Information Modelling and Knowledge Bases XXXIII
M. Tropmann-Frick et al. (Eds.)

doi:10.3233/FAIA210485

Conceptual Framework for the Conversion of Text Document into TIL-Script

Martina ČÍHALOVÁ[a] and Marek MENŠÍK[b]
[a] *Palacký University Olomouc, Department of Philosophy, Czech Republic*
[b] *VSB – Technical University of Ostrava, Department of Computer Science, Czech Republic*

Abstract. The paper deals with the introduction of TILUS tool for the needs of appropriate textual information sources retrieval and natural language processing. TILUS tool presupposed up to now that all the data are formalized in TIL-Script, the computational variant of Transparent Intensional Logic (TIL). We outline the general proposal of utilizing the Stanford typed dependencies representation for semi-automate conversion of natural language into TIL-Script. In order to be able to correctly solve this problem, we also introduce our universal conceptualization which is able to cover the thematic variations of processed texts.

Keywords. Information sources, Stanford typed dependencies, Transparent Intensional Logic, TIL-Script, Ontology

1. Introduction

In this paper we briefly introduce the current state of development of TILUS tool for the needs of appropriate textual information sources retrieval and natural language processing. The paper is organized as follows. In the section 2, there is a brief introduction of TILUS and his functions. TILUS presuppose that all the data are previously formalized in TIL-Script, the computational variant of Transparent Intensional Logic (TIL). However, this automated conversion of natural language texts into TIL-Script was not incorporated within TILUS. The aim of this paper is to outline the basis of the strategy to automate or at least semi-automate the conversion of natural language into TIL-Script. In order to be able to correctly solve this problem, we introduce our general conceptualization which is able to cover the thematic variations of processed texts in the section 3. Finally, we outline how to utilize the Stanford dependencies relations for automate conversion of natural language into TIL-Script in section 4. Concluding remarks on further research can be found in Section 5.

2. The brief introduction of TILUS for appropriate textual information sources retrieval

We are developing the TILUS tool for the needs of appropriate textual information sources retrieval and natural language processing. This tool has many modules that can be divided into two main parts. The first part is the module of logical deduction in order to be able to infer the conclusions of premises and verify the validity of arguments. The inferential mechanism is based on natural deduction. This part was introduced in [1], [2],

[3]. The second part concerns machine learning and searching for relevant text sources which was introduces in [4], [5], [6], [7]. Now our system presuppose that all the inputs are transformed into TIL-Script. However, a module for the automatic transformation of natural language texts into TIL-Script is not developed yet.

We vote for TIL-Script, the computational variant of Transparent Intensional logic (TIL), because its procedural semantics is close to natural language and has the great expressive power. TIL was developed by Pavel Tichý who introduced its main principles in [8]. The complete works from Tichý can be found in [9]. TIL has further been developed for instance by M. Duží, P. Materna, B. Jespersen, J. Raclavský and a great deal of the TIL contemporary research can be found in [10]. The outcomes of this book are especially relevant to the inter alia conceptual modelling area. TIL-Script as well as TIL is hyperintensional, typed, partial specification language with ramified hierarchy of types.

The aim of this paper is not to introduce the current system in details, but to outline a solution for the conversion of texts in natural language into TIL-Script. However, in order to give the reader an overall view of the current state, we firstly summarize the functionality of the individual modules of the TILUS tool in brief.

The deduction module is a module for the conclusion inference of premises and for the verifying the validity of arguments using natural deduction. Due to the fact that TIL-Script is a robust system with great expressive power, our deductive system had to been modified and new deduction rules had to been introduced, such as beta-conversion rule, the rules for working with lambda quantifiers, and so on. The specification of the deduction rules extension can be found in [1].

In the module for 'machine learning and text source recommendation', which is crutial for the proposals of solutions in this paper, we deal with how to generate an *explication* in TIL-Script of an atomic concept. The term *explication* can be understood as a brief description of the sought term in a particular source text. We generate an explication from each source text and present it to the user of our application. The user then selects a particular explication based on his preferences and our system then recommends him or her other possible relevant documents.

The theoretical framework for the recommandation module design is the method of association rules and on the FCA theory, where the relevant documents are ordered according to their significance to the user, for more see [11]. First, we need to analyse textual natural language sources to obtain their logical formalization in TIL-Script constructions. The conversion from natural language to the TIL-Script is a rather demanding process and requires logical-linguistic analysis. Hence, all our contributions published up to now presumed the cooperation with the Department of Computational Linguistics to obtain the conversion of text into TIL-Script constructions. The set of relevant propositional constructions is then selected from the formalized set; namely those where the concept to be explicated occurs. The scheme of preparation of the text source is depicted in the fig. 1.

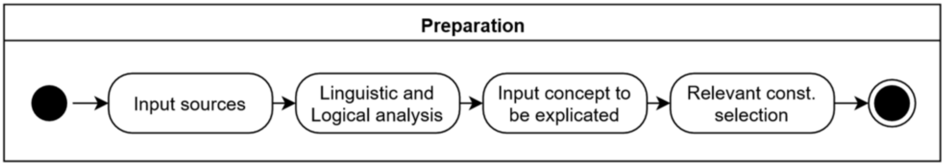

Figure 1. Pre-processing and formalisation of textual resources.

Next, the set of the selected constructions serves as an input for machine learning techniques, in particular, the *Inductive Heuristics module*, for more see [4]. This module produces hypothetic molecular concepts that should explicate the simple concept to be learned. Since there are several resource texts documents from which the hypothetic concepts are extracted, we obtain several explications of the learned simple concept. The inductive heuristics module is schematically specified by the following Fig. 2.

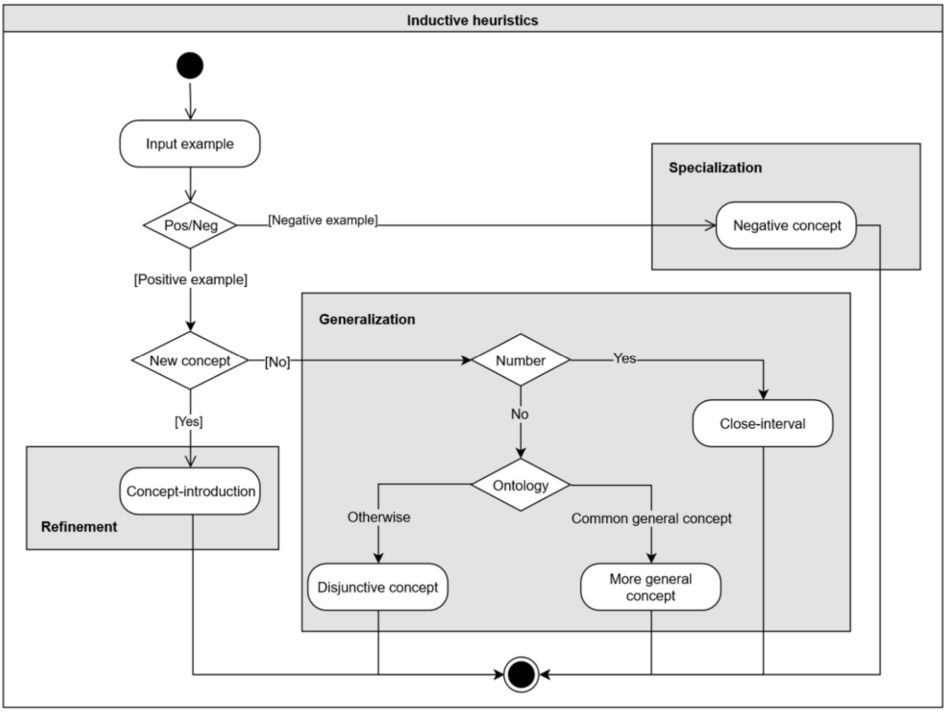

Figure 2. Inductive heuristics.

The last module of *Relevant Source Selection* is still work in progress. It is the module that deals with hypotheses processing and their evaluation. There are several functionalities that might be realised here. They include, inter alia, filtering out irrelevant sources according to the additional user-defined linguistic and logical criteria, search for inconsistencies among the hypotheses such as *contrarieties* and *contradictions*.

As already mentioned, the TILUS tool presuppose that all the data are formalized in TIL-Script, the computational variant of TIL. The next step is to automate or at least semi-automate the conversion of natural language into TIL-Script and this is as the crutial part of TILUS tool development. In order to be able to correctly solve this problem, we also introduce our general conceptualization which is able to cover the thematical variations of processed texts in the following section.

3. Basic ontological types and their logical types

In this section, a general conceptual framework of our tool is proposed.[1] This conceptual framework was designed as general and basically universal ontology close to natural language so that it could also be successfully utilized for the appropriate textual information sources retrieval.

The starting point of building a general conceptual framework is to distinguish between 'static' objects (static entities) in domain of interest like assigning properties of individuals, and 'dynamic' entities such as *activities* which are detected by some special type of verbs. The proposed analysis makes use of Tichý's formulation where such verbs are called *episodic verbs*. Tichý [13, pp. 263–296] draws a distinction between *episodic* and *attributive* verbs. Episodic verbs (e.g. *drive*, *tell*, etc.) express the actions of objects or people as opposed to attributive verbs (e.g. *is heavy*, *looks speedy*) that ascribe some empirical properties. Both static and dynamic entities are characterized by their further specification. At first we will focus on analysis of activities and their characteristics in more details.

Our analysis of activities proceeds from John Sowa's linguistically oriented approach to logical specification of natural language sentences. We focus on his analysis of sentences containing verbs, specifically episodic verbs according to Tichý's formulation. This analysis is based on the linguistic theory of verb valency frames.

The semantics of the respective verb is provided via its valency frame. In general, verb valency is the ability of a verb to bind other formal units, i.e. words, which cooperate to provide its meaning completely. These units are so-called *functors* or *participants* or *thematic roles* (according to Sowa's approach). Thus the valency of a verb determines the number of arguments (participants) controlled by a verbal predicate. This ability of verbs results from their meaning rather than from formal aspects. Despite this fact, lexical and grammatical valency can be distinguished. Grammatical valency contains information about the formal aspects of a verb, such as the grammatical case, and lexical valency contains information about the semantic character of particular participants. Grammatical valency depends on the language that is being analysed. Lexical valency, in contrast, is language independent, since it is established semantically. More details on the theory of valency frames can be found, for instance, in [14]. Attention to lexical valency has to therefore be paid in order to build up ontology independent of a used language. There are several types of valency. An impersonal (avalent) verb has no subject or a dummy subject. "It rains." is a typical example. Here the grammatical subject 'it' is just a dummy subject because it does not refer to any concrete object. An intransitive (monovalent) verb has just one argument, the subject S; "John (S) is singing." A transitive (divalent) verb has two arguments, an agent (A) and a patient (P), as in "John (A) kicked the ball (P)." A ditransitive verb has three arguments, and so on.[2] Consider an example of the sentence *John is going to Ostrava by train*. The activity *is going* has the following participants: actor (*John*), instrument of transport (*train*) and destination (*Ostrava*).

Verb valency frames also determine the obligatory and facultative arguments of a given verb, together with their types. Facultative arguments can be missing, of course. One might consider, for example, the verb *chastise*. This verb has two obligatory

---

[1] This conceptual framework is also introduced in [12] as general conceptual framework for logical classification of Wh-questions and possible answers to such questions in multi agent system.

[2] For details, see [15].

participants *who* (agent) and *whom* (patient). In addition, this verb can be connected with other facultative participants which express inter alia locality and time such as in the following sentence: A teacher chastises a student in the school early in the morning. It would be useful to classify verb participants into types according to their semantics. There are many classifications, however, of the participant types described in the literature. Three approaches to verb valency classification, according to the two valency dictionaries for the Czech language VALLEX and VerbaLex[3] and John Sowa's approach, are briefly compared in [18].[4]

Sowa uses the term *thematic roles* for the verb valency participants. His summary of all the types of thematic roles can be found in [20, pp. 506-510] or in his web source *Thematic roles* [21]. Sowa developed the system of conceptual graphs that are specified in [22, p. 187] as the system of logic for representing natural language semantics. Unlike predicate calculus, which was designed for studies in the foundations of mathematics, conceptual graphs were designed to simplify the mapping to and from natural language. They are based on a graph notation for logic first developed by the philosopher and logician Charles Sanders Peirce. Conceptual graph is represented as a labelled bipartite graph. Nodes in the set *C* are called concepts and nodes in the set *R* are called conceptual relations. Thematic roles are represented by conceptual relations that link the concept of a verb to the concepts of the participants in the occurrent expressed by the verb. Besides the graph notation, there is an equivalent linear notation for conceptual graphs where boxes are represented by square brackets, and the circles are represented by parentheses. Sowa distinguishes several types of thematic roles, for instance:

- *Agent* as an active animate entity that voluntarily initiates an action, example: *Eve bit an apple*:
  [Person: Eve] ← (Agnt) ← [Bite] → (Ptnt) → [Apple],
- *Theme* as an essential participant that may be moved, said or experienced, but is not structurally changed, example: *Billy likes the Beer*:
  [Person: Billy] ← (Expr) ← [Like] → (Thme) → [Beer: #],
- *Beneficiary* as a recipient that derives a benefit from the successful completion of event, example: *Diamonds were given to Ruby*:
  [Diamond: {*}] ← (Thme) ← [Give] → (Benf) → [Person: Ruby],
- *Destination* as a goal of a spatial process, example: *Bob went to Danbury*:
  [Person: Bob] ← (Agnt) ← [Go] → (Dest) → [City: Danbury],
- *Instrument* as a resource that is not changed by an event, example: *The key opened the door*:
  [Key: #] ← (Inst) ← [Open] → (Thme) →[Door: #],
- *Location* as an essential participant of a spatial nexus, example *Vehicles arrive at a station*:
  [Vehicle: {*}] ← (Thme) ← [Arrive] → (Loc) → [Station],
- *Patient* as an essential participant that undergoes some structural change as a result of event, example: *The cat swallowed a canary*:
  [Cat: #] ← (Agnt) ← [Swallow] → (Ptnt) → [Canary: #],
- and so on. For details, see [20, pp. 508-510].

[3] See, for instance [16] and [17].
[4] A very detailed comparison of these three classifications was provided in [19].

From the logical point of view, we deal with the verb phrases as denoting a function that is applied to its arguments. The number of arguments is controlled by the verb valency content. Dynamic entities relating to activities can be characterised by the special relationships in intensions between activities and their participants modelled as functions of TIL-Script. In our background theory (TIL) we view α-intensions as functions mapping possible worlds (of type ω) to a type β. The type β is frequently the type of a chronology of the elements of type α. These α-chronologies are, in turn, functions mapping time (of type τ) to the type α. Thus, α-intensions are usually mappings of type $(\omega \rightarrow (\tau \rightarrow \alpha))$, or in TIL notation $((\alpha\tau)\omega)$, $\alpha_{\tau\omega}$ for short and in TIL-Script notation ((α Time)World)). Another frequent type of intensions is the *property of individuals*, an object of type $(o\iota)_{\tau\omega}$, in TIL-Script notation (((Bool Indiv)Time) World). Consider for example the above mentioned sentence *John is going to Ostrava by train*. The specification of this sentence in TIL-Script is the following one:[5]

\w\t [‘And [[‘Agent@wt ‘John ‘Is_going][[‘And [‘Destination@wt ‘Ostrava ‘Is_going] [‘instrument@wt ‘train ‘is_going]]][6]

The types are the following:

- Agent, Destination / (((Bool Indiv (((Bool Indiv)Time)World))Time)World)
- Instrument / (Bool (((((Bool Indiv)Time)World) (((Bool Indiv)Time)World))Time)World)
- Ostrava, John / Indiv
- Is_going, train / (((Bool Indiv)Time)World)

Now let’s return to the conceptual characteristics of the ‘static’ entities. Static entities can be characterised by their properties and attributes. From the linguistic point of view, the properties assigned to them are usually denoted by a copular verb + adjective. Hence, the typical form of a sentence characterizing a static object is “S Cv Adj”, where S is a subject, Cv a copular verb and Adj an adjective. Typical copular verbs are *is*, *am*, *are*, …, *appear*, *seem*, *look*, *sound*, *smell*, *taste*, *feel*, *become* and *get*.

In the conceptual analysis of a given domain, it is useful to distinguish between two basic classes of characteristics of static objects. The former are relatively stable characteristics of objects and they remain unchanged over the time of objects’ existence, the second are dynamic empirical facts about these objects. The former can be called ‘substantive’ properties and the latter ‘accidental’ properties.

For instance, according to the laws of physics and biology, if an individual is born as a person, then during its life-time span it cannot become, say, a dog or a vase. Hence, being a person is a *substantive* property of such an individual. On the other hand the property of being a student is possessed by individual only contingently and the individual may acquire or lose this property depending on time. *Substantive properties* are those that individuals have nomically necessarily and are stable during the existence of respective individual in the respective possible world, while *accidental properties* are possessed by individuals purely contingently and can change during the time in the respective possible world.

One might consider traffic domain, for example, as the domain of interest. Then the substantial property is that of being a car with its subsumed properties such as the type

[5] Note the specific notation character @wt. This character expresses an intensional descent to the respective possible world and time to obtain the final value of this intension in this possible world and time.

[6] “\w\t” represents lambda abstraction over variables of possible worlds and times.

of car, namely a Sedan, Minivan, Hatchback, etc. According, of course, to the ISA hierarchies, each car is the vehicle. So all the properties connected via ISA hierarchy with the respective substantive property are substantive properties as well. The accidental properties of an individual car include, inter alia, the speed limit, weight, colour, date of manufacture, etc. In addition, each substantive property is mostly associated with certain accidental attributes of the individuals representing the given substantive property. The substantive property of car is associated, for example, with the above accidental properties, and also its owner, etc.

Substantive properties are from the logical point of view the functions which input is the respective possible world and individual and the output is the truth value according to the fact if the respective individual is/is not the owner of this property. The logical type of accidental properties is the function from possible worlds to time and the output is also in many cases the truth value (for instance the properties of being a student, employee and so on). However, some accidental properties such as the weight, speed limit, etc. are the functions with a possible world, time and individual as an input and the output is the number or magnitude.

This conceptual framework is within its universality able to cover the thematic variations of processed texts. But how to achieve the automated or at least semi-automated conversion of the natural language text into TIL-Script? In the following section we examine the possibility to utilize Stanford typed dependencies representation for this purpose.

4. The utilization of Stanford typed dependencies representation for the conversion into TIL-Script

Universal Dependencies (UD) is a framework for consistent annotation of parts of natural language texts across different human languages. It was designed to provide a straightforward description of grammatical relations for any user who could benefit from automatic text understanding. UD is an open community effort with over 300 contributors producing nearly 200 treebanks in over 100 languages. It is a project which is developing cross-linguistically consistent treebank annotation for many languages, with the goal of facilitating multilingual parser development, cross-lingual learning, and parsing research from a language typology perspective.[7] The annotation scheme is based on an evolution of (universal) Stanford dependencies see [23], Google universal part-of-speech tags, see [24], and the Interset interlingua for morphosyntactic tagsets, see [25]. The theoretical framework for the types of syntactical relations are Stanford dependencies which map straightforwardly onto a directed graph representation, in which words in the sentence are nodes in the graph and grammatical relations are edge labels. This representation contains approximately 50 grammatical relations and in [26], you can find the list of all the types of them with their specifications and natural language examples. There are some types of grammatical relations with examples below:

- *acomp*: adjectival complement. An adjectival complement of a verb is an adjectival phrase which functions as the complement (like an object of the verb). Example: *She looks very beautiful*, where the term *beautiful* is related with the term *looks*.

[7] For more see the web source https://universaldependencies.org/

- *advmod:* adverb modifier. An adverb modifier of a word is a (non-clausal) adverb or adverb-headed phrase that serves to modify the meaning of the word. Examples: *Genetically modified food*: advmod(modified, genetically), *less often*: advmod(often, less).
- *attr* is a relation intended for the complement of a copular verb such as "to be", "to seem", "to appear".
- *nsubj*: nominal subject. A nominal subject is a noun phrase which is the syntactic subject of a clause. The governor of this relation might not always be a verb: when the verb is a copular verb, the root of the clause is the complement of the copular verb, which can be an adjective or noun. Example: *Clinton defeated Dole*: nsubj(defeated, Clinton), *The baby is cute*: nsubj(cute, baby).
- *num*: numeric modifier. A numeric modifier of a noun is any number phrase that serves to modify the meaning of the noun with a quantity. Examples: *Sam ate 3 sheep*: num(sheep, 3), *Sam spent forty dollars*: num(dollars, 40).
- *agent*: agent. An agent is the complement of a passive verb which is introduced by the preposition "by" and does the action. This relation only appears in the collapsed dependencies, where it can replace prep by, where appropriate. It does not appear in basic dependencies output. Examples: *The man has been killed by the police*: agent(killed, police), *Effects caused by the protein are important*: agent(caused, protein).

These relations provide a straightforward description of grammatical relations in the analysed text and we will demonstrate that we can extract also some important semantic information from this Stanford typed dependencies representation. We decide to use the tool *Explosion* for the text annotation which is based right on Stanford dependencies. This is a software company specializing in developer tools for artificial intelligence and natural language processing (NLP). They are the makers of spaCy (the leading open-source library for advanced NLP) and Prodigy (an annotation tool for radically efficient machine teaching). The next step will be the precise specification of the rules for the exact syntactical relations conversion into logical types and TIL-Script. In the next paragraph there is the rough outline of how we are going to proceed.

In many cases, we are able to syntactically recognize the substantive properties and accidental properties. For the comparison, let's have the sentence *The cat is a fast mammal*. The relation *advmod* (adverb modifier) determines the accidental property and the *attr* relation determines the substantive property.

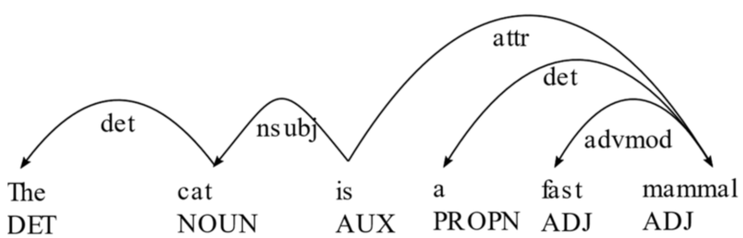

Figure 3. The output of *Explosion* for the sentence *The cat is a fast mammal*.

However, we are aware of some exceptions. For instance in the sentence *Peter is a student* an *attr* relation determines the *accidental* property. For these reasons we have to build *substantive* properties into ontology and provide respective rules for these exceptions.

In many cases, we are also able to detect the activity and its participants in general on the base of Stanford typed dependencies relations. Let's have the above mentioned example of the sentence *John goes to Ostrava by train*. The verb valency participants of the activity *going* are the following: *John* is the agent of the activity, *Ostrava* is the final destination and *train* is the instrument of transport. For the comparison, on the Fig. 4 is the respective output of this sentence from *Explosion*.

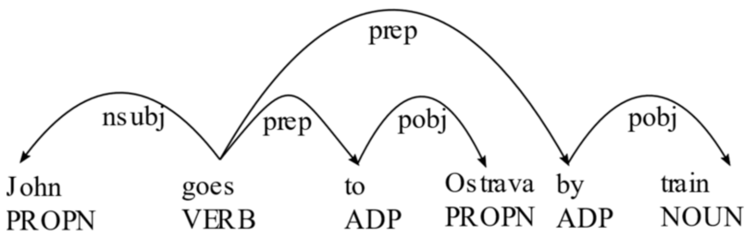

Figure 4. The output of *Explosion* for the sentence *John goes to Ostrava by train*.

We are able to recognize the activity determined by the verb and the agent of this activity which is determined by the proper name *John* and is in the relation *nsubj* with the verb. The terms *Ostrava* and *train* are detected as consecutive relations *prep* (preposition) and *pobj* (object of preposition). Hence, we can recognize in an automated way via analysis of dependencies that these terms are the verb valency participants in general, but not the respective categories of these participants (destination, instrument). For this purpose, we are considering the possibilities of utilization of the verb valency dictionaries. Another possibility could be the querying the user to fill in the particular type of participant manually.

5. Conclusion and future work

The aim of this paper was to outline the basis of the strategy to automate or at least semi-automate the conversion of natural language into TIL-Script, the computational variant of TIL. This strategy was outlined for the currently developed TILUS tool for the needs of appropriate textual information sources retrieval and natural language processing. To grasp this goal, we need to map the concepts in natural language into the respective logical types at first. We introduce for these purposes the general conceptualization which is able to cover the thematical variations of processed texts and the respective logical types of the main concepts. We also outline how the Stanford typed dependencies representation can be utilized for semi-automated conversion of natural language into TIL-Script. We are currently working on the precise specification of the rules for the conversion of outputs from *Explosion* to TIL-Script.

Acknowledgements

The work on this paper was supported by the project 'JG_2020_005 Times, events, and logical specification' of Palacký University and Grant of SGS No. SP2021/87 and also by the projekt VSB – Technical University of Ostrava, Czech Republic, 'Application of Formal Methods in Knowledge Modelling and Software Engineering IV'.

References

[1] Duží, M., Menšík, M. Inferring knowledge from textual data by natural deduction. *Computacion y Sistemas*, 2020, 24(1): 29-48. ISSN 2007-9737. DOI: 10.13053/CyS-24-1-3345

[2] Duží, M., Menšík, M., Pajr, M., Patschka, V. Natural deduction system in the TIL-script language. In: Endrjukaite T.,Jaakkola H.,Dudko A.,Kiyoki Y.,Thalheim B.,Yoshida N., editors. *Frontiers in Artificial Intelligence and Applications*, vol. 312; 2019, Amsterdam: IOS Press, p. 237-255. DOI: 10.3233/978-1-61499-933-1-237

[3] Duží, M., Menšík, M. Vích, L. Deduction system for TIL-2010. In: Horák, A., Rychlý, P., editors. *6th Workshop on Recent Advances in Slavonic Natural Language Processing, RASLAN 2012*; 2012, Brno: Masaryk University, p. 49-53. ISBN: 978-802630313-8.

[4] Menšík, M., Duží, M., Albert, A., Patschka, V., Pajr, M. Refining concepts by machine learning. *Computación y Sistemas*, 2019, (23)3: 943–958. DOI: 10.13053/CyS-23-3-3242

[5] Menšík, M., Duží, M., Albert, A., Patschka, V., Pajr, M. Machine Learning Using TIL. In: A. Dahanayake, J. Huiskonen, Y. Kiyoki, B. Thalheim, H. Jaakkola, N. Yoshida, editors. *Frontiers in Artificial Intelligence and Applications*, vol. 321; 2019, Amsterdam: IOS Press, p. 344-362. DOI: 10.3233/FAIA200024

[6] Menšík, M., Duží, M., Albert, A., Patschka, V., Pajr, M. Seeking relevant information sources. In: *2019 IEEE 15th International Scientific Conference on Informatics*, 2019, Poprad (Slovakia): IEEE, p. 271-276. DOI: 10.1109/Informatics47936.2019.9119332

[7] Albert, A., Duží, M., Menšík, M., Pajr, M., Patschka, V. Search for Appropriate Textual Information Sources. In: B. Thalheim, M. Tropmann-Frick, H. Jaakkola, N. Yoshida, Y. Kiyoki, editors. *Frontiers in Artificial Intelligence and Applications*, vol. 333; 2021, p. 227-246, Amsterdam: IOS Press. DOI: 10.3233/FAIA200832

[8] Tichý, P. *The Foundations of Frege's Logic*. 1988, Berlin, New York: De Gruyter.

[9] Tichý, P. *Collected Papers in Logic and Philosophy*. 2004, V. Svoboda, B. Jespersen (eds.).

[10] Duží, M., Jespersen, B. and Materna, P. *Procedural Semantics for Hyperintensional Logic. Foundations and Applications of Transparent Intensional Logic.* 2010, Berlin, Springer, series Logic, Epistemology, and the Unity of Science, vol. 17.

[11] Menšík, M., Albert, A., Patschka, V. Using FCA for Seeking Relevant Information Source. In *RASLAN 2020*; 2020, p. 47-54, Brno: Tribun EU. ISBN 978-80-263-1600-8, ISSN 2336-4289.

[12] Číhalová, M., Duží, M. Modelling dynamic behaviour of agents in a multi-agent world; logical analysis of *Wh*-questions and answers. Submitted to the *Logic Journal of the IGPL*.

[13] Tichý, P. The semantics of episodic verbs. *Theoretical Linguistics*, vol. 7, 1980; p. 263-296. Reprinted in (Tichý 2004: 411–446).

[14] Fischer, K., Ágel, V. Dependency grammar and valency theory. In: *The Oxford handbook of linguistic analysis*; 2010, p. 223-255.

[15] Dixon, R. M. W. A Typology of Causatives: Form, Syntax, and Meaning. In: R. M. W. Dixon, A. Y. Aikhenvald, editors. *Changing Valency: Case Studies in* Transitivity, 2000, p. 30-41, New York, NY: Cambridge University Press.

[16] Lopatková, M., Žabokrtský, Z., Kettnerová, V. *Valenční slovník českých sloves*. 2008, Praha: Karolinum. ISBN 978-80-246-1467-0.

[17] Hlaváčková, D., Horák, A. VerbaLex - New Comprehensive Lexicon of Verb Valencies for Czech. In: *Computer Treatment of Slavic and East European Languages*; 2006, p. 107-115, Bratislava, Slovakia: Slovenský národný korpus.

[18] Číhalová, M. Event ontology specification based on the theory of valency frames. In: T. Welzer, H. Jaakkola, B. Thalheim, Y. Kiyoki and N. Yoshida, editors. *Frontiers in Artificial Intelligence and Applications*; 2016, p. 299-313, Amsterdam, Berlin, Tokyo, Washington DC: IOS Press.

[19] Číhalová, M. *Jazyky pro tvorbu ontologií* (*Languages for ontology building*). 2011, Ph.D. Thesis, VŠB-Technical University of Ostrava, Ostrava, The Czech Republic.

[20] Sowa, J. F. Knowledge representation (logical, philosophical, and computational foundations). 2000, Pacific Grove, CA: Brooks Cole Publishing Co.

[21] Sowa, J. F. *Thematic roles*. http://www.jfsowa.com/ontology/roles.htm. Accessed October 1, 2021.

[22] Sowa, J. F. Towards the expressive power of natural language. In: J. F. Sowa, editor. *Principles of semantic networks. Explorations in the representation of knowledge,* 1991, p. 157-189, San Mateo, CA: Morgan Kaufmann. DOI: org/10.1016/C2013-0-08297-7

[23] De Marneffe M., C., Manning, Ch. D. *Stanford typed dependencies manual*. 2021. https://nlp.stanford.edu/software/dependencies_manual.pdf. Accessed January 1, 2021.

[24] Petrov, S., Das, D., McDonald, R. A universal part-of-speech tagset. In: *Proceedings of the 8th International Conference on Language Resources and Evaluation*, *LREC 2012*; 2012, p. 2089-2096, Istanbul, Turkey: European Language Resources Association (ELRA). ISBN 978-295174087-7.

[25] Zeman, D. Reusable Tagset Conversion Using Tagset Drivers. In: *Proceedings of the 6th International Conference on Language Resources and Evaluation*, *LREC 2008*; 2008, p. 213-218, Marrakech, Morocco: European Language Resources Association (ELRA). ISBN 978-295174084-6.

[26] de Marneffe, M. C., Dozat, T., Silveira, N., Haverinen, K., Ginter, F., Nivre, J. Manning, Ch. D. Universal Stanford Dependencies: A cross-linguistic typology. In: *Proceedings of the 9th International Conference on Language Resources and Evaluation*, *LREC 2014*; 2014, p. 4585-4592, Reykjavik, Iceland: European Language Resources Association (ELRA). ISBN 978-295174088-4.