

MAIL-25-07-05

Připravili jsme pro vás dvě nové úlohy: **DIR3-do2** a **DIR13-start-mezi**. Začnou se vám dříve nebo později objevovat v aplikaci. Protože si k nim nebudeme dělat schůzku, představím je zde v mailu.

Nejprve trochu obecně. Anotujeme prostorová určení s významy: *kde* (LOC), *odkud* (DIR1), *kudy* (DIR2), *kam* (DIR3), která jsou vyjádřena nějakou předložkovou frází. Anotujeme v zásadě dva typy úloh.

(i) jedna předložka

= úlohy, ve kterých anotujeme kombinaci jediného funktoru a jediné polyfunkční frekventované předložky (LOC-v6; LOC-na6, LOC-mezi -- a teď nově DIR3-do2). V těchto úlohách vedle *konkrétních lokálních významů* (zpravidla jeden dva subfunktory) podrobněji anotujeme i některé abstraktní významy (*institute, value, function...*). Special label je zde spíše výjimečný.

(ii) více předložek

= úlohy, ve kterých anotujeme určení vyjádřené různými formami (LOC-start -- a nová DIR13-start-mezi). V těchto úlohách určujeme především *konkrétní lokální významy* a abstraktní významy označujeme jen hrubě pomocí *special labels* (a k jejich rozřazení na *institute, value, level* apod. se dostaneme později).

Zadání jednotlivých úloh (uspořádání funktorů a forem) pro představu v příloze (snad je to trochu názorné).

Uspořádání forem (předložek) a významů/funkcí (subfunktorů) vykazuje na jedné straně systém a na druhé straně "chaos" (fuzziness, vagueness). Jsou případy jasné (hned víte) a na druhé straně případy, kdy se se hodí víc krabiček nebo žádná (zaseknete se a nevíte, jaký label vybrat). Tenhle dojem již z anotace již jistě máte. Jednotlivé (zvl. abstraktní) subfunktory definuju na základě toho, že se v našich datech takový význam jeví jako častý, jako zobecnitelná krabička. Repertoár navržených subfunktorů rozhodně nepokrývá všechny významy v datech/v jazyce obsažené. Moje představa je, že dáváme definované subfunktory tam, kde "sedí", podivné, nejasné případy označujeme *abstract, IDK* ap. Chtěla bych, aby výsledkem anotace byly hezké krabičky (jako *function, institute, value, work* ad.) a vedle toho hromada **fuzzy** případů, kde je význam nejasný.

A teď konkrétněji k těm novým úlohám:

DIR13-start-mezi (tabulka v příloze a na webu)

Úloha typu (ii). Pro vás, co jste již anotovali LOC-start a LOC-mezi, bude tato úloha mnohem srozumitelnější, anotujeme **základní prostorové lokalizace na vertikální a horizontální ose** (*front -- behind, above -- below* a některé další *outside, around, beside* atp.). U abstraktních významů vyjádřených příslušnými předložkami odlišujeme speciálně *quantity* (pro případy, které známe z jiných úloh jako *value (nad 3 procenta), level (pod průměrnou úroveň)*). Řada významů jako *institute, function* odlišovaných v úloze typu (i) zde dostane jen special label "**abstract**" a jejich jemno-zrná klasifikace se provede později. Vaše komentáře zde budou velmi užitečné.

DIR3-do2 (tabulka v příloze a na webu)

Úloha typu (i). Jedná to "směrová" určení "paralelní" ke statickým urcům (LOC) vyjádřeným předložkou v+6.

Kočka je v krabici. - Dej kočku do krabice. (inside)
Studoval ve škole. -- Nastoupil do školy. (institute)
V knize se píše o násilí. --- Do knihy vepsal násilí. (work)
atp.

Poznámky:

towards - na rozdíl od *inside* jde o pouhé směřování, jehož výsledkem není umístění v cíli (*cesta do Tanzánie*-*inside* (cestující jsou nakonec v Tanzánii) vs. *cesta do schodů/kopce*-*towards* (cestující jsou nakonec ve schodech / kopci);
okna do ulice-*towards* (okna nejsou nakonec v ulici). Patří sem *dívání, pohled*.

event vs. **time** -- z *eventu* oddělují *time*. Jde o podobné subfunktory.

-- *event* je pro případy, kdy podstatné jméno po předložce vyjadřuje děj (a proto i čas a místo -- to může být rozvitím): *do prosincového, pražského slosování*

-- *time* je pro případy, kdy substantivum po předložce vyjadřuje přednostně čas děje (*do druhého kola slosování*). Tato určení hraničí s časovými -- mají však fuktor DIR3, protože se na ně ptáme "kam" (a nikoliv kdy/dokdy).

aim - účel, záměr, chodit někam s nějakým úmyslem, „aby“ (*chodit do houslí* = abych se naučil hrát na housle). Jde i o záměr přimět někoho jiného k nějaké činnosti (*zatlačit Slavii do obrany*).

state - dostat se nebo dostat něco do nějakého stavu (*upadli jsme do starých chyb*). Mezi *aim* a *state* je někdy úzká hranice. Oproti *aim* jde zde více o výsledný stav, kdežto v *aim* jde více o proces, o vykonávání nějaké činnosti.

Dále upozorňuju, že podle PDT-vallexu (<https://lindat.mff.cuni.cz/services/PDT-Vallex4.5/PDT-Vallex4.5.html>)

je v *investoval do zlata, vstoupit do věcí veřejných, zapojit se do operace, promítnout se do cen* určení do+2 valenční. Prosím označit jako **skip-valency**.

Slučuju případy **skip-SECPREP a PART** v obou úlohách pod jeden label.

Uvědomuji si, že v hromadě těch úloh a labels může začít být trochu zmatek. Proto prosím opravdu o **pečlivost a poctivost. Pravidelná a častá anotace** pak vzhledem k proměnlivosti zadání skutečně vede k rychlejší anotaci a větší kvalitě a konzistenci. Moc děkuju.

Text tohoto emailu též jako příloha (mail-DIR13-do2)

Přeju hodně zdaru, ať vás anotace baví. Prosím piště, ptejte se...

Zdravím vás i za Báru a Honzu, bez kterých by nové úlohy nebyly.
maruska

Kdyby někoho hodně zajímal ten **systém**, který v uspořádání subfunktorů vídím, může se podívat do knížky, kterou jsem napsala s paní profesorkou Panevovou: *Formy a funkce okolnostních určení v češtině. Určení prostorová a časová* (<https://ufal.mff.cuni.cz/books/2021-mikulova>).

A kdyby někoho hodně zajímal ten **chaos**, tak o tom jsme napsali článek s Honzou Štěpánkem: *Label Bias in Symbolic Representation of*

Meaning <https://sigann.github.io/LAW-XIX-2025/pdf/2025.lawxix2025-1.12.pdf>

(Pro ans, krh, ivh, toh -- ten článek vychází z experimentu, na kterém jste se podíleli, poděkování anotátorům v Acknowledgments platí pro vás! Výsledky budu prezentovat letos v létě na konferenci ve Vídni. Děkuje.)