Introduction to Cross-reference Clauses in Legal Texts

F. STUDNICKI - A. LACHWA - J. FALL - E. STABRAWA*

1. FROM THE AUTHORS

This paper reports an introductory part of our monograph, devoted to automatic interpretation of the cross-references occurring in source legal texts, published in 1991 by the Jagiellonian University Press (in Polish). The monograph was a result of the research project conducted by the team of the authors of this paper for several years. The author appearing as the first one was the manager of the whole group and did the greater part of that work.

2. PERSPECTIVE AND AIM OF WORK

Although cross-references to contents located in other places of the same or other texts appear practically in all the source legal texts, the problems connected with them are mostly treated by legal theory only marginally. They are usually touched upon only when general theoretical questions are discussed, in particular when the problems of the technique of legislation are dealt with. Apart from that, information on the problems mentioned above – usually very short and fragmentary – can be found in handbooks of theory of law and in works devoted to the most general problems of dogmatics. Only two works devoted to a detailed discussion of cross-references are known to the authors of this book. They are: the large monograph of H.-U. Karpen (1970) and a work of J. Wróblewski (1964). These works, however, discuss the problems connected with cross-references from the point of view different from ours. In particular, they

* Jagiellonian University, Kraków, Poland.
1 Cf. F. Studnicki et al. (1990).
are par excellence legal works showing the problem from the point of view of the theory of law and legislative technique.

A different point of view and – consequently – different perspective have been chosen by the authors of this work. Our attention has been focused above all on such linguistic properties of legal cross-references as can be helpful in organizing of the processes of automated retrieval of information inherent in legal texts. The already existing systems of legal information retrieval are incapable of searching out information inherent in fragments of texts referred to by the cross-reference clauses that occur in legal texts. Identification of this information is left for the users of the above systems. Thus, the user of the retrieval system is supposed to perform consequently two operations. The first consists in making use of the system in order to retrieve the fragment containing a reference clause, while the second – in searching out the fragments it refers to. Such course of affairs leads up to additional and unnecessary burdening imposed upon the user. In our work we aim at relieving him of that burden.

3. CROSS-REFERENCE CLAUSES

In source legal texts, such as texts of statutes, decrees, regulations etc., there often appear the so-called cross-reference clauses. By this term we mean here linguistic clauses referring to contents expressed in other fragments of the same text or in other source legal texts in such a way that interpretation of the fragments containing these clauses is not possible without taking those contents into account.¹

¹ J. Wróblewski (1964), p. 3, understands references broader, which is evident from his definition of the «referring clause sensu largo». The definition sounds as follows «Let us call ‘referring regulations sensu largo’ all the legal regulations in force, which ‘refer’ to other regulations belonging to the system of law or to rules non-belonging to the system of law». Both J. Wróblewski (1964) and H.-U. Karpen (1970) differentiate references in the large sense from references in the narrow sense. J. Wróblewski distinguishes in particular «(intra-) system referring regulations», i.e., those which refer other legal regulations, from extra-system referring regulations. Cf. also W. Lang, J. Wróblewski & S. Zawadzki (1986), p. 350. The differentiation between cross-references and other references/quotations to some rules or principles is presented in H.-U. Karpen (1970), p. 21 and further. In the present study we focus exclusively on – following the above terminology – «(intra-) system referring regulations». Hence, we are not occupied also with the so-called assumptive/implicative references, for instance, with the phenomenon consisting in this that regulations from the general part of some code, and among them also the defining regulations of that part, find their applications in situations regulated in the detailed part of the code. Cf. H.-U. Karpen (1970), p. 34. The authors quoted above single out references of two levels. In particular, J. Wróblewski (1964) considers a regulation which «regulates a scope of being in force or
4. DISTINCTION FROM OTHER SIMILAR CONSTRUCTIONS OF LEGAL LANGUAGE

It is not so that in a source legal text every employment of a clause containing a denomination of the other text or of its part may be treated as an employment of a cross-reference clause. For instance, we will not use the term «cross-reference clause» for characteristic clauses appearing in the so-called validating regulations, i.e., in regulations that enforce some new rules into the legal order. We will treat in the same way characteristic clauses appearing in regulations which repeal the legal force of some texts or some parts of them. The reason is that the regulations of these two kinds may be properly interpreted without taking the contents expressed by texts or text parts indicated by those regulations into account.

Our demands addressed towards cross-reference clauses are not fulfilled by expressions called in jurisprudence the general clauses. We mean here the expressions referring to such sources of norms and/or validations that are non-identical with the law, for instance, «the principles of community life», «the good will requirements», «the principles of the honest (trade) turnover» etc. Those expressions do not refer to contents comprised in source legal texts. Therefore, study of them is beyond the scope of our interest.

Proper performance of cross-reference clauses is promoted by a peculiar structure of source legal texts, which may be described as mosaic (cf. further in point 7).

5. STRUCTURE OF CLAUSE-REFERENCE CLAUSES

Cross-reference clauses appear in source legal texts in a great variety of forms. Some of them have quite complex formal and semantic structures.
Our work is devoted to the study of those structures. Nevertheless, our concern has been directed first of all towards reconstruction of those of their properties that may be employed in operations of automatic resolving of cross-reference clauses. By the term «resolving» we mean here operations aiming at retrieval of such fragments of proper texts that comprise the contents mentioned above, i.e., the contents to which the clauses of our interest refer.

6. Direct and indirect information retrieval.

Our study is directed mainly to applying of its results in the systems of the so-called indirect legal information retrieval. The point is in the retrieving whose result consists in providing a user of a given system with (or, at least, indicating to him) such source legal texts or fragments of such texts that express contents that are relevant to him. In particular, the retrieved text fragments should serve the user as a basis to be applied in appropriate interpretation procedures resulting in reaching the answer to legal problems he is curious about in a given case.

The performance of such systems is founded on the assumption that all the contents expressed by texts are located in these texts in such a way that those contents may be rendered accessible to the users of the retrieval system by providing them with proper text fragments. The above assumption, which we call 'the content localization assumption', is one of the so-called idealizing assumptions and, similarly to others of that kind, it somehow simplifies the reality.

This simplification consists in taking no account of closer and further contexts influencing the form taken by the content in individual text fragments. It is unpropitious to the use of the so-called system (legal) interpretation by lawyers applying the methods of indirect information retrieval. The term «system (legal) interpretation» means here such (legal) interpretation that takes into account various content connections relating the currently interpreted regulations to other regulations of the same legal order.

To some extent all the methods of indirect legal information retrieval suffer from the same defect. However, it cannot result in their total rejection, for the information losses caused by it are excessively rewarded by the advantages following the use of those methods. Nonetheless the existence of the losses binds us to apply some remedial measures. One of such measures consists in providing the retrieval systems with facilities enabling them to reveal to the user at least some of the contents connections men-
tioned above. The measures of revealing of connections between different regulations that have been introduced by a legislator with the use of cross-reference clauses, i.e., the measures of automatic identification and resolving of such clauses may also belong there.

Some of the operations taking place during indirect legal information retrieval, in particular the operations consisting in identification of such text fragments that fulfill demands of system user’s retrieval criteria, may also be applied in some of the direct retrieval systems, i.e., in systems that give (direct) answers to the legal questions asked by the system users. This regards such of those systems whose data bases are provided with reproductions of source legal texts. In such systems retrieving of appropriate text fragments composes the introductory stage, followed by the stage in which the answer is generated. Text fragments identified in the first stage establish one of the sources of information employed then by the system in its second stage. Therefore, the usefulness of the results of our work, may spread out also on the functions of the direct information retrieval systems.

7. AUTHENTIC TEXT DIVISION AND INFORMATION RETRIEVAL

The text division introduced by a legislator (henceforth: the authentic division), provides source legal texts with a structure that may be described as a modular or mosaic one. A peculiarity of that structure consists in the fact that legal texts are divided into fragments that may be distinguished with the use of formal measures. Every such fragment develops an individual, though connected with the others, thread of the content. The fragments are in nearly all cases supplied with denominations taking the form of numbers and/or titles. Otherwise the fragments may be distinguished in a text by referring to punctuation or/and to their typographic form. The division of a text into fragments established by the legislator (authentic division) facilitates identification of these parts of the text, in which the contents indicated by cross-reference clauses are located.

It is obvious that a content located in a given text fragment is at the same time located in all the fragments of which that fragment is a part. For

---

10 About the system of headings in source legal texts cf. F. Studnicki (1985), passim. This author understands the system of headings as a set of headings appearing in a given text, together with relations established on that set, like, for instance, the relation of precedence or the relation of domination.
instance, the content located in a given article of a statute is at the same time located in the proper chapter, and also in the proper section; if fragments composing chapters and sections have been distinguished in the statute. The comprised in a cross-reference clause information about where the content to which that clause refers is located may be more or less detailed. In other words, the cross-reference clause «encircles» the above content more or less «tightly». The lower position in the hierarchy of fragments of a given text the fragment indicated in the clause takes, the tighter the encircling is. For example, the encircling obtained by having mentioned a particular article in a given clause is tighter than the encircling that would be obtained by mentioning the chapter of which the article is a part.

If we assumed that the result of solving of a cross-reference clause was always to consist in providing a user with such a text fragment that was mentioned in the clause – irrespective of how tight the encircling composed by that clause was, we would have to put up with the fact that in many cases the user would receive very large fragments, e.g., complete volumes, sections or chapters. Such state of affairs would not be profitable to the user, for he would then be obtaining the information of his concern together with a great amount of irrelevant information, i.e., the result with a high level of noise. Therefore, we have to assume that the function of cross-reference clauses that create loose encirclings does not consist in direct indicating of proper contents to the user, but only in identifying of such regions of the database in which subsequent retrieval may take place.

Clearly, it is reasonable to undertake such retrieving only in cases when the information reached previously does not satisfy the actual information demands of the user, for instance, when he is not satisfied with the information that the contents of his concern are located in a given volume of some code. When this is the case – and it may occur most often – the user will usually undertake subsequent retrieving, in which he will use the information acquired previously. It will allow him to limit his retrieval only to that region of the database that has been distinguished by the resolved cross-reference clause.

8. TEXT FRAGMENTS AND CONTENTS LOCATED IN THEM

In the previous discussion we used the term «a text fragment» without a clear definition of its meaning. Now we explain that by the above term we understand a part of a text that may be distinguished with the use of formal criteria, not smaller than a grammatical sentence, and considered with taking its content properties into account. Text fragments, with excep-
tion of fragments of the lowest level, consist of fragments of lower level. For instance, the fragment being a chapter will usually hold in itself fragments being paragraphs, they in turn will contain in themselves fragments being sentences and containing in themselves (ex definitione) no fragments.

This state of affairs gives the text the character of a hierarchic structure of fragments. Fragments of different levels may be distinguished in that structure and therefore it is possible to talk about subordination of some fragments to other fragments. The relation of superiority of level of fragment \( F \) to fragment \( F' \) occurring between two fragments \( F \) and \( F' \) is accompanied with the relation of subordination of fragment \( F' \) to fragment \( F \) if and only if fragment \( F' \) is a proper fragment of fragment \( F \). For example, article \( A \) is subordinated to chapter \( R \) if and only if the former composes a proper part of the latter.

Dividing of a text into fragments by its producer is motivated by his aiming at such content distribution inside the produced text in which the contents bound with each other and composing certain threads of content are located close to one another and opposed to the contents grouped in a similar way in the following fragments. It is obvious that where one fragment ends, several following fragments may begin. However, each of such fragments must be of another level. For instance, in the place, where a given article ends, the next article, the first of the paragraphs composing this article, the first of the passages of this paragraph and the first sentence of this passage may begin. It is similar, mutatis mutandis, to the situation with the place in which a given fragment ends. In particular, in such places also other fragments may end. Yet, each of these fragments must be of a different level. The above-mentioned opposition between the contents located in a given text fragment and the contents located in the fragments following it grows according to the level of that of fragments directly following the given fragment to which the highest level may be assigned. In other words: the higher the difference between maximum levels of adjoining fragments is, the stronger their contents opposition should be.

The opposition discussed here is usually stressed by the producer of a text with the use some graphic endeavors that result in clear separation of adjoining fragments. These endeavors consist in applying of particular punctuation, in use of indentation and/or in providing fragments of some levels with special denominations.

We will speak of the location of a definite content in a given sentence when the content in question cannot be reproduced without taking that sentence into account. A definite content may be located in one sentence or in many sentences. We will say that a given content is located in many
sentences when its reproducing from the text requires taking all these sentences into account.

The sentences composing the locations in the above sense of the contents indicated by a given cross reference clause are called nests of these contents. If a given content may be reproduced while taking only one sentence into account, we call this sentence a self-contained nest of that content. On the other hand, when its reproduction requires considering of two or more sentences (directly adjoining or separated ones), the set of these sentences will be called a colony of the nests of that content. It is obvious that none of the sentences composing together a given colony can be a self-contained nest of that content.

The contents expressed by source legal texts may be comprised not only in sentences appearing in the so-called main line of a text, but also in inscriptions composing titles of some text fragments. In the texts constituting the empirical material of our research we have not found any titles that would have the form of a sentence (except one case). Neither have we found such clauses that would refer to contents located wholly in a certain title. Hence, we can decide that the inscriptions composing titles will be counted among the nets of some contents. However, we cannot count them among the nets of some contents. However, we cannot count them among the self-contained nets. The above formulation leads to the conclusion that contents comprised in titles may only supplement or modify the content located in sentences of the main line of a source legal text. Nonetheless, the contents comprised in the titles should not be left out of account in operations of resolving of cross-reference clauses. It is possible to avoid omitting them. Namely, by applying an operation in which the system would retrieve not only the text fragments expressing the contents to which a given cross-reference clause recalls but also the titles of the fragments (if the fragments are somehow headed), and all (or at least some) of the titles of the fragments superior to the retrieved fragments.

Parts of a text equivalent to text fragments but considered only by taking their formal properties into account are called text segments. Therefore, it is evident that the same part of a text may be considered as a text fragment or as its segment.

As it is already known, in source legal texts there usually appears a text division coming from the legislator and called in this work the authentic division. Parts of a text distinguished by that division fulfill our requirements.

\[11\] On the notion of the main line and the distinction between the main line and other structural lines of a text (nonidentical with its graphic lines) cf. F. Studnicki (1985), p. 550.
set for the fragments. The existence of the authentic division is no obstacle to another text division to be used when needed for retrieval purposes. However, parts of the text distinguished by the other divisions must fulfill the demands mentioned above.

The authentic division looks variously in various source legal texts, and even in various parts of the same text. For instance, in Polish civil code the fragments of a directly higher level for fragments being sentences are usually passages. Further there appear paragraphs, articles, chapters, divisions, titles and volumes as the fragments of higher and higher levels. In some parts of the aforementioned code this division is simplified, which means that some of its constituents do not appear in those parts. For example, in some parts of Polish civil code the divisions are divided directly into articles, with omission of chapters. It is similar with some other source legal texts.

9. ILCLOCUTIONARY STATUS OF CROSS-REFERENCE CLAUSES

Cross-reference clauses appearing in source legal texts have their particular and significantly complex illocutionary status. Its specific character comes not only from the fact that those clauses are constituents of a specific kind of texts, namely of normative texts. The other factor here is that the illocutionary status characteristic to all the expressions comprised in normative texts is complemented by a specific illocutionary feature that may by rested in all the referring expressions irrelevant of the kind of text in which they appear. Therefore, the illocutionary status of cross-reference clauses occurring in source legal texts constitutes a combination of the two kinds of illocution mentioned above.

In considerations whose results are to be used first of all for organization of the processes of automated legal information retrieval, the normative tinge of the discussed clauses does not play a significant role. Mentioning that tinge we mean that use of such clauses in a source legal text usually imposes certain duties on definite subjects, first of all – the duty of respecting, in appropriate situations (described in hypotheses of regulations with cross-reference clauses), of the contents indicated by those clauses. However, the existence of that duty is completely irrelevant when a «retrieval» point of view is taken into account. For referring to that duty does not in any

12 In some regulations of the code the so-called points are the units of direct subdivision of passages or other authentic fragments. The points are marked with Arabic numerals or distinguished by typographic measures. The problem of points is left out of the scope of our work because it shows no peculiarities.
sense enriche the criteria which make identification of the contents under consideration in appropriate text(s) possible. Therefore, the problem of normative character of cross-reference clauses appearing in source legal texts, i.e., the fact that their functioning consists among others in imposing of certain duties on definite subjects, may in our considerations be almost completely disregarded.

10. **Objects indicated by cross-reference clauses**

We regard this function of cross-reference clauses which consists in indicating of certain objects as their basic function. The question of what objects are indicated by those clauses arises. In our introductory remarks concerning the definition of a cross-reference clause we assumed that these objects are certain contents, namely the contents to which a given clause refers. However, is not this formulation arbitrary? In the clauses under consideration no contents are mentioned, but usually only some parts of a text or texts, e.g., articles and/or paragraphs denominated with the use of numbers. From this fact one may conclude that the objects indicated by cross-reference clauses are not these or those contents but definite text segments, or definite text fragments, in particular the segments (fragments) located in a text under positions mentioned in the above clauses. One may also state that it would be justified to see the objects indicated by cross-reference clauses as these or those contents if and only if the contents were directly indicated by the clauses, e.g., by their verbalization.

The standpoint at which one sees objects indicated by cross-reference clauses as being text segments or fragments and not certain contents is called formal standpoint13. The arguments against the above standpoint are in our opinion as follows. Stressing of the fact that cross-reference clauses do not usually direct to definite contents but only mention – with the use of different measures – positions taken in a text by some of its fragments, does not take one issue into account. Namely, the above fragments are distinguished by a clause not because of their formal features but exclusively because of a content expressed by them. Hence, it should be assumed that in cross-reference clauses quoting of positions of those fragments is only a medium whose sense consists in indicating of proper contents. Thus, while considering the functions fulfilled by cross-reference clauses we should see the difference between objects indicated by those clauses, i.e., certain contents – if our way of thinking is to be followed – and the manner in

which such indication is made. The same objects (in our case: the same contents) may be indicated with the use of different ways of indicating. In particular, we distinguish seven types of the elementary indications (internal, external, deictic, semantic, semaphoric, recursive and reflective) and various compound indications.

11. FUNCTION OF CROSS-REFERENCE CLAUSES

The use of cross references is one of the traditional measures in legislation technique. The aim of that measure is first of all to eliminate the needless redundancy from the corpus of source legal texts. By this term we mean such redundancy that may be eliminated from texts without the fear of arising disturbances in their proper reception, i.e., without an excessive increase of the level of noise during their reception. That measure must be used with moderation. Its abuse may lead to a situation in which apart from - or even instead of - the needless redundancy the text is deprived of a «needed» redundancy. For certain limited level of redundancy plays a useful role in a text in the sense of providing it with a sufficient noise-resistance, or - in other words - in the sense of being easier to comprehend.

The above function of cross-references is important from the point of view of communication. However the discussed measure of legislative technique plays also other roles. We will mention here only the increase of the content cohesion of individual source legal texts and the intensification of content connections between particular texts of a given legal order.

12. CONTENT CONNECTIONS IN TEXTS

The above-mentioned connections may have various character and various intensity. As regards the connections within the limits of one and the same text, it is normal that the semantic contents of fragments occurring in succession from longer or shorter content sequences. In these sequences information located in earlier text fragments is joined by information located in further fragments that supplements or modifies the former. However, Nonetheless, the use of cross-references in source legal texts is often criticized because of some disadvantages of that measure. In particular, it was stressed that cross-references make a text less clear, and at the same time less comprehensible, especially for non-lawyers. The disadvantage may be decreased while using legal information retrieval systems enriched with software facilities able to resolve cross-references in an automated way. On advantages and disadvantages of the use of cross-references cf. H.-U. Karpen (1970), p. 11-18. On an influence of cross-references on text cohesion cf. E. Stabrawa (1986), p. 103-106.
the relationship of supplementing and modifying may also exist between contents of fragments that do not adjoin in a text, or even between contents of fragments from different source legal texts.\footnote{On the content and formal connections in a text cf. A. Łachwa (1988), p. 15-22.}

In both cases the connections may lie in this, e.g., that principle $Z$ expressed in fragment $F$ and referring to a situation from class $K$ which has been characterized in that fragment, is expanded by another fragment $F'$ into situations non belonging to that class.

The connections under discussion may also consist, among others, in this that from a definite class $K$ of situations (characterized in fragment $F$) to which a given principle $Z$ applies, a certain subclass $K'$ is eliminated according to fragment $F'$. Subclass $K'$ is characterized in fragment $F$ but it comprises situations in which principle $Z'$ (expressed in fragment $F'$) should be applied instead of principle $Z$.

In all the above cases we will deal with strong content connections which take place between fragments $F$ and $F'$. One of the measures for obtaining such connections in source legal texts is the use of cross-reference clauses.

The structure of connections obtained by employing of such clauses needs more detailed consideration. We will begin the discussion with an analysis of the simplest cases, namely, the cases in which a given fragment of a source legal text, e.g., a definite article of a statute contains a cross-reference clause to the content of another fragment of the same text, for example, to another article of the same statute.

Let us use the following example (the original Polish version from the Penal Code):  

1) «Art. 11. § 1. He who intends with malice to commit an unlawful act and with his behavior aims directly at committing it but fails in it, is responsible for the attempt.  
§ 2. It is also an attempt when the perpetrator is not conscious that the act is impossible because of the lack of an object which might be used as a goal of the offence or because of having used a measure that has been improper for reaching of the intended result.  
Art. 12. § 1. The court of justice metes out punishment for an attempt within the limits of statutory punishability for a given offence.  
§ 2. In the case described in art. 11 § 2 the court of justice may employ extraordinary mitigation of punishment or even renounce from its infliction.»

There is no doubt that there is a content connection between the fragment of the penal code denominated as art. 12 § 2 and the fragment of
the same code denominated as art. 11 § 2. The connection consists in the fact that the former, i.e., art. 12 § 2 contains a content gap and at the same time it informs that this gap may be filled by the content located in the latter, i.e., in art. 11 § 2, what, in fact, takes place. The roles played in this connection by the two participating text fragments are by no means equivalent. It is art. 12 § 2 that cannot be sufficiently comprehended without taking the content of art. 11 § 2 into account. By insufficiency we mean here that there exist such questions connected with the matter that is spoken of in art. 12 § 2 that cannot be answered without taking the content of art. 11 § 2 into account. In particular, we can answer the question whether art. 12 § 2 entitles the court of justice to employ extraordinary mitigation of punishment or even to renounce from its infliction without taking the content of art. 11 § 2 into account. However, we cannot answer the question in what situations the court of justice has the above competence. Some additional information is given by the direct context of the article 12 § 2, namely that this regulation is located in a direct neighborhood of the regulations about the attempt. It enables us to state with great probability that the court of justice is entitled to that competence in certain cases of attempt. However, this information is not sufficient to state precisely what cases of attempt are concerned. This information is only given by the content of art. 11 § 2, that is, the content indicated by the cross-reference clause comprised in article 12 § 2.

As it was mentioned earlier, the content dependence understood in the above way may appear not only between text fragments connected with each other by the use of a cross reference clause in one (or in both) of them. For instance, the standard case of its appearance is the connection between a text fragment expressing a general rule (lex generalis) and a fragment expressing a rule that constituting a particular case of the general rule (lex specialis). In such cases the fragment that expresses a general rule is in content dependence upon the fragment or fragments expressing particular rules. For example, article 199 § 1 of the Polish penal code reads as follows:

2) «Art. 199 § 1. Whoever seizes common property fraudulently is liable to penalty of 6 months to 5 years of deprivation of liberty».

However, the code contains also several regulations that modify this rule in particular situations, e.g., in the situation when the perpetrator administered the appropriated common property in connection with his post or a function fulfilled by him (art. 200 § 1), in the situation when property of great value is concerned (art. 201 § 1), etc. The dependence of the fragment denominated as art. 199 § 1 from the fragment denominated
as art. 200 § 1, or from the fragment denominated as art. 201 § 1 consists
in this that there exist such questions connected with the matter mentioned
in art. 199 § 1 that cannot be answered without taking the content of art.
200 § 1 and art. 201 § 1 into account. (In order to simplify the description
we do not talk about other articles that might also be engaged here).
Among such questions there is one about punishment which the court of
justice should mete out in a given case of fraudulent seizure of common
property. The fact that an act of a certain person fulfills the characteristics
mentioned in article 199 § 1 does not settle the question. For it is not
excluded that the act under discussion fulfills at the same time the
characteristics discussed in one of the modifying regulations, e.g., in art. 200
§ 1 or in art. 201 § 1 (or even the characteristics of both regulations). In the
latter cases the answer to the question under discussion must be different.

According to what was stated above on content connections occurring
between fragments of source legal texts, we may assume the following:

a) Content dependence between a definite fragment F of a source legal
text and another fragment F' of the same or another source legal text
constitutes one of the content connections that may have place between
those fragments.

b) The above dependence appears when at least one question connected
with the matter mentioned in fragment F may be indicated, which cannot
be answered without taking the content located in fragment F' into account.

c) Introducing of a cross-reference clause into a given fragment of a
source legal text creates context dependence of this fragment on another
fragment or on other fragments of the same text, or on other fragments of
texts of the same kind, in particular on a fragment or fragments expressing
the contents to which fragment F refers. Such dependence is called anaphoric
dependence. The other kinds of content dependence are called non-anaphoric
dependence.

b) The most common cases of non-anaphoric dependence are:

1. Content dependence occurring between a definite fragment F' of a
source legal text and another fragment F of the same or another text of the
same kind, where the former expresses a principle composing lex specialis
in relation to a principle expressed by the latter. This kind of dependence
is called here inclusive dependence;

2. Content dependence occurring between fragment F of a source legal
text that expresses a definition of a certain term T and another fragment
of the same or another text of the same kind in which that term is used.
We call this kind of dependence definitional dependence.

Content dependence does not exhaust the whole class of content connec-
tions while the dependence mentioned above in $d_1$ and $d_2$ does not exhaust the class of non-anaphoric content dependence that may appear between some fragments of source legal texts.

In all the above examples content dependence of fragment $F$ of a given source legal text on fragment $F'$ of the same or another text have occurred, while the reverse relation (content dependence of fragment $F'$ on fragment $F$) has not taken place. A case like these will be called non-reciprocal content dependence. However, there may also appear cases in which a definite fragment $F$ is content dependent on another fragment $F'$, and at the same time fragment $F'$ is content dependent on fragment $F$. Such cases are called reciprocal content dependence.

The above terms «anaphoric dependence» and «non-anaphoric dependence» obviously refer to much broader problems, namely, to the question of anaphoric expressions (cf. 14 & 15).

13. Systematics of Content Dependence

In the light of the above facts the systematics of content connections may be presented with the use of the following diagram:

\[
\begin{array}{c}
P \\
\downarrow & \downarrow \\
U & \sim U \\
\downarrow & \downarrow \\
A & \sim A \\
\end{array}
\]

where: symbol $P$ designates content connection;
symbol $U$ designates connection that composes content dependence;
symbol $\sim U$ designates connection that does not compose content dependence;
symbol $A$ designates anaphoric content dependence, and
symbol $\sim A$ designates non-anaphoric content dependence.
14. ANAPHORIC EXPRESSIONS

The problems connected with the question of cross-reference clauses appearing in natural language texts, and — among others — in source legal texts bear certain similarity to the problems connected with the so-called expressions, and to the problems connected with the linguistic phenomenon labelled by linguists as text *deixis*. Therefore, in our further discussion we will pay some attention to these two groups of problems.

The feature of being an anaphoric expression is among others ascribed to pronouns used for expressing of connection between a text fragment in which the given pronoun has appeared and a discourse object that has occurred earlier in the text or has been present in the context. We use a neutral term «discourse object» because we do not want to settle too early the question of what is the object to which the pronoun refers. Especially that the choice is usually supposed to be between:

1) a text fragment located above the fragment in which the pronoun has appeared,
2) the content expressed by the above part of a given text,
3) the content located in memories of participants of a given communication process. Other objects might be also taken into account.

A good example of a pronoun used in the above role is the word «it», which appears in the second sentence of the following text: «Rysy is the highest mountain of the Polish part of the Tatra. It is 2499 m high». In this example anaphoric connection crosses the border between the two sentences though often it is limited to one sentence.

What is more, anaphoric reference may be directed forward, and not backwards. It is so, because a pronoun may express a relation with a discourse object that will only appear in a text or will accompany it a moment later. The pronouns used in such a way are sometimes called *kataphors*, and the term «anaphora» is then restricted to those pronouns which refer backwards.

Other expressions employed in order to receive connections similar to those described above are also counted among anaphoric expressions. Among others, the expressions referring to certain events or situations (for instance «then» or «at that time» used in a proper way) or to places of their action (for instance «there» used in a proper way) belong there. Problems connected with the issue of non-pronominal anaphoric expressions,

16 Non-pronominal expressions employed for creating anaphors in English were analyzed by G. Hirst (1981), in Polish — by J. Fall (1988). These expressions, and especially the so-called topo- and chrono-deictic expressions have been studied by M.A. Conte (1981), passim.
similar to those connected with the anaphoric use of pronouns arise. It is obvious that all these questions must be considered separately for every natural language taken into account. In computational linguistics most studies engaged in these matters have been made for English.

Anaphoric expressions have lately become an object of intensive interest among linguists, particularly among authors engaged in computational linguistics. One of the major reasons for this state of affairs were some obstacles met by the theory of machine translation. Namely, certain problems emerged while trying to construe such software solutions that would enable a computer to interpret texts comprising anaphoric expressions properly, in particular, to identify relevant objects having content connections (as described above) with expressions under discussion.

Hitherto it has been characteristic for formation of opinions on substance and functions of anaphoric expressions that new solutions (some of them pretending to be universal ones) appear from time to time but – as a rule – they fail rather quickly. In particular, after a short time it occurs that not all cases of anaphoric expressions are susceptible to interpretation procedures based on the proposals of the innovators. Nonetheless, at least some of the innovations leave lasting traces in the theory of anaphoric expressions. Though they do not supersede former formulations totally, yet, they enrich the theory and improve its explanatory force.

15. Directions of research on anaphoric expressions

15.1. Two main trends appeared in research on anaphoric expressions at the start. The first, which may be called syntactic, recommended investigating of these expressions with the exclusive use of formal analysis. The so-called theory of formal substitution might serve as a pattern example of that syntactic approach. It assumed that an anaphorically used pronoun substituted in a given fragment of a text an expression used earlier in the same text. The latter expression was usually called the antecedent (of the pronoun). It was also assumed that a pronoun used in an anaphoric way may be – without a change in the meaning of the text fragment involved – replaced with its antecedent. The theory under consideration was falsified by quoting of examples in which such replacement let to a change in the meaning of the text fragment in question.\(^7\)

\(^7\) In English one of such cases is the replacement of the sentence «Every man is happy when he is in love» with the sentence «Every man is happy when every man is in love». Cf. J. Hintikka & J. Kulas (1985), p. 59.
For some time the essence of anaphoric expressions and, first of all, of anaphorically used pronouns was seen in the fact that they were alleged to have the status of quantifier-bound variables. The antecedents of these pronouns were supposed to play the role of these quantifiers\(^\text{18}\). However, it appeared afterwards that not all anaphoric pronouns could be treated as variables. The notional apparatus referring to predicate calculus was then replaced almost completely by another in which the term «coreference» was treated as the central notion.

According to this approach an anaphoric pronoun is always bound with its antecedent in such a way that they both (co)refer to the same object of the extra-linguistic reality. This approach, which may be called theory of coreference, constituted the other trend, namely the semantic trend in research on the essence and functions of anaphoric expressions. This trend was afterwards weakened by indicating of the examples in which anaphoric expressions undoubtedly appeared though the above-mentioned relation of coreference did not take place\(^\text{19}\).

An undeniable contribution of approaches from the semantic trend was disclosing of the fact that elucidation of the essence and functions of cross-references is not possible without the use of semantic analysis. The conflict between the syntactic approach and the semantic approach brought about – what is quite normal in such situations – some middle-course ideas\(^\text{20}\).

15.2. In recent years some new approaches have appeared in the theory of anaphoric expressions. They are comprehensively called pragmatic approaches. These approaches postulate reaching beyond the limits of the two types of analysis described above. They recommend looking for the

\(^{18}\) A convincing critique of this approach is given by H. Hiz (1968), passim. Lately, the approach under consideration was strongly attacked by the work cited in note 17. Its authors state that quantifying used in formal languages is considerably different from quantifying that appears in natural languages.

\(^{19}\) R.G. Reilly (1984) following C.L. Sidner (1983) gives one of such examples: «My neighbor has an Irish Wolfhound. They are really hunge but friendly dogs». The word «they» used in the second sentence of this example is undoubtedly an anaphoric pronoun. This pronoun refers to a certain class of dogs, namely to the class of Irish Wolfhounds, though its «antecedent» is another object, namely one of the elements of the above class. The coreferential approach was also criticized by J. Hintikka & J. Kulas (1985). These authors indicate particularly the vagueness of the notion of «coreference» as used in linguistics. Cf. especially p. 82.

\(^{20}\) G. Leech’s conception may be counted among such ideas. This author assumes that an anaphoric reference is a reference to a syntactic entity. However, that entity is identified with regard to semantics.
objects to which anaphoric expressions refer in certain states occurring inside the participants of communication processes, and particularly – but not exclusively – in the states occurring in them as a result of producing and receiving of a text. The theories in question assume in particular that transition of a text to a receiver results on his side in occurrence of a cognitive structure that represents the domain of extra-linguistic reality the text talks about. When the process is successful the structure mentioned above resembles the structure that appeared previously on the side of the transmitter. Pragmatic theories often use the term «discourse models» for the described structures. It is often stated that a discourse model consists of certain cognitive structures created by the participants of a communication process. These structures correspond with entities of the above-mentioned extra-linguistic reality and appear thanks to the processes of creation and reception of the text. Apart from such structures discourse models are composed of other structures that have been established in the knowledge of participants of a communication process earlier.

Adherents of pragmatic approaches caution against identifying of a discourse model with the whole content of the memory of an actor of a communication process. Besides, they assume that a discourse model changes while the process is going on. Some of the authors tend to associate the problems connected with the discourse model with the question of the so-called focus, i.e., that part of a text on which the receiver's attention is currently concentrated. In particular, they identify the content of a discourse model with such a part of the content of memory (consciousness) that is activated by the focus in a given moment. While the focus moves to the text end, it also updates the content of memory. Namely, the parts of memory content that have been brought up to date for some time retire into the background, or «leave the scene». They withdraw making room for other parts, and at the same time they leave the sphere in which they could be accessible by references with the use of pronouns or other anaphoric expressions. The latest fact imposes a restriction on effective employment of anaphoric expressions. It consists in this that an attempt to use them while referring to discourse objects that have already left the

---

21 Cf. B.L. Webber (1979).
22 Cf. B.L. Webber (1979), p. 29. This shows that the term «model» has a different meaning from that used in logical semantics where signifies (in simplification) the domain of extra-linguistic reality to which a given discourse refers.
23 Cf. B.L. Webber (1979), p. 29 and below.
above-mentioned sphere of accessibility may cause that the text will not be understood by the receiver.

15.3. There is a novelty that distinguishes all the pragmatic approaches from the earlier approaches. It is the thesis stating that employment of an anaphoric expression by the text producer is not a reference to this or that component of the text but a reference to the proper constituent of a discourse model (not identical with the text). Therefore, as adherents of that approach state – it is not essential for an anaphor pronoun to have a reference identical with the reference of another expression appearing earlier in the text. It is essential, however, that it refers to a definite – and still accessible for such reference – component of the discourse model.

We may state that pragmatic approaches give evidence of their superiority to earlier approaches. This advantage results from this that they explain more completely the work of mechanisms of identifying (by the participants of communication processes) of the objects to which anaphoric expressions occurring in the text refer. In particular, pragmatic approaches show that apart from making use of the criteria that may be reconstructed while examining this part of the text, which precedes the occurrence of an anaphoric expression, those mechanisms apply also other criteria. Namely, they apply extra-textual criteria, i.e., the criteria grounded in the knowledge of the actors of communication process about that domain of reality, which is the topic of the text. These criteria are reconstructed from the memory (consciousness) of the actors of communication process. Applying of these extra-textual criteria becomes necessary when the criteria reconstructed from the text itself do not suffice for performing of the identification mentioned above. It takes places particularly when several objects compete with each other to play the role of the object of reference and the competition among them cannot be resolved solely on the basis of the criteria yielded by a text.

15.4. The changes that have taken place in understanding of the essence and performance of anaphoric expressions have not brought along a decisive progress in constructing of programs that might interpret text fragments comprising such expressions automatically. Though different solutions have been given they are generally deceptive and work only in very limited

---

It is underlined in the pragmatically oriented literature on anaphoric expressions that not all discourse objects are accessible as objects of reference with the use of these expressions. The term «anaphoric islands» is used for objects not accessible to such references. The examples of anaphoric islands are given by G. Leech (1974).
spheres of application. Lack of uniformity of conceptions in the theory of anaphoric expressions causes that the constructors of required programs are inclined to apply eclectic solutions, i.e., to build programs that employ, simultaneously or in sequence, solutions based on various theoretic bases. Those programs work most often in such a way that their performance begins with application of procedures utilizing only formal criteria. When the results received at that stage appear to be unsatisfactory, other procedures, particularly the procedures utilizing semantic criteria (especially those taken from a lexicon of a given system), are applied. When the application of semantic criteria also proves a failure, the program passes on to applying of criteria taken from the system knowledge on an appropriate domain of extra-linguistic reality.

16. ANAPHORIC EXPRESSIONS AND CROSS-REFERENCE CLAUSES

Not all cross-reference clauses appearing in legal texts may be counted among anaphoric expressions. Vacillation and diversity of views on anaphoric expressions cause that it is not easy to establish criteria application of which would make it possible to decide whether a given cross-reference clause belongs or does not belong to the aforementioned expressions. It seems, however, that there exists a specific feature that settles the question of including of a specific cross-reference clause in the class of anaphoric expressions. This feature consists in a characteristic way of informing about the location of the content to which a given clause refers. In particular, in cross-reference clauses being also anaphoric expressions the information on that location is given indirectly, namely by informing about the position taken by the fragment in which that content is located, in relation to the position of the fragment in which the cross-reference clause appears. It manifests itself in occurrences of phrases like «in the previous article», «the next article», «foregoing paragraphs», etc. We will not treat as anaphoric expressions such cross-reference clauses that do not fulfill the above criterium, i.e., the clauses in which the information on the location of proper content in the text takes another form, for example, the form of a quotation of the direct address where a given fragment may be found in the text, will not be treated as anaphoric expressions.

17. CROSS-REFERENCE CLAUSES AND TEXT DEIXIS

As regards relations between the problems connected with the issues of cross-reference clauses and the so-called text deixis, we cannot discuss them
without expressing our views on the scope of the latter. For the phenomenon of deixis is a matter of argument among text theorists. In particular, some of them tend to see a deixis in all the cases in which in a text there appears an expression indicating another fragment of the same text. Others restrict the phenomenon of deixis to such cases only in which the indication takes a specific characteristic form consisting in application of the so-called topo-deictic or chrono-deictic expressions. Phrases like «in the previous chapter», «in the next paragraph», «above», «below», «earlier», «lately», are counted among those expressions.

C.J. Fillmore seems to hold the former, more general, standpoint. In his articles devoted to text deixis he does not state that an indication of a proper fragment has to employ the expressions mentioned above. M.E. Conte represents the other position in this matter.

It is very important to assume a definite attitude in this matter, for, if we shared the former standpoint we would have to include practically all the types of cross-reference clauses presented below into the category of text deixis. Clauses with a direct address of the fragment to which a given clause refers, e.g., an address indicating the number of a proper article or paragraph, would also be treated then as text deixis. We reject such an attitude and assume that one of the constitutive features of every deixis (text deixis included) is that the meaning of expressions in which a deixis appears is dependent on specific circumstances of their use. In the case of text deixis those «circumstances» mean, among others, that a proper expression has been used in a definite place of a text and the use of this very expression, e.g.: «in the previous chapter» – in another place of the text would have a completely different meaning.

Therefore, we join the latter position. Consequently, we include only two from among several types of cross-reference clauses described below in the category of text deixis. These two will be called cross-reference clauses with deictic indication and cross-reference clauses with semaphoric indication.

18. CONCLUSION

Our considerations concerning particular types of references are usually ended with conclusions that may serve as a direct basis for applying of proper software solutions. Those conclusions are thus equivalent to algori-

thms that may be formulated in proper programming languages. The solutions offered in our work devoted to cross-reference clauses in legal texts may also be of some use in legislative drafting. They can play the role of instructions, complying with which will ensure formulation of properly built reference clauses. As it may be noticed the work was arranged so that its results could be applied both in constructing of programs for automated interpretation (resolution) of cross-reference clauses and in legislative drafting.

Our work refers to empirical material taken from Polish source legal texts constituting a survey of 200 randomly chosen normative acts that were published in the Dziennik Ustaw [Journal of Law of the Polish People's Republic] or in the Monitor Polski [Polish Official Gazette] in the years 1944-1979. Apart from that a body of some Polish codes quoted here came into the material mentioned above. We do not think that restricting of our research only to Polish text material will make the employment of the proposed solutions to other national legal systems impossible. Our opinion is based on the fact that the structural features of the Polish legal texts taken into account as significant for cross-references' analysis – for instance the features connected with division of a text into hierarchically configured fragments of different levels, and the features connected with formulating and allocating of headings or applying of a definite enumerating method – are repeated (possibly slightly modified) in text material of many legal systems, especially in the European (continental) law systems. Therefore – in our opinion – the direction which we have tried to indicate, may also appear useful for the researchers working in the field of cross-reference of text material different from the Polish example.

REFERENCES


