the process of creating complete sentences from elliptical lists that can be very useful when we would like to have automatized process of referring. We also showed that this transformation

can influence an adequate logical description of the transformed paragraphs' content. To handle this influence we need to consider propositional logic and deontic logic too. Moreover, sometimes we have to know these in order to correctly interpret some natural language utterances, even in a natural language register so formal as law is. That is—as our examples showed—in order to that the computer properly understand the legal texts it is necessary to use logical apparatus.

## 8. ACKNOWLEDGMENT

Support provided by the research projects K83887 and K109456 of the Hungarian Scientific Research Fund is gratefully acknowledged.

## 9. REFERENCES

- [1] 61/2009. (XII. 14.) decree of the Minister of Justice on regulating the legislative drafting (A jogszabályszerkesztésről szóló 61/2009. (XII. 14.) IRM rendelet)
- [2] *Hungarian Civil Code, Act V. of 2013* (2013. évi V. törvény a Polgári törvénykönyvről)
- [3] Friedl, J.E.F. *Mastering Regular Expressions*, O'Reilly Media (2006)
- [4] Goody, J.: Language and Writing. In: *The Interface between the Written and the Oral*, Cambridge University Press (1987) 258-289
- [5] Jennings, R.E.: Can There Be a Natural Deontic Logic? *Synthese*, (1985), 65, 2, 257-273.
- [6] Kamp, H.: Free Choice Permission. *Proceedings of the Aristotelian Society*, New Series, Vol. 74 (1973-1974), (1974) 57-74
- [7] Makinson, D.: On a Fundamental Problem of Deontic Logic. In: P. McNamara, H. Prakken (eds.). *Norms, Logics and Information Systems. New Studies in Deontic Logic and Computer Science*, IOS Press (1999) 29-53