From the Prague Dependency Treebank to the Uniform Meaning Representation: Gold-Standard Czech UMR Data and Partial Automatic Conversion

Markéta Lopatková*, Hana Hledíková, Jan Štěpánek and Daniel Zeman

Charles University, Faculty of Mathematics and Physics, Institute of Formal and Applied Linguistics, Malostranské náměstí 25, Prague, Czechia

Abstract

Uniform Meaning Representation (UMR) is a semantic framework designed to capture the meaning of texts in a structured and interpretable manner. In this paper, we present the Czech gold-standard UMR data and analyze the inter-annotator agreement on a sample annotated in parallel by two human annotators. Instances of disagreement are identified, the main sources of ambiguity are highlighted, and potential resolution strategies are discussed. Furthermore, we briefly describe the main principles of the automatic conversion procedure that maps data from the Prague Dependency Treebank (PDT-C) into the UMR framework. We illustrate the interaction of multiple linguistic phenomena, which contributes to the overall complexity of the (still partial) conversion process. Finally, we quantitatively evaluate the output of the conversion system against the gold-standard data.

Keywords

PDT, UMR, gold-standard UMR data for Czech, partial automatic conversion, quantitative evaluation

1. Motivation and Goals

In our previous work [1], we explored and compared two different approaches to semantic representation: the Uniform Meaning Representation (UMR) and the meaning representation used in the Prague Dependency Treebank (PDT). Focusing on how each of these approaches models sentence-level meaning, we examined structural aspects of sentence graphs as well as linguistic phenomena such as the event/entity distinction, verbal predicates and their argument structures, lexical and abstract predicates, and coreference chains. Special attention was paid to the challenges and strategies for automatically converting PDT-style annotations into the UMR format.

Despite the availability of extensive language resources for Czech, the findings highlighted the limited feasibility of a fully automatic translation between the two representations. We showed that an effective conversion requires careful heuristic design and significant manual annotation.

Looking ahead, we identified several key areas for further research, bearing in mind their materialization in a working automatic procedure:

- Transforming graph structures, which stems mainly from different treatment of coordination, coreference relations, relative clauses, and raising and control verbs.
- Changing node labels, reflecting the turn from PDT deep syntactic to UMR conceptual representation.
- Converting PDT deep syntactic roles into UMR semantic relations through the verb-specific mapping of arguments (whenever available) or through the default mapping table (for both arguments and adjuncts) when such specific mapping is not provided.

In addition, UMR nodes should be enriched with selected semantic attributes—including aspect, polarity, number, and person—using PDTgrammatemes (i.e., semantic morphological features), as discussed in more detail in [2].

An implementation of any PDT to UMR conversion procedure requires not only a thorough understanding of both underlying representation frameworks, but also a deep familiarity with the complex and richly structured data schemata (particularly that of PDT). Furthermore, an appropriate evaluation of the conversion output necessitates not only visual checking the outputs and their comparison against ad-hoc manually annotated Czech data (which is indispensable to refine the conversion of individual linguistic phenomena), but also the availability of gold-standard Czech UMR annotations. Only such data can serve as a reliable reference and show overall progress, bearing in mind complex and interlinked structures of a natural language.

The purpose of this paper is to introduce (a small portion of) the gold-standard Czech UMR annotations, together with an analysis of the inter-annotator agreement on a sample annotated in parallel by two human annotators. Instances of disagreement are identified, the main sources of ambiguity are highlighted, and potential resolution strategies are discussed (Sect. 3). Furthermore, we briefly describe the main principles of the automatic conversion procedure that maps the PDT data into the UMR framework. We illustrate the interaction of multiple linguistic phenomena, which contributes to the overall complexity of the (still partial) conversion process, and quantitatively evaluate the output of the newest version of the conversion against the gold-standard (Sect. 4). The statistics cited here are taken from [3].

2. Introducing PDT and UMR

PDT and UMR represent two distinct yet complementary approaches to meaning representation.

PDT. PDT[1] (namely its tectogrammatical layer) is a richly structured deep syntactic annotation scheme tailored to Czech, capturing the underlying predicate-argument structure through a dependency tree with labeled functors. Morphosyntactic and semantic features, including tense, aspect, and modality, are encoded as grammatemes, offering fine-grained linguistic insight specific to the inflectional nature of Czech [4, 5, 6, 7].[2] In particular, the PDT annotation reflects linguistically structured meaning, i.e., its deep syntactic structures more-or-less directly refer to the annotated text, and as such, it is less abstract than UMR.

UMR. UMR[3] is a graph-based semantically grounded framework designed for cross-linguistic applicability, abstracting away from surface syntax to encode concepts (entities and events represented as graph nodes), their relations (graph edges) and attributes through a normalized, language-independent format [10, 11, 12]. In particular, all syntactic variants of a statement are represented uniformly (contrary to PDT). However, at the same time, it allows more alternative annotations. This feature is challenging especially from an evaluation point of view, as it artificially deteriorates the resulting figures.

3. Towards Czech Gold-Standard UMR Data

The PDT-C corpus contains a substantial amount of Czech data in a variety of genres.[4] As a basis for gold-standard data, we selected a sample of six files from its development set for manual annotation. This sample represents the main (coarse-grained) genres stored in PDT-C: both texts (especially general journalistic and technical styles) and spoken language, both original and translated. Further, the selected files contain predefined linguistic phenomena that are likely to present challenges during conversion—such as implicit events, not overtly expressed concepts (entities, events), coordinated structures (esp. those with common modifiers), coreference chains, relative clauses, negation, particular functors, and discourse relations. We also ensured that these files do not contain (large) tables or similar structured texts, as these

---

[1] https://ufal.mff.cuni.cz/pdt-c

[2] However, extensive PDT-like resources for other languages, such as Latin [8] and English [9], prove that its applicability is not limited to Czech.

[3] https://umr4nlp.github.io/web/

[4] The latest version of the data, PDT-C 2.0, is available through the Lindat repository, http://hdl.handle.net/11234/1-5813.

Table 1
Statistics for manually annotated data ('Annot1 / Annot2' indicates results for 2 annotators).

Gold-standard data: (sub)corpus	sentences	tokens	tokens per sentence	PDT-C nodes	UMR nodes	UMR per PDT-C
PDT	25	467	18.7	378	375	0.99
PDTSC	50	374	7.5	321	442	1.38
PCEDT	16	474	29.6	400	307	0.77
total	91	1315	14.5	1099	1124	1.02

Parallel annotations: (sub)corpus	sentences	tokens	tokens per sentence	PDT-C nodes	UMR nodes Annot1 / Annot2	UMR per PDT-C (avg.)
PDT	11	192	17.5	151	153 / 150	1.00
PDTSC	10	63	6.3	68	75 / 71	1.26
total	21	255	12.1	209	228 / 221	1.07

pose specific challenges; in the case of PDT and PCEDT subcorpora, the lengths of the documents were also considered (with preferences for shorter documents). The selected data set includes:[5]

- Two complete documents from the core PDT subcorpus, consisting of Czech newspaper texts from 1992–1994 (11 + 14 sentences);
- Two files from the PDTSC subcorpus that contains spontaneous dialogues; 25 sentences from each file were annotated;[6]
- Initial parts of two documents from the Czech portion of the PCEDT subcorpus, comprising translations of the Penn Treebank (Wall Street Journal texts, all translated from English by professional translators); this subcorpus contains mostly business and finance news (6 + 10 sentences).

Basic data statistics. For basic statistics, see the upper part of Table 1.

The table demonstrates that the genres represented in individual subcorpora of PDT-C differ significantly in their basic characteristics, such as sentence length. The shortest sentences are found in spontaneous dialogues in PDTSC, while written newspaper texts from PDT exhibit sentences that are, in the selected samples, 2.5 to 2.7 times longer. The most complex sentences occur in translations from the PCEDT, where sentence lengths (measured in tokens) are approximately four times greater than in the spoken data.

Furthermore, although on average PDT and UMR represent data using graphs with a comparable number of nodes, the individual subcorpora again show substantial differences. In annotating dialogues from PDTSC, annotators added higher-level graph structures to represent individual speakers, resulting in a higher number of UMR nodes than PDT-C nodes (in PDTSC, this information is included in the metadata). Conversely, the PCEDT, due to its focus on finance and economics, contains a large number of company names. These are represented in the original data as entire subtrees (with nodes for individual tokens) but are merged into single UMR nodes. As a result, the UMR structures contain 23% fewer nodes.

Parallel annotated data. A subset of these data (Table 1, lower part) was annotated in parallel by two annotators with deep knowledge of the PDT framework and trained to understand the UMR principles. Their annotations were then carefully compared—differences were thoroughly discussed, oversights corrected, and (some) challenging cases resolved. This reconciliation phase aimed to ensure a consistent

---

[5]The Czech UMR data described and compared in the paper (both the manual UMRs and the automatically converted structures) are available through the Lindat repository, see http://hdl.handle.net/11234/1-5951.

[6]PDTSC files contain 50 sentences each, they typically include several short dialogues (but a dialogue can be split into more files).

Table 2
Quantitative comparison of manually double-annotated UMRs (PDT+PDTSC); the table is taken from [3]. (* Unmapped nodes are ignored. ** Unmapped nodes all counted as incorrect.)

UMR node mapping:					F_1
Annot1 nodes	Annot2 nodes	mapped	recall	precision	
228	221	215	94%	97%	96%

Concept and relation comparison (only mapped nodes):*					F_1
Annot1 triples	Annot2 triples	match	recall	precision	
633	644	595	94%	92%	93%

Concept and relation comparison:**					$ju{:}mæʧ$ = F_1
Annot1 triples	Annot2 triples	match	recall	precision	
663	659	595	90%	90%	90%

interpretation of the UMR rules (which are often complex and are not always documented in sufficient detail).

A quantitative comparison of the parallel annotated data (in terms of inter-annotator agreement) is discussed in Sect. 3.1, a qualitative analysis of differences in Sect. 3.2

3.1. Quantitative Comparison

UMR graphs can be represented as a set of triples (x, y, z), where either x is a node, y a name of a relation (edge) and z is the respective child node, or x is a node, y its attribute and z a value of this attribute.

When comparing two graphs,[7] the corresponding nodes must first be identified. Following [3], the mapping algorithm $ju{:}mæʧ$ is used here. The algorithm primarily aligns nodes linked to the same word(s); for nodes without word alignment (representing esp. overtly unexpressed concepts that are restored in PDT and/or UMR graphs as nodes), it requires concept identity. The algorithm outputs a symmetric one-to-one mapping of nodes whenever possible (with some nodes left unmapped).

Finally, the similarity of the UMR graphs is measured as the F_1 score of these triples.

The quantitative comparison of the Czech parallel data is presented in Table 2. The figures ($ju{:}mæʧ$ = 90%, with 96% of nodes mapped) indicate that the reconciliation results in relatively high inter-annotator agreement. This success level can be seen as an upper bound for what the automatic conversion procedure can achieve.

3.2. Qualitative Analysis

Even after the reconciliation phase, the parallel annotations exhibit 10% differences in UMR triples, reflecting either legitimate and well-grounded variations in text interpretation or individual annotators' preferences in representing certain phenomena. This aligns with the UMR specification,[8] which—as repeatedly noted—permits multiple valid annotations of the same meaning.

An analysis of differences in manual annotations, despite being limited to the small data sample, reveals several linguistic phenomena whose representation tends to be inconsistent. These can be roughly classified into several larger groups: those related to events and their structure, ellipses, granularity of concept classification, and attributes. The aim of the analysis is to identify phenomena where clearer specifications could help reduce variability in annotations.

Events and argument structure. UMR conceptually distinguishes entities, states, and events, regardless of their surface (morphological) forms. However, the crucial **boundary between events and**

---

[7]Our comparison is limited to sentence-level UMR graphs as the scripts do not consider document-level triples.
[8]https://github.com/umr4nlp/umr-guidelines/blob/master/guidelines.md

non-events remains unclear, as already pinpointed in [1]. This ambiguity appears in parallel data as well, as exemplified in (1): While for one annotator, the concept *schůze* 'meeting' is still seen as "actional" enough to be considered an event (and therefore is annotated with the corresponding verb *sejít-se* '(to) meet' and event attributes), the other annotator sees this concept as an entity (and thus annotates the number attribute).

(1)   *Včera to připustil člen komise poslanec Pavel Severa … **po schůzi** orgánu.*
        'Yesterday, commission member MP Pavel Severa … admitted this after **a meeting** of the body.'

```
Annot1:                                    Annot2:
(s4s1 / sejít-se-001   `(to) meet'         (s4s2 / schůze   `meeting'
  :aspect performance                        :refer-number singular
  :modal-strength full-affirmative
  :ARG0 (s4o1 / orgán   `body'              :mod (s4o1 / orgán   `body'
      :refer-number singular))                  :refer-number singular))
```

Improvement possibility: We can apply a morphological criterion to determine which concepts should be treated as events (those morphologically derived from a verb). However, while this approach can improve an inter-annotator agreement, it represents a departure from the core UMR principles.

Even if both annotators agree that a particular concept in the given context should be considered an entity, they can differ in assigning **argument vs. non-argument roles**: one of them can gives it argument structure anyway, while the other limits the use of arguments to events and uses the non-argument roles for entities. This is exemplified in (2) with the *podnět* 'complaint' concept and its roles (ARG0, ARG1 vs. source, regard).

(2)   *Ačkoli … před týdnem ukončil vyšetřování **podnětů** ODA vůči kontrarozvědce …*
        'Although … (it) closed its investigation into the ODA's **complaints** against counterintelligence a week ago … '

```
Annot1:                                    Annot2:
(s3p3 / podnět   `complaint'               (s3p2 / podnět   `complaint'
  :refer-number plural                       :refer-number plural
  :ARG0 (s3o2 / organization                 :source (s3p5 / political-organization
      :wiki "Q1807830"                           :refer-number singular
      :name (s3n3 / name :op1 "ODA"))            :name (s3n2 / name :op1 "ODA"))
  :ARG1 (s3k2 / kontrarozvědka               :regard (s3k2 / kontrarozvědka
      :refer-number singular))                   :refer-number singular))
```

Improvement possibility: For entities denoted by words (morphologically) related to verbs, annotators should be instructed to consult the (PDT-C-related) valency lexicon of Czech verbs PDT-Vallex [13, 14] and adhere to the corresponding verbal argument structure whenever possible.

Another source of disagreement related to events comes from an **incomplete argument structure**. The UMR guidelines do not specify whether verbs' argument structure should be completed when its arguments are not expressed overtly in the sentence. Thus, one annotator may add the unexpressed argument (e.g., in (3), ARG0 of the verb *nachromovat* '(to) chrome' is identified as the abstract entity *person*), while the other may not add it.

(3)   *Nechal jsem si **nachromovat** … lampu, roury, teleskopy.* 'I had the lamp, pipes, and telescopes **chromed** …'

```
Annot1:                                          Annot2:
(s3n2 / nachromovat-001   `(to) chrom'          (s3n2 / nachromovat-001   `(to) chrom'
  :aspect performance                             :aspect performance
  :modal-strength full-affirmative                :modal-strength full-affirmative
                                                  :quote s3s1
  :affectee s3e1                                  :affectee s3p1
  :ARG0 (s3e3 / person
        :refer-person 3rd)
  :ARG1 (s3a1 / and                               :ARG1 (s3a1 / and
        :op1 (s3l1 / lampa   `lamp'                     :op1 (s3l1 / lampa   `lamp'
              :refer-number singular)                         :refer-number singular)
        :op2 (s3r1 / roura   `pipe'                     :op2 (s3r1 / roura   `pipe'
              :refer-number plural)                           :refer-number plural)
        :op3 (s3t1 / teleskop  `telescope'              :op3 (s3t1 / teleskop  `telescope'
              :refer-number plural)))                         :refer-number plural)))
```

Improvement possibility: Given the fact that the PDT-C annotation is supported by the PDT-Vallex valency lexicon [13, 14], annotators should be instructed to use the lexicon and complete the argument structure of verbs whenever relevant.

Ellipses. While the treatment of unexpressed arguments can be harmonized (see example (3) above), **ellipses** and their reconstruction represent a problem in general. For example, in (4), with *vydání* 'edition', one annotator may reconstruct the full structure and add the elided modifier from a previous context (*vydání novin* 'edition of newspapers'), while the other may not.

(4)   *Cena pátečního **vydání** … zůstává.*
        'The price of Friday's **edition** … of remains the same.'

```
Annot1:                                          Annot2:
(s5v1 / vydání   `edition'                       (s5v1 / vydání   `edition'
  :refer-number singular
  :mod (s5p1 / date-entity                         :mod (s5d1 / date-entity
       :weekday (s5p2 / pátek)) `Friday'                :weekday (s5p1 / pátek))) `Friday'
  :mod (s5n1 / noviny   `newspapers'
       :refer-number singular))
```

Improvement possibility: It is not possible to formulate exhaustive guidelines for when and how to reconstruct ellipses. The situation may improve partially once coreference relations are established at the document level, though even then, systematically verifying such reconstructions will remain formally complex.

Granularity of named entity classification. UMR uses a relatively rich **hierarchy of named entities (NE)**. However, it provides varying levels of granularity for different types of NEs, and these are not always clearly described or exemplified, making their use potentially ambiguous. For example, in (2), *ODA* is characterized as an organization (and further specified through its wikidata item) by one annotator, whereas the other annotator classifies it as a political organization, without anchoring it in the Wikipedia (thus, even with a finer level of the NE classification, the annotation is less specific).

Improvement possibility: Anchoring to a corresponding wikidata item wherever possible may help address this issue; however, formal inconsistencies in the data are likely to persist nonetheless.

Relations vs. attributes and their values. Relations between two concepts are represented as graph edges, both in PDT-C and UMR. In addition, UMR also employs attributes to characterize individual concepts. For instance, quantified entities such as *three dogs* are represented as a single node (here the concept *dog*) with the quant attribute specifying the quantity (here *three*). This approach offers a clear and efficient representation for numerical expressions.

However, quantity can also be expressed through quantifying operators such as *all*, *almost nothing*, or *several* (for Czech, e.g., *všechen*, *věškerý*, *téměř žádný*, *několik*). Since comprehensive inventories of quantifying expressions for Czech are lacking (and even existing annotations in English show inconsistency in this respect), annotators may adopt varying strategies, as illustrated in (5): while one annotator considers *veškerý* 'all' a concept (represented as a separate node, with quant relation), the other represents it as a quantifying operator (the quant attribute with value all).

(5)   *Stále prý jde o to, zda tajná služba **veškeré** údaje mohla získat z otevřených zdrojů.* 'The issue is still whether the secret service could have obtained **all** the data from open sources.'

```
Annot1:                              Annot2:
(s6u1 / údaj  `data'                 (s6u1 / údaj  `data'
  :quant (s6v1 / veškerý)) `all'        :quant all)
```

Improvement possibility: At least a tentative inventory of quantifying expressions would enhance inter-annotator agreement; nevertheless, no such list can be entirely comprehensive and would need to be continually expanded as additional data are processed.

Attributes and their annotation.   Another source of disagreement arises from the annotation of attributes. The annotators may either disagree on which attributes a given concept should bear, or they may agree on the presence of a specific attribute but diverge on its value. The former case is illustrated by examples (2) and (4) (in both cases, one of the annotators omitted the refer-number attribute, value singular). The latter case is exemplified in (6), where annotators disagreed on whether the event denoted by the verb *dokončit* 'complete' should be characterized as fully affirmed (the attribute modal-strength with value full-affirmative) or merely probable (value partial-affirmative).

(6)   *Komise se shodla na tom, že dokončí šetření, …*
        'The Commission agreed to complete investigations, … '

```
Annot1:                              Annot2:
(s5d1 / dokončit-001                 (s5d1 / dokončit-001
  :aspect performance                  :aspect performance
  :modal-strength partial-affirmative  :modal-strength full-affirmative
  :ARG0 ...                            :ARG0 ...
  :ARG1 ...)                           :ARG1 ...)
```

Improvement possibility: A general solution is difficult to define; however, data preprocessing and identifying expected attributes in advance may help, along with encouraging annotators to consistently include relevant attributes.

4. PDT to UMR Conversion

4.1. Conversion Principles

The conversion algorithm recursively traverses the PDT tree (specifically, its t-layer structure), and incrementally builds the corresponding UMR graph. During traversal, each node and edge are examined to identify and apply the necessary structural changes, relabeling operations, and insertion of UMR-specific attributes.[9]

Although the basic idea of conversion is conceptually straightforward, the handling of individual linguistic phenomena necessarily draws on various types of information provided in PDT-C. The conversion process accounts for the following:

- The original *syntactic structure*, including differences in the representation of coordination structures (see also [1, 2]) and named entity structures;
- The *lexical values* of individual nodes;

---

[9]The Czech UMR data described and compared in the paper (both the manual UMRs and the automatically converted structures) are available through the Lindat repository, see http://hdl.handle.net/11234/1-5951.
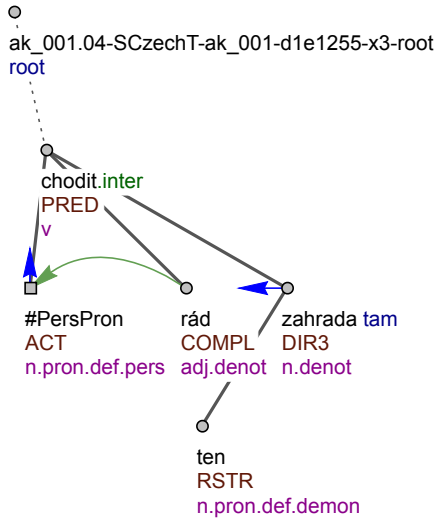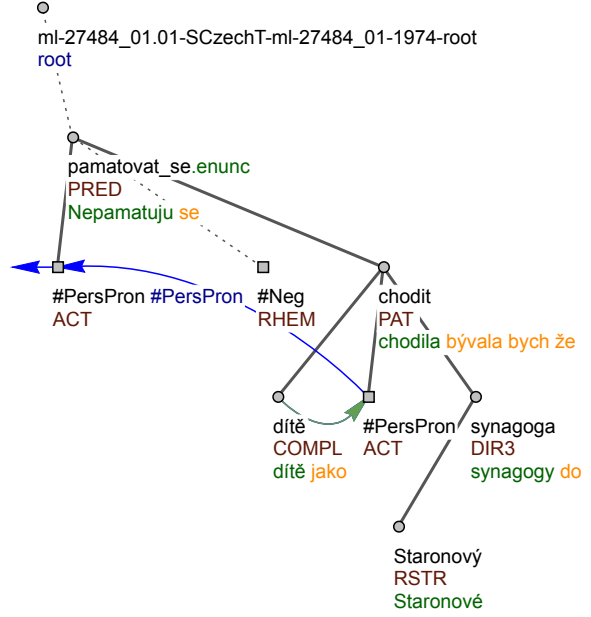
(a) COMPL, example (7).   (b) COMPL combined with coreference, example (8).

Figure 1: Verbal complement COMPL in PDT; green arrows represent the compl.rf links.

- The *semantics of morphological categories* (i.e., grammatemes), where available (fully provided only in the PDT subcorpus), otherwise relying on morphological features;
- The differing representation of *coreferential nodes*.

Moreover, these linguistic phenomena often interact, which further increases the complexity of the conversion process.

For example, let us see how the complement functor COMPL is converted. In accord with Czech syntactic tradition, a complement depends on two nodes: a predicate that's used as the complement's parent in PDT, and a noun with whom it agrees in gender, number, and case, represented in PDT by an arrow (a link of type compl.rf) (see Fig. 1a). The tree converted to UMR uses the relation manner for the complement (based on the deep syntactic part of speech, it could also be mod if the parent is a noun), the secondary relation is converted to a mod-of relation.

(7)   *Chodíte ráda do té zahrady?*
       'Do you like going to the garden?'

```
(s11c1 / chodit-006 `go'
     :ARG1 (s11e1 / entity)
     :manner (s11r1 / rád `glad'
        :mod-of s11e1)
     :goal (s11z1 / zahrada `garden'
        :mod (s11t1 / ten) `that'
        :refer-number singular)
     :aspect activity)
```

This seems rather straightforward, until we try to convert the whole data and notice that coreference interferes with the rule: The target of the secondary relation might have been removed from the UMR tree earlier in the conversion because it was an elided personal pronoun, see Fig. 1b and the resulting UMR:

(8)   *Nepamatuju se, že bych jako dítě bývala chodila do Staronové synagogy.*
       'I don't remember going to the Old-New Synagogue as a child.'

```
(s29p1 / pamatovat-se-001 `remember'
      :ARG0 (s29e1 / entity)
      :ARG1 (s29c1 / chodit-006 `go'
          :manner (s29d1 / dítě `child'
              :mod-of s29e1
              :refer-number singular)
          :ARG1 s29e1
          :goal (s29s1 / synagoga `synagogue'
              :mod (s29s2 / Staronový) `Old-New'
              :refer-number singular)
          :aspect activity)
      :aspect activity
      :polarity -)
```

The secondary relation from COMPL leads to the personal pronoun serving as an actor to the verb *chodit* 'go', but this node gets removed from the tree in an earlier step of the conversion, and the corresponding ARG1 role is satisfied by its antecedent, the pronominal subject of the verb *pamatovat se* 'remember'. Therefore, we have to keep track of the removed nodes and reroute the secondary complement relations accordingly (see the mod-of relation in example (8)).

The conversion of coordination might complicate the process even further.

4.2. Quantitative Evaluation

The overall quantitative evaluation of the conversion procedure is presented in Table 3. The agreement between automatically converted data and manually annotated data is calculated using the same scripts as those used to assess inter-annotator agreement. Therefore, the figures in Table 3 can be compared directly with those in Table 2. It is evident that even node alignment poses a major challenge, with only less than three quarters of the nodes (72%) successfully mapped automatically.

Table 3
Quantitative comparison of manual and automatic UMRs; the table is taken from [3].
(* Unmapped nodes are ignored. ** Unmapped nodes all counted as incorrect.
MAN stands for the manual annotation, AUTO for the automatic conversion.)

UMR node mapping:						
corpus	MAN nodes	AUTO nodes	mapped	recall	precision	F_1
PDT	375	349	284	76%	81%	78%
PDTSC	442	305	235	53%	77%	63%
PCEDT	307	327	244	79%	75%	77%
total	1124	981	763	68%	78%	72%

Concept and relation comparison (only mapped nodes):*						
corpus	MAN triples	AUTO triples	match	recall	precision	F_1
PDT	844	819	502	59%	61%	60%
PDTSC	622	633	352	57%	56%	56%
PCEDT	714	588	342	48%	58%	53%
total	2180	2040	1196	55%	59%	57%

Concept and relation comparison:**						
corpus	MAN triples	AUTO triples	match	recall	precision	*juːmæʧ* = F_1
PDT	1082	916	502	46%	55%	50%
PDTSC	1318	770	352	27%	46%	34%
PCEDT	916	757	342	37%	45%	41%
total	3316	2443	1196	36%	49%	42%

This is (at least partially) caused by UMR abstract concepts: since they do not have direct counterparts in PDT-C, their reliable identification in the source data and correct transformation represent a challenging task. Of the mapped nodes, less than 60% triples (consisting of (parent node, relation, child node) or (node, attribute, value)) are correctly converted. Since the conversion is only partial and covers only selected linguistic phenomena, the results seem promising.

5. Final Remarks

The paper presented our efforts to create manually annotated *Czech UMR gold-standard data*. Such data are essential for evaluating experiments that aim to convert existing language resources into a meaning representation based on the UMR framework. The inter-annotator agreement reaches 90%, and we analyzed examples to highlight the challenges of producing such complex annotations.

We used this dataset *to evaluate a conversion procedure that transforms selected linguistic phenomena from the PDT-C corpus into the UMR representation*. Despite being a partial conversion, the method achieved 53–60% accuracy on the aligned nodes, depending on the data type. In the upcoming months, we plan to address (some of) currently uncovered phenomena.

We are convinced that such automatic conversion is an essential first step that enables the otherwise extremely demanding manual annotation of (at least some) UMR phenomena. Although our experience from the PDT-C project fully supports this hypothesis [15], we currently lack any experimental evidence confirming the usefulness of the partial automatic conversion for the UMR task.

Declaration on Generative AI

During the preparation of this work, the authors used OpenAI's ChatGPT (GPT-5, free access tier) in order to: Grammar and spelling check, Paraphrase and reword. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the publication's content.

References

[1] M. Lopatková, E. Fučíková, F. Gamba, J. Štěpánek, D. Zeman, Š. Zikánová, Towards a conversion of the Prague Dependency Treebank data to the Uniform Meaning Representation, in: L. Ciencialová, M. Holeňa, R. Jajcay, T. Jajcayová, F. Mráz, D. Pardubská, M. Plátek (Eds.), Proceedings of the 24th Conference Information Technologies – Applications and Theory (ITAT 2024), Univerzita Pavla Jozefa Šafárika v Košiciach, Košice, Slovakia, CEUR-WS.org, Košice, Slovakia, 2024, pp. 62–76. URL: https://ceur-ws.org/Vol-3792/paper7.pdf.

[2] M. Lopatková, E. Fučíková, F. Gamba, J. Hajič, H. Hledíková, M. Mikulová, M. Novák, J. Štěpánek, D. Zeman, Š. Zikánová, UMR 2.0 - Czech: Release Notes, Technical Report TR-2025-74, ÚFAL MFF UK, Prague, Czechia, 2025. URL: https://ufal.mff.cuni.cz/techrep/tr74.pdf.

[3] J. Štěpánek, D. Zeman, M. Lopatková, F. Gamba, H. Hledíková, Comparing Manual and Automatic UMRs for Czech and Latin, in: Proceedings of the Sixth International Workshop on Designing

Meaning Representations (DMR 2025), Association for Computational Lingustics, Stroudsburg, PA, USA, 2025, pp. 1–12. URL: https://aclanthology.org/2025.dmr-1.1/.

[4] P. Sgall, E. Hajičová, J. Panevová, The Meaning of the Sentence in Its Semantic and Pragmatic Aspects, Reidel, Dordrecht, 1986.

[5] E. Hajičová, A. Abeillé, J. Hajič, J. Mírovský, Z. Urešová, Treebank Annotation, in: N. Indurkhya, F. J. Damerau (Eds.), Handbook of Natural Language Processing, second edition ed., Chapman & Hall/CRC Press, Boca Raton, FL, USA, 2010, pp. 167–188.

[6] J. Hajič, E. Hajičová, M. Mikulová, J. Mírovský, Prague Dependency Treebank, in: N. Ide, J. Pustejovsky (Eds.), Handbook on Linguistic Annotation, Springer Handbooks, Springer Verlag, Berlin, Germany, 2017, pp. 555–594.

[7] J. Hajič, E. Bejček, A. Bémová, E. Buráňová, E. Fučíková, E. Hajičová, J. Havelka, J. Hlaváčová, P. Homola, P. Ircing, J. Kárník, V. Kettnerová, N. Klyueva, V. Kolářová, L. Kučová, M. Lopatková, D. Mareček, M. Mikulová, J. Mírovský, A. Nedoluzhko, M. Novák, P. Pajas, J. Panevová, N. Peterek, L. Poláková, M. Popel, J. Popelka, J. Romportl, M. Rysová, J. Semecký, P. Sgall, J. Spoustová, M. Straka, P. Straňák, P. Synková, M. Ševčíková, J. Šindlerová, J. Štěpánek, B. Štěpánková, J. Toman, Z. Urešová, B. V. Hladká, D. Zeman, Š. Zikánová, Z. Žabokrtský, Prague Dependency Treebank - Consolidated 2.0 (PDT-C 2.0), 2024. URL: http://hdl.handle.net/11234/1-5813, LINDAT/CLARIAH-CZ Digital Library, ÚFAL, MFF UK, Prague, Czechia.

[8] M. Passarotti, From Syntax to Semantics. First Steps Towards Tectogrammatical Annotation of Latin, in: K. Zervanou, C. Vertan, A. van den Bosch, C. Sporleder (Eds.), Proceedings of the 8th Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities (LaTeCH), Association for Computational Linguistics, Gothenburg, Sweden, 2014, pp. 100–109. URL: https://aclanthology.org/W14-0615/. doi:10.3115/v1/W14-0615.

[9] S. Cinková, J. Toman, J. Hajič, K. Čermáková, V. Klimeš, L. Mladová, J. Šindlerová, K. Tomšů, Z. Žabokrtský, Tectogrammatical Annotation of the Wall Street Journal, The Prague Bulletin of Mathematical Linguistics (2009) 85–104. URL: https://ufal.mff.cuni.cz/pbml/92/pbml92.pdf.

[10] J. van Gysel, M. Vigus, J. Chun, K. Lai, S. Moeller, J. Yao, T. O'Gorman, J. Cowell, W. Croft, C.-R. Huang, J. Hajič, J. Martin, S. Oepen, M. Palmer, J. Pustejovsky, R. Vallejos, Designing a uniform meaning representation for natural language processing, KI - Künstliche Intelligenz 35 (2021) 343–360. doi:10.1007/s13218-021-00722-w.

[11] J. Bonn, M. J. Buchholz, J. Chun, A. Cowell, W. Croft, L. Denk, S. Ge, J. Hajič, K. Lai, J. H. Martin, S. Myers, A. Palmer, M. Palmer, C. B. Post, J. Pustejovsky, K. Stenzel, H. Sun, Z. Urešová, R. Vallejos, J. E. L. Van Gysel, M. Vigus, N. Xue, J. Zhao, Building a broad infrastructure for uniform meaning representations, in: N. Calzolari, M.-Y. Kan, V. Hoste, A. Lenci, S. Sakti, N. Xue (Eds.), Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024), ELRA and ICCL, Torino, Italia, 2024, pp. 2537–2547. URL: https://aclanthology.org/2024.lrec-main.229/.

[12] J. Bonn, C. Bonial, M. Buchholz, H.-J. Cheng, A. Chen, C. Chen, A. Cowell, W. Croft, L. Denk, A. Elsayed, E. Fučíková, F. Gamba, C. Gomez, J. Hajič, E. Hajičová, J. Havelka, L. Havenmeier, A. Kilgore, V. Kolářová, L. Kučová, K. Lai, B. Li, J. Li, M. Lopatková, M. MacGregor, M. Mikulová, J. Mírovský, A. Nedoluzhko, S. Myers, M. Novák, T. O'Gorman, P. Pajas, A. Palmer, M. Palmer, J. Panevová, B. Post, J. Pustejovsky, P. Sgall, J. Song, L. Song, M. Ševčíková, J. Štěpánek, Z. Urešová, H. Sun, Y. Sun, R. Vallejos Yopán, J. Van Gysel, M. Vigus, K. Wright-Bettner, J. Wu, N. Xue, D. Xing, K. Xu, Z. Xu, L. Yue, D. Zeman, J. Zhao, Š. Zikánová, Z. Žabokrtský, Uniform meaning representation 2.0, 2025. URL: http://hdl.handle.net/11234/1-5902, LINDAT/CLARIAH-CZ digital library at the Institute of Formal and Applied Linguistics (ÚFAL), Faculty of Mathematics and Physics, Charles University.

[13] J. Hajič, J. Panevová, Z. Urešová, A. Bémová, V. Kolářová, P. Pajas, PDT-VALLEX: Creating a large-coverage valency lexicon for treebank annotation, in: Proceedings of The Second Workshop on Treebanks and Linguistic Theories, volume 9 of *Mathematical Modeling in Physics, Engineering and Cognitive Sciences*, Vaxjo University Press, Vaxjo, Sweden, 2003, pp. 57–68.

[14] Z. Urešová, A. Bémová, E. Fučíková, J. Hajič, V. Kolářová, M. Mikulová, P. Pajas, J. Panevová,

J. Štěpánek, PDT-Vallex: Valenční slovník češtiny propojený s korpusy 4.5 (PDT-Vallex 4.5), 2024. URL: http://hdl.handle.net/11234/1-5814, LINDAT/CLARIAH-CZ Digital Library, ÚFAL, MFF UK, Prague, Czechia.

[15] M. Mikulová, M. Straka, J. Štěpánek, B. Štěpánková, J. Hajič, Quality and Efficiency of Manual Annotation: Pre-annotation Bias, in: N. Calzolari, F. Béchet, P. Blache, K. Choukri, C. Cieri, T. Declerck, S. Goggi, H. Isahara, B. Maegaard, J. Mariani, H. Mazo, J. Odijk, S. Piperidis (Eds.), Proceedings of the 13th Conference on Language Resources and Evaluation (LREC 2022), European Language Resources Association, Marseille, France, 2022, pp. 2909–2918. URL: https://aclanthology.org/2022.lrec-1.312/.