An Analysis of Plane Task Text Ellipticity and the Possibility of Ellipses Reconstructing Based on Cognitive Modeling Geometric Objects and Actions

Xenia Naidenova[1], Sergei Kurbatov[2], and Vjacheslav Ganapol'skii[3]

[1] Military Medical Academy, Saint Petersburg, Russian Federation
ksennaidd@gmail.com
[2] Research Center of Electronic Computer Engineering, Moscow, Russian Federation
curbatow.serg@yandex.ru
[3] Military Medical Academy, Saint Petersburg, Russian Federation
Ganvp@mail.ru

Abstract. The article describes the processing of ellipses in an automated system of solving planimetric tasks according to their description in natural language. An approach is proposed to processing ellipses basing on cognitive semantics. The resolution of ellipses is based on using syntactic structures and semantics of geometry in parallel. The types of ellipses most frequently encountered in geometric tasks are revealed. A new approach to recognizing and resolving ellipses in the framework of cognitive semantics is offered.

Keywords: ellipsis resolution, cognitive semantics, planimetric task, text understanding.

1 Introduction

The ambiguity of natural language caused by homonymy has long been studied by computer linguistics. However, the ambiguity associated with the omission of a thinkable language unit (ellipsis) in text has been actively analyzed in natural language processing relatively recently [1], [2]. Although in theoretical linguistics ellipticity got enough coverage [3], [4], restoration of ellipses in systems of syntactic text analysis is clearly developed not enough. Firstly, this is largely due to the fact that eliminating ellipticity is subordinate to actual syntactic analysis and, secondly, this is caused by complexity of resolving ellipses.

The complexity is explained by the necessity to consider a number of contexts: current sentence, adjacent sentences, already established syntactic relations and, finally, semantics of the text. This work is divided into two parts. In the first part, it is described how to handle ellipticity in a specific holistic system of solving plane geometry tasks described in natural language. This system has been implemented in the framework of the INTEGRO project (INTEGRating Ontology) [5]. The second part proposes a new approach to the processing of ellipses based on cognitive semantics.

2 Resolving ellipses in the texts of geometrical tasks

2.1 Syntactical analysis

The architecture and principles of functioning of the system for solving geometrical problems are described in [6] and its general scheme is illustrated by Fig. 1.

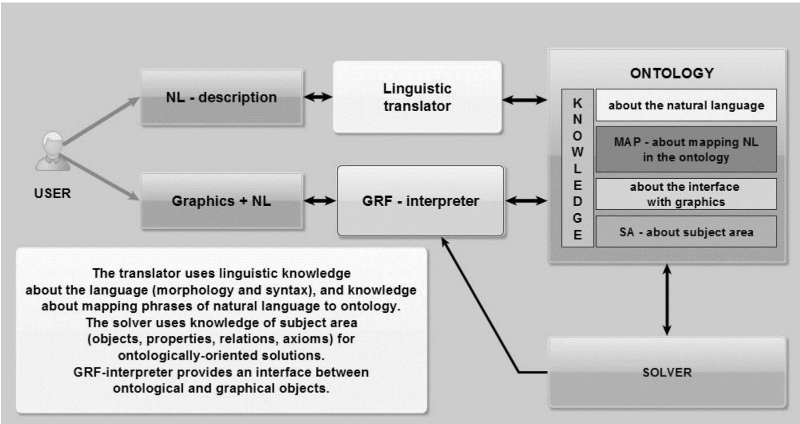

Fig. 1. Scheme of the system for solving geometrical tasks.

The system includes the following blocks: linguistic translator, ontology, solver, and graphical module for displaying and explaining the results (drawing NL-explanation of the solution process). The solver receives the ontological structure of the task and forms a chain of basic operations using knowledge of the subject area. In this section, we concentrate on the extension of the system to correctly interpret elliptical (incomplete) sentences.

The language translator creates a syntactic structure and determines that some of its elements violate the language rules. For example, there is no noun for the adjective, the pretext is at the end of the sentence, the number does not have a mandatory measuring unit, and so on. The basic criteria for determining ellipticity are studied by linguists [7, 8]. Based on these criteria recorded in the ontology, the translator identifies the fragments of the syntactic tree that admittedly contain ellipticity. Next, with the use of algorithms described in short below in section 2.2, the identified ellipses are restored. Specifically, in sentence “the radius of the first circle equals 12 cm, and the second 10 cm”, the elements “second” and “10” define the ellipticity. As a result, two syntactical structures are formed:

• The radius of the first circle equals 12 cm;

• The radius of the second circle equals 10 cm.

These structures are further processed by the system mechanisms of paraphrasing to obtain an ontological representation of sentence in the formal terms of the subject area [6]. The concept “paraphrasing” has been proposed by the well-known Russian linguist Apresyan in [9]. In our system, we use an adaptive variant of this concept.

The conception of paraphrasing assumes that any class of sentences corresponding to one and only one sense can be reduced to the simplest or canonical phrase composed only of the lexemes expressing most clearly the basic concepts of sentences. Thus, paraphrasing is based on the following proposition in [9]: “One of the fundamental properties of human languages consists in the fact that if there are several synonyms, in the broad sense, to express some concept, then only one of them turns out to be privileged, canonical, or prototypical for expressing this concept”. In particular, such canonical concepts in plane geometry are, for example, the point, the line, the plane and to belong, to lie between, and to be congruent. Thus the rules of paraphrasing provide only one canonical form for a group of sentences having the same sense. For example, sentences “a point located on the straight line”, “the straight line passing through a point”, “a point belonging to the line”, “a point lying on the line segment” etc. are reduced to the following canonical phrase “point belongs to straight line”. This canonical phrase is mapped to its ontological representation in the form of the following triplet “point lies line”. It should be stressed that the members of the triplet (objects and relations between them) are not dependent on a language. Therefore the corresponding rule of paraphrasing contains, in its left part, the objects and relations depending on language, but, in its right part, the formal objects and relations invariant in different languages.

The rules of paraphrasing are divided into two classes; the first one consists of rules in which both parts are some generalized syntactic structures; the second one consists of rules having canonical descriptions in their left parts and semantic descriptions in their right parts. The second class of rules can be used for transforming ontological structures into corresponding natural language texts. It is reasonable to apply the rules of the first class to equivalent synonymic transformations of synthesized structures to retrieve texts in the most appropriate manner in a considered application domain.

2.2 Algorithm for resolving ellipticity

The algorithm for treating ellipses is based on the ontology knowledge reflecting the semantic hierarchy of word forms in the syntactic structure and the norms of natural language. To a first approximation the algorithm can be described as follows:

- to segment a syntactic structure into two segments: a complete one without ellipticity and the other one containing ellipticity (generally, it is a set of noun groups (NG));
- in the elliptical segment, to reveal the elements that are supposed to be used for resolving ellipticity;
- in the full syntactic structure, to reveal the candidates to be replaced by the elements found in the previous step;
- to perform the replacement and obtain the complete syntactic structure.

In the example given in section 2.1 “*first*” is replaced by “*second*” and “*12*” by “*10*” because they correspond to the same concepts of ontology. Here we have different objects and the same type of attribute (length). In the sentence “*the perimeter of triangle is 37 cm and the area – 20 cm*” we have the same object and different types of attributes. This seemingly simple algorithm allows to successfully recover not only geometrical ellipses, but several others, described, for example, in [2]: in the sentence "*twenty years of such dance form the age, forty – the history*" "*twenty*" is replaced by "*forty*" and "*age*" is replaced by "*history*".

2.3 Limitations

Of course, many cases of ellipticity cannot be processed by the algorithm above. Example: "There are seven circles. Radius of one 5 cm, two others − 3 cm, and the others − 10 cm". We have multiple ellipticity in this example. A similar example from [2]: "Anemones discard tentacles, crabs − claws, lizards − tail". In many cases, ambiguity arises at the level of comparison. Two options were analyzed: 1) to move forward with analyzing the situation and eliminating ambiguity at the stage of semantic processing; 2) to complement the ontology by the rules of preferences when choosing a candidate for replacement (substitution). It should be noted that the question of clear ellipticity criteria and methods for restoring the full structure of sentences has not been fully resolved within the framework of a generally accepted linguistic theory. Resolving ellipses in natural language texts remains one of the most difficult and unsolved tasks in linguistics, despite the abundance of proposed methods based on syntactic-semantic parsing of sentences. Syntax reveals the structure of the ellipsis and the similar part of the sentence without it; semantics deals with word values. However, as the example from [11, page. 62] shows, resolving ellipses is based on the understanding of context (text theme), the sense of words and collocations: "Charles makes love with his wife twice a week. So does John".

2.4 Testing the algorithm

The algorithm performs the ellipses' resolution with the accuracy equal to 100% in simple cases when the noun phrase in a sentence consists of only one word. It is important to note that resolving ellipses is directly connected with the correct functioning the system ontology, since the ontology supports the process of sentence understanding. In more complex cases with the composite noun phrases or incomplete ontology, the accuracy of the algorithm declines to 70 %. In any case, difficult texts of some planimetric tasks require the special analysis and the solution ad hoc.

Currently, several hundred of simple ellipses and several tens of complex ones have been tested.

In general, it should be anticipated that the vast majority of sentences contains several types of ellipses or some number of ellipses of the same type. This fact implies the search for some new approach to reconstructing ellipses covering not only the ontology and linguistic knowledge but also the model of human plausible reasoning and cognitive model of practical geometrical situations. Ellipsis resolution must be based on cognitive semantics.

3 Ellipsis classification in geometrical tasks

To study the typology of ellipses in geometric tasks we used a body of texts containing more than 1000 planimetric tasks. We have revealed the following types of ellipses: ellipses using dash "–" (ellipses with skipped predicate or verb), ellipses without "–" (ellipses with skipped verb, noun, pronoun, or predicate. Consider the structure of these ellipses. We will give only fragments of tasks containing ellipses.

Skipped predicate: *In triangular ABC there are given R and r – radii of circumscribed and inscribed circles. A_1, B_1, C_1 – points of crossing the bisectors of triangle ABC with the circumscribed circle.*

Structural components of these ellipses are Noun Phrase (NP) and Prepositional Phrase (PP). Revealing NP and PP is realized in the system OntoIntegrator [12] in the framework of the project on creating World Digital Mathematical Library – WDML.

Consider this type of ellipses in greater detail:

a) < NP > < – > < Designation(s) (Bases of perpendiculars dropped from B and D on AC – M and N);

b) < Designation(s) > < –> < NP > (O1, O2, O3, O4 – centers of circles; D – arbitrary point of the plane; BD – the side of rectilinear pentagon inscribed in this circle);

c) < NP > < – > < NP > (The points of their intersection lie on the same circle – the circle of nine points; This quadrangle is a diamond; Every parallelogram inscribed in a circle – rectangle; Every diamond inscribed in a circle – square);

d) <NP> < – > <PP> (Center of circle – inside the quadrangle; C – between A and F);

e) < NP > < – > <The property expressed by adjective> (Angle C – right; To find a point on a given line such that the sum of distances from it to two points A, B – minimal).

The resolution of these ellipses can be carried out according to the scheme:

to select NPs; to identify the heads of NPs as geometrical objects; to identify designations; to localize the dash between the designation(s) and the NPs; to check (according to the rules of the ontology) the conformity between the designations and the heads of the NPs; to restore ellipses. In these cases, the dash is replaced by the forms "is" or "are" of the verb "to be".

The dash in the Russian language is put in a variety of situations. In situation c), the dash is put between the subject and the predicate in the absence of a link between them [13], if both members are expressed as nouns in the form of the same case, for example, "Loneliness in a creative work – a hard thing", "The next station is Mytishchi". In geometric problems, situation c) has the nature of a logical definition (geometry – a section of mathematics) or identity, when the subject and the predicate are expressed by the same concept. If the subject and the predicate are not expressed by the same word, then it is necessary to check the predicative relation through logical inference in the ontology.

In view of our consideration of Verb Phrase Ellipsis in the previous section we confine ourselves to one of difficult cases of this ellipsis.

Skipped verb (ellipsis with dash): *In triangle ABC there are taken points M, N and P: M and N – on sides AC and BC, P – on line segment MN.*

In this sentence, we have an incomplete VP: *In triangle ABC there are taken points M, N and P* (presupposition), this VP is prolongated by the follow way:

In triangle ABC there is taken point M on side AC;

In triangle ABC there is taken point N on side BC;

In triangle ABC there is taken point P on line segment MN.

Restoration of this sentence is supported by a thinkable geometric situation, (let us call it a **cognitive model of a geometric situation**). And the restoration goes on sequentially, but with simultaneous creation of different relationships: temporary (earlier, later), referential (the designation refers to an object, the pronoun refers to an object), spatial (in the triangle, on the side), linguistic (links of relationships, objects, properties with certain word forms and expressions), quantitative. So, in our example we have ($\rightarrow$ means a reference):

In triangle ABC $\rightarrow$ triangle $\rightarrow$ designation = ABC;

Triangle ABC $\rightarrow$ one $\rightarrow$ it $\rightarrow$ it is given $\rightarrow$ this $\rightarrow$ in it;

Triangle ABC $\rightarrow$ side AC $\rightarrow$one, side BC $\rightarrow$ two, side AB $\rightarrow$ three

In triangle ABC there are taken points M, N, and P;

Point one $\rightarrow$ designation M $\rightarrow$ first, point two $\rightarrow$ designation N $\rightarrow$ second;

Point three $\rightarrow$ designation P $\rightarrow$ third;

In triangle ABC there is taken point M; in triangle ABC there is taken point N; in triangle ABC there is taken point P;

Now we need a model of acting: “to take point in a tringle” and **generating hypotheses** “Where?”. In accordance with one of the hypotheses the following cases are:

In triangle ABC there is taken point M (one) on side AC (one);

In triangle ABC there is taken point N (two) on side BC (two);

By analogy:

In triangle ABC there is taken point P (three) on line segment MN.

Line segment $\rightarrow$ designation MN $\rightarrow$ it joins points M and N (supported by knowledge about how a segment of a line is generated).

As a result, we can restore the full text of this task: *In triangle ABC there is taken point M on side AC*; *there is taken point N on side BC***,** *and there is taken P on line segment MN.*

The process of binding objects during their construction is supported by cognitive models of objects and operational knowledge. As D. Suleymanov [14] noted, “it is necessary to go not from the text, but from the task”. All cognitive models can be explicitly defined based on geometric semantics and they are associated with speech parts and typical collocations with their grammatical categories at the sentence level.

Restoration of the full text requires reasoning by analogy and understanding the meaning of actions with geometric objects. Exactly, similar actions are supposed with similar objects, and therefore the words are skipped. In practice, most skipped words are redundant for understanding the sense of sentences. People omit words consciously. However, if the missing information is not redundant, understanding texts represents a problem that is resolved by analyzing geometric situations.

The following sentences give the examples of ellipses without dashes.

Skipped verb, ellipsis without dash: *The vertices of parallelogram $A_1B_1C_1D_1$ lie on the sides of parallelogram ABCD: point A1 lies on AB,* <u>*B_1 on BC*</u>*, etc.).* (Word "*lies*" after B_1 is skipped)

Skipped noun: *Prove that the value of angle with the vertex inside a circle equals the half-sum of the angular values of two arcs of which one is enclosed between the sides of this corner and the other between the prolongation of* <u>*sides*</u>*.* (Words "*of this corner*" after word "*sides*" are skipped).

Skipped pronoun: *In a circle of radius R, two chords AB and AC are drawn. On AB or* <u>*on its extension, point*</u> *M is taken. Analogically, on AC or on* <u>*the extension*</u>*, point N is taken.* ("*of it*" is skipped after "*the extension*").

Skipped predicate: *Side BC of triangle ABC is equal to* <u>*a*</u>*, radii of a circumscribed circle* <u>*r*</u>*.*

4 The structure of cognitive models of objects and actions

Cognitive structures correspond to the semantic structures of situations described in the text. They should be aligned with the narrative structures of sentences. A word can have multiple values, but only one sense, at least in mathematical texts. Ellipsis (omitting words, economy of text) is possible because the preceding text determines unambiguously (uniquely) the meaning of each word and situation, and these meanings remain unchanged. In cognitive models of objects, the following relationships are important:

- object can perform some actions;
- object can be subjected to actions of other objects;
- object can have spatial and temporal relationships (earlier, later, already built, already given) with other objects;
- object can be composed of some other objects;
- object can be a part of some other object (objects);
- object has properties, some of which (call them actant ones) are related to the actions that the object commits (intersects – intersecting, lies – lying) or the actions that are committed over it (has been given – given, has been formed – formed, cut of, embedded). Thus, the actant properties of objects are directly displayed in the morphological forms of words describing these properties;
- the relationships between the properties of one geometrical object and the properties of others.

These relationships are in agreement with the universals described by D. Suleymanov [15]. The properties between an object and its parts are realized through implications: if center, then a circle; if radius, then a circle; if circle, then circumscribed about or inscribed in; if inscribed in, then in an object; if bisector, then bisector of an angle; if bisector of angle, then the vertex of angle from which it originates; if bisector, then the angle from which it comes is divided in half; if bisector, then it is the axis of symmetry of angle divided in half by this bisector.

The interaction of cognitive models and the analyzed text should provide the principle of "cognitive expectation" and "determinism of context" [14].

Creation of cognitive models of objects and actions for plane geometry, in the proposed approach, is performed in a step-by-step mode by the use of a given text corpus. Some fragments of cognitive model “Bisector” are shown in Tables 1 and 2. It is also a problem of considerable interest to apply a plausible reasoning for resolution of ellipses, including analogy, generalizations, specialization, use of implications, forming hypotheses and many others.

Table 1. Noun Phrases with “Bisector”.

Bisector	Hyperlink to object (to NP)	Hyperlink to object (to PP)
Bisector of	angle	
Bisector of	angle	in (of) triangle
Bisector of	acute angle	in (of) rectangular triangle
Bisector of	inner angle	in (of) triangle
Bisector of	angle	at base of isosceles triangle
Bisector coming from	vertex	of inscribed triangle
Bisector of	angles adjacent to one side	in (of) parallelogram
Bisector of		in (of) triangle
Bisector of	inner angle	in (of) parallelogram
Bisector of	angle	in (of) convex quadrilateral
Bisector of	angle	in (of) rectangle

Table 2. Verb Phrases with “Bisector”.

Bisector	Hyperlink to object (to NP)	Hyperlink to object (to PP)
Dividing	To divide	Side of triangle
Perpendicular	To be perpendicular	Median of triangle
Splitting, cutting in	To split, to cut in	Side of parallelogram in segments
Intersecting	To intersect	Bisector of triangle
Intersecting	To intersect	Circle
Restricting	To restrict	Area of quadrangle
Coming across	To come across	Circle in points
Containing	To contain	Points of intersection
Lying on	To lie	Straight line

Within the proposed approach, text analysis becomes cognitive-driven, and the parser plays a subordinate role (Fig. 2). If ellipsis resolution is based on cognitive models, then it is possible to synthesize a text describing a geometric situation and compare this text with the text to be analyzed. The ontology contains theoretical knowledge in the area to solve geometry tasks of various types (computational, for construction, for proof). The ontology takes the burden of solving tasks and visualizing solutions. The

Cognitive Analyzer runs incrementally and transmits a converted and meaningful text to the ontology in the form required by it.

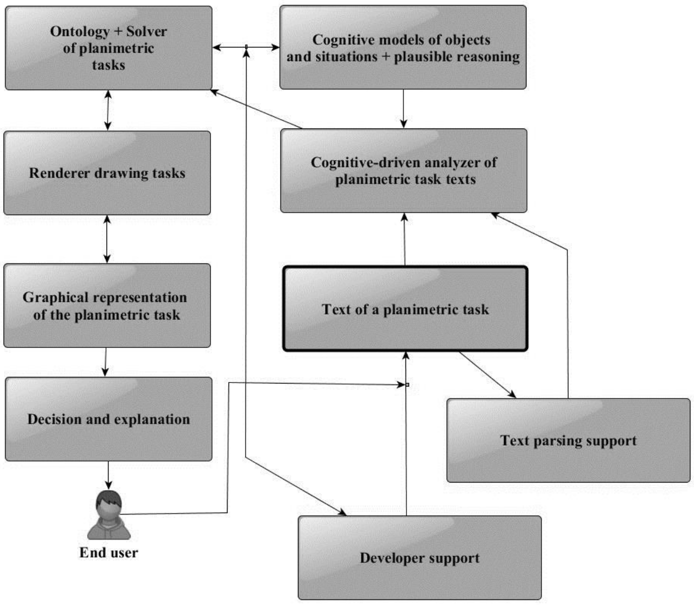

Fig. 2. Scheme of a cognitive-controlled analysis of a text.

5 Related works

Verb Phrase Ellipsis is a well-studied topic in theoretical linguistics but has received little attention as a computational problem and a task of human reasoning except the paper [16]. Exhaustive linguistic analysis of ellipses for different languages performed in many sources: for example, [8], [17- 24].

In spite of the fact that a lot of works deal with resolution of ellipses, the significant results are obtained only for some special types of them, namely for the verb ellipses (VE) in the framework of syntactical-semantic analysis.

Detection and resolution of Verb Phrase Ellipsis (VPE) are considered in the articles [25-30] but only for some special cases: resolving elided scopes of modality and ellipses with auxiliary verbs. In [26], the authors have proposed a method of automatic ellipsis resolution without preliminary processing or annotation of texts. This work is carried out within the OntoSem language processing system of the OntoAgent cognitive architecture. OntoAgents carry out deep semantic and pragmatic language analysis, yielding ontologically grounded text meaning representation that populate agent memory and subsequently support agent reasoning [27].

The text with the VE has the following structure consisting of 2 parts standing on the right and left of the "dash" (both parts are in the same sentence). The verb is skipped in the right part, the left part (the antecedent) contains the verb. The right part is complemented by the verb from the left part. Example: *She can* ***go*** *to Hawaii but he can't (She can go to Hawaii but he can't go)*.

The resolution of such an ellipsis consists of three stages:

- Recognizing the occurrence of ellipsis, localizing it, and selecting its parts;
- finding the nearest to the left verb in the antecedent;
- resolving ellipsis.

The paper [28] describes a system ViPER (VP Ellipsis Resolver) that detects and resolves VP ellipsis, relying on linguistic principles such as syntactic parallelism and modality correlations. The system ViPER has been incorporated into the OntoSem2 incremental semantic analysis system that provides language analysis capabilities to OntoAgents.

In [27], a novel approach is presented to detecting and resolving VPE by using supervised discriminative machine learning techniques trained on features extracted from an automatically parsed, publicly available dataset. Additionally, this approach uses the Margin-Infused-Relaxed Algorithm for antecedent identification. It is proposed a decomposition of the overall resolution problem into three tasks – target detection (ellipsis detection), antecedent head resolution, and antecedent boundary detection.

The features used for antecedent head resolution and/or boundary determination try to capture aspects of both tasks. The features are roughly grouped by their type. **Labelsfeatures** make use of the parsing labels of the antecedent and target; **Treefeatures** are intended to capture the dependency relations between the antecedent and target; **Distancefeatures** describe distance between them; **Matchfeatures** test whether the context of the antecedent and target are similar; **Semanticfeatures** capture shallow semantic similarity; there are a few **Otherfeatures** which are not categorized.

In [30], a new method is proposed to resolve multiple ellipses in such sentences as:

- Unemployment has reached 27.6% in Azerbaijan, 25.7% in Tadzhikistan, 22.8% in Uzbekistan, 18.8% in Turkmenia, 18% in Armenia and 16.3% in Kirgizia;

In this paper, sentences lack an overt predicate. The authors present two methods for reconstructing elided predicates within the Universal Dependencies (UD) framework. The first method adapts an existing procedure for parsing sentences with elided function words [31], which uses composite labels that can be deterministically turned into dependency graphs in most cases. The second method is a novel procedure that relies on the parser only to identify a gap. Then an unsupervised method is used to reconstruct the elided predicates and reattach the arguments to the reconstructed predicate. The both methods work with very high accuracy (from 81,69 to 90,57 %) and significantly exceed the recently proposed constituent parser by Kummerfeld and [32]. The types of ellipses reconstructed are:

(1) Single predicate gaps:

John **bought** books, and Mary____ flowers.

(2) Contiguous predicate-argument gap (including ACCs):

Eve **gave flowers** to Aland Sue_____ to Paul.

Eve **gave** a CD to Al and____ roses to Sue.

(3) Non-contiguous predicate-argument gap:

Arizona **elected** Goldwater **Senator**, and Pennsylvania_____ Schwelker____.
(4) Verb cluster gap:
I want to try to begin to write a novel and ... Mary _____a play. ...
Mary _____to write a play. ...
Mary ______to begin to write a play. ...
Mary ______to try to begin to write a play.

The core characteristic of resolving ellipses is that there is a clause that lacks a predicate (the gap) but still contains two or more arguments or modifiers of the elided predicate. In most cases, the remnants have a corresponding argument or modifier in the clause with the overt predicate. The UD frame work aims to provide cross-linguistically consistent dependency annotations that are useful for NLP tasks. The UD defines two types of representation: the basic UD representation which is a strict surface syntax dependency tree and the enhanced UD representation [33] which may be a graph instead of a tree and may contain additional nodes.

See [34] and [35] for a more comprehensive overview of cross-linguistically attested gapping.

The major advantage of this approach is that the dependency tree contains information about the types of arguments and so it should be straightforward to turn dependency trees into enhanced UD graphs. For most dependency trees, one can obtain the enhanced UD graph by splitting the composite relations into its atomic parts and inserting copy nodes at the splitting points.

A crucial step is the third step, determining the highest-scoring alignment. This can be done with the algorithm presented by Needlemann and Wunsch [36] in which one defines a similarity function *sim*(g,f) that returns a similarity score between the arguments g and f. Defining *sim* based on the intuitions that often, parallel arguments are of the same syntactic category, that they are introduced by the same function words (e.g., the same preposition), and that they are closely related in meaning.

Seeker et al. [31] compared three ways of parsing with empty heads: adding a transition that inserts empty nodes, using composite relation labels for nodes that depend on an elided node, and pre-inserting empties before parsing. These papers all focus on recovering nodes for elided function words such as auxiliaries; none of them attempt to recover and resolve the content word elisions of gapping.

6 Conclusion

Processing ellipses is given in a specific system of plane geometry tasks described in natural language. Ellipsis resolution is based on using in parallel the syntax structures of sentences and the geometry semantics. A broader approach to ellipses processing based on cognitive semantics has been proposed. The approach gives a classification of ellipses (across a geometric text corpus) and introduces the concept of a cognitive model of geometry objects and actions. This approach allows to view the structure of automated analysis of geometric texts as a cognitively controlled parsing.

Acknowledgments. The research was partially supported by Russian Foundation for Basic Research, research project No. 18-07-00098A.

References

1. Malkovsky, M. et al.: Restoring ellipse as a task of automatic text processing. Program Products and Systems 3(107), 32-36 (2014) (in Russian).
2. Kobzareva, T., Epifanov, M., Lakhuti D.: Finding the antecedent of the ellipsis of a fragment containing predicate (automatic analysis of Russian sentence). In: Proceedings of 15th National Conference on Artificial Intelligence with International Participation (CAI-2016),vol. 2, pp. 56-62 (2016) (in Russian).
3. Adametc, P.: Semantic interpretation of “significant zeros” in Russian sentences. Language and verse in Russia. In: Proceedings of papers in honor of Dean C. Worth for his 65 Anniversary, pp. 9-18. East Literature, Russian Academy of Sciences, Moscow (1995) (in Russian).
4. Vardul, I.: To the question of the ellipse phenomenon. In: Proceedings of the conference “Invariant Syntactical Values and Structure of Sentence”, pp. 59-70 (1969) (in Russian).
5. Kurbatov, S., Naidenova, X., Khakhalin G.: Integrating intelligent systems of analysis/synthesis of images and text: project outlines INTEGRO. Contours of the INTEGRO project. In: Proceedings of the International Scientific Conference “Open Semantic Technologies for Intelligent Systems” (OSTIS-2011), pp. 213-232 (2011).
6. Kurbatov, S., Vorobyev, A.: Ontological solver of geometry problems in natural language description. In: Proceedings of 15th National Conference on Artificial Intelligence with International Participation (CAI-2016), vol. 1, pp. 56-63 (2016) (in Russian).
7. Larkina, A.: About the lingustic investigations of ellipcises in French languauge. Bulletin of the Chelyabinsk State University 5(143), 7-29 (2009) (in Russian).
8. Minjaylov, V.: The peculiarities of elliptical construction in Russian language. World of Science and Innovation 1(1), 27-31 (2015).
9. Apresyan Y., Boguslavsky, I., Iodmin, L., et al.: Linguistical maintenance of the system ETAP-2. Nauka, Moscow (1989) (in Russian).
10. Lobzin, À., Khakhalin, G., Kurbatov, S., Litvinovich, A.: Integration based on natural language and image ontology in the system Text-To-Picture. In: Proceedings of 8th Scientific-Practical Conference “Integrated Models and Soft Computing in Artificial Intelligence, pp. 296-305 (2015) (in Russian).
11. Eco, U.: The role of the reader. Exploration in the semiotics of texts. Publishing House “ACT”: CORPUS, Moscow (2016) (in Russian).
12. Nevzorova, O. and Nevzorov, V.: Terminological annotation of the document in a retrieval context on the basis of technologies of system “OntoIntegrator”. International Journal “Information Technologies& Knowledge”, 5(2) (2012).
13. Rosental, D.: Reference book on Russian language: orthography and punctuation. Publishing House “ACT”, Moscow (2013) (in Russian).
14. Suleymanov, D.: Pragmatic-oriented linguistic models as a basis of systems and technologies for processing natural language. Analytical Review. In: Sosnin P.I. and Nevsorova O.A. (eds.) “Formal Models and Systems in Computational Linguistics”, Chapter 1 (pp. 8 - 59). Republic Tatarstan Academy of Sciences, Kazan (2016) (in Russian).
15. Suleymanov, D. and Gatiatullin, A.: A system of semantical universals and their implementation in the form of relational-situational frame. In: Stephanyuk V. and Taysina E.

(eds.) "Cognitive-Semiotics Aspects of Modelling in Humanitarian Sphere" (pp. 185 - 210). Republic Tatarstan Academy of Sciences, Kazan (2017) (in Russian).
16. Kubota, Y. and Levine, R.: Gapping as hypothetical reasoning. Natural Language and Linguistic Theory 34(1), 107-156 (2016).
17. Kobozeva, I.V.: Linguistical semantics. Publishing House "URSS", Moscow (2000) (in Russian).
18. Mitkov, R.: Outstanding issues in anaphora resolution. In: Gelgukh A (ed.) "Computational Linguistics and Intelligent Text Processing" (pp. 110-125). Springer (2001).
19. Nielsen, L. A.: A corpus-based study of verb phrase ellipsis identification and resolution. Ph.D. thesis. King's College, London (2005).
20. Nielsen, L. A.: Verb phrase ellipsis detection using automatically parsed text. In: Proceedings of the 20th International Conference on Computational Linguistics (COLING 2004). V. 1, pp. 1093-1099 (2004).
21. Bengtson, E., and Roth, D.: Understanding the value of features for coreference resolution In: EMNLP'08 Proceedings of the Conference on Empirical Methods in Natural Language Processing, pp.294-303 (2008).
22. Carberry, S.: A pragmatics-based approach to ellipsis resolution. Computational Linguistics 15(2), 75-96 (1989).
23. Carbonell, J. G.: Discourse pragmatics and ellipsis resolution in task-oriented natural language interfaces In: ACL'83 Proceedings of the 21st annual meeting on Association for Computational Linguistics, pp. 64-168 (1983).
24. Ivan, S. and Hankamer, J.: Toward a theory of anaphoric processing. Linguistics and Philosophy 7, 325-345 (1984).
25. McShane, M. & Babkin, P.: Automatic Ellipsis Resolution: Recovering Covert Information from Text. In: Proceedings of the Twenty Ninth AAAI Conference on AI, pp. 572-578 (2015).
26. McShane, M., Nirenburg, S., Beale, S., Johnson, B.: Resolving Elided Scopes of Modality in OntoAgent. Advances in Cognitive Systems 2, 95-112 (2012).
27. Kenyen-Dean, K., Cheung, J.C.K., Precup, D.: Verb Phrase Resolution Using Discriminative and Margin-Infused Algorithms. In: Proceedings of the 2016 Conf. "Empirical Methods in NLP", pp. 1734-1743 (2016).
28. McShane, M., Babkin, P.: Detection and Resolution of Verb Phrase Ellipsis. LiLT 13(1), 1-36 (2016).
29. Liu, Z., Gonzalez, E., Gillick., D.: Exploring the steps of VPE. In: Proceedings of the Workshop on Conference Resolution Beyond OntoNotes (CORBON 2016), co-located with NAACL, pp. 32-63 (2016).
30. Schuster, S., Nivre, J., and Manning, Ch. D.: Sentences with Gapping: Parsing and Reconstructing Elided Predicates. arXive: 1801.06922v1 [cs.CL] 18 Apr 2018.
31. Seeker, W., Faras, R., Bohnet, B., Schmid, H., and Kuhn, J.: Data-driven dependency parsing with empty heads. In: Proceedings of COLING, pp. 1081-1090 (2012).
32. Kummerfeld, J. K. and Klein, D.: Parsing with traces: An O(n4) algorithm and a structural representation. Transactions of the Association for Computational Linguistics 5, 1-36 (2016).
33. Schuster, S. and Manning, C. D.: Enhanced English Universal Dependencies: An improved representation for natural language understanding tasks. In: Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016), pp. 2371-2378 (2016).

34. Schuster, S., Lamm, M., and Manning, C.D.: Gapping constructions in Universal Dependencies v2. In: Proceedings of the NoDaLiDa 2017 Workshop on Universal Dependencies (UDW 2017), pp. 123-132 (2017).
35. Johnson, M.: A simple pattern-matching algorithm for recovering empty nodes and their antecedents. In: Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL 2002), 136-143 (2002).
36. Needleman, S. B. and Wunsch, C. D.: A general method applicable to the search for similarities in the amino acid sequence of two proteins. Journal of Molecular Biology 48(3), 443-453 (1970).