Tutorial title
From Syntax to Semantics: Introducing UMR for NLP Annotation

Description

Uniform Meaning Representation (UMR) is a cross-linguistic semantic representation framework designed to encode sentence meaning in a structured and interpretable way. Building on the foundations of Abstract Meaning Representation (AMR), UMR extends semantic coverage to events, participants, semantic roles, temporal/aspectual information, modality, and discourse links. It is language-agnostic and therefore suitable for multilingual exploration.

This tutorial provides a beginner's introduction to UMR aimed at an audience with no prior experience with AMR, UMR, or meaning representations. The tutorial begins with a simple introduction to the essentials of Universal Dependencies (UD) needed to understand how UMR graphs can be constructed from syntactic information.
Using simple Portuguese examples, the tutorial illustrates how basic UD structures guide the creation of UMR graphs. Participants will leave with a foundational understanding of what UMR is; how it relates to syntax and semantic roles; how to create minimal UMR graphs, and how Portuguese UD treebanks can support UMR annotation.

Relevance to the PROPOR community

PROPOR has a strong tradition in linguistic annotation, UD parsing, resource creation, and linguistically informed NLP for Portuguese. At the same time, UMR remains virtually unknown in Portuguese NLP, despite its potential to support:

- semantic parsing,
- text simplification and accessibility,
- machine translation evaluation,
- event and information extraction,
- semantic alignment across languages,
- summarization and document understanding,
- interpretable NLP systems.

As Portuguese UD treebanks are already widely used in the community, UMR offers a way to add a semantic layer that builds on existing annotated resources rather than requiring new annotation initiatives from scratch.

Intended audience: computational linguists, linguists interested in annotation, students learning about semantic methods, NLP practitioners integrating semantics into pipelines.

Outline of the tutorial structure
Total Duration: 3 hours (including 10-min break)

Welcome & Motivation — 10 min
What is semantic representation? Why UMR? Why introduce it now?

Part 1 -- Minimum UD Essentials -- 25 min
A compact introduction for participants unfamiliar with UD using simple Portuguese UD examples and focusing on:

- What dependency syntax is
- Heads, dependents, and relations
- Core relations relevant for meaning (nsubj, obj, iobj, obl)

Part 2 -- UMR Basics -- 35 min

- What UMR graphs look like
- Events, entities, semantic roles
- Representing "who does what to whom"
- Simple tense, aspect, and modality

Break — 10 min

Part 3 -- How UD Informs UMR -- 30 min

- Syntactic cues for semantic interpretation.
- Identifying the main event (head verb)
- Mapping UD relations to UMR roles
- Adding basic modifiers (time, place, manner)

Part 4 — Guided Hands-On Exercise -- 35 min
Participants receive two simple Portuguese sentences annotated with UD.
Working step-by-step, they will:

Identify the main predicate
Identify core participants
Assign semantic roles (e.g., agent, patient)
Identify basic adjuncts such as time or location

Instructors will present a live demonstration covering:

Parsing the two sentences with UDPipe to obtain a CoNLL-U representation
Running a simple conversion script that uses UD structure to generate a UMR skeleton
Uploading the generated UMR graph into a tool for visualization and editing
Comparing the automatic UMR output with the manually analysed sentences

Part 5 -- Conclusion -- 10 min

How UMR can benefit Portuguese NLP
Tools and guidelines for continued learning
Opportunities for collaboration within PROPOR

Tutorial instructors

Adriana S Pagano - Universidade Federal de Minas Gerais (UFMG) - apagano@ufmg.br
https://orcid.org/0000-0002-3150-3503

Adriana S. Pagano is a Full Professor of Applied Linguistics and Translation Studies at Universidade Federal de Minas Gerais (UFMG). She holds a BA in Translation (UNLP), an MA in English Language and Literature (UFSC), and a PhD in Linguistic and Literary Studies (UFMG). She has led and collaborated on several NLP projects involving translation and post-editing, natural language understanding, and natural language generation. She currently coordinates dependency syntax annotation projects in the healthcare domain and collaborates with annotation initiatives at the Center for Artificial Intelligence (C4AI). Her research interests include systemic-functional grammar, NLU/NLG, and linguistically informed approaches to annotation.

Magali Sanches Duran - Universidade de São Paulo (USP) - magali.duran@gmail.com
https://orcid.org/0000-0002-3843-4600

Magali Sanches Duran holds a degree in Translation Studies from UNESP (1985), an MBA from FGV–São Paulo (1992), and a Master's (2004) and PhD (2008) in Linguistics from UNESP. She completed multiple postdoctoral research projects at the Núcleo Interinstitucional de Linguística Computacional (NILC/USP–São Carlos), working between 2009 and 2025 on initiatives funded by Microsoft, Samsung, and IBM at the Centro de Inteligência Artificial (C4AI). Her expertise includes syntactic and semantic corpus annotation, sentiment analysis, text complexity metrics, and the development of lexical resources for NLP.

Federica Gamba - Charles University - gamba@ufal.mff.cuni.cz
https://orcid.org/0000-0003-3632-0594

Federica Gamba is a PhD candidate at the Institute of Formal and Applied Linguistics (ÚFAL), Charles University, specializing in semantic and syntactic annotation with a focus on Uniform Meaning Representation (UMR) and Universal Dependencies (UD). She has worked on multilingual resource development as a Visiting Researcher at the University of Colorado Boulder and previously as a Research Fellow at the Institute of Computational Linguistics (CNR-ILC) in Pisa. Her background includes work on lexical and textual resources in low-resource and historical languages, supported by advanced training at the University of Pavia, IUSS Pavia, and the Université Paris-Sorbonne.

Requirements for tutorial attendees: notebook and Internet access

Requirements for tutorial instructors: Internet access and multimedia projector