Mapping PDT-Vallex functors to PropBank arguments

Source:

1/ EngVallex 2.0 - https://lindat.mff.cuni.cz/repository/xmlui/handle/11234/1-3526

2/ SynSemClass – synsemclass\_eng\_cms.xml r.15196 from https://svn.ms.mff.cuni.cz/trac/SynSemClass/browser/data/anotace/BEST

=> file ssc\_role\_pb\_arg\_mapping\_with\_pb\_links\_control.txt

Example for the class vec00026

Class roles: Mover, Area, Vehicle

Valid english classmembers (cms):

- drive (EngVallex-ID-ev-w1063f3)

link to the PropBank: drive/drive.01

valency frame: ACT, ?PAT, ?DIR1, ?DIR2, ?DIR3, ?INTT

mapping ACT→Mover, DIR1→Area, DIR2→Area, DIR3→Area, PAT->Vehicle

- pedal (EngVallex-ID-ev-w2230f3)

link to the PropBank pedal/pedal.01

valency frame: ACT, ?DIR3

mapping ACT→Mover, DIR3→Area, #sth->Vehicle

- ride (EngVallex-ID-ev-w2758f1)

link to the PropBank ride/ride.01

valency frame: ACT, PAT

mapping ACT→Mover, #swh→Area, PAT->Vehicle

- run (EngVallex-ID-ev-w2806f8)

no mapping to the PropBank

valency frame: ACT, ?PAT

mapping ACT→Mover, #swh→Area, PAT->Vehicle

Links from EngVallex functors to the ProbBank arguments

(<http://ufal.mff.cuni.cz/biblio/publications/2006-cinkova-p116560102333242.html>)

drive (ev-w1063f3)

<element functor="ACT" type="oblig">

<externallink type="PB role" rolesetref="drive.01" n="0" frequency="19"/>

</element>

<element functor="PAT" type="non-oblig">

<externallink type="PB role" rolesetref="drive.01" n="1" frequency="8"/>

</element>

<element functor="DIR1" type="non-oblig"/>

<element functor="DIR2" type="non-oblig"/>

<element functor="DIR3" type="non-oblig"/>

<element functor="INTT" type="non-oblig"/>

pedal (ev-w2230f3)

<element functor="ACT" type="oblig">

<externallink type="PB role" rolesetref="pedal.01" n="0" frequency="1"/>

</element>

<element functor="DIR3" type="non-oblig">

</element>

ride (ev-w2758f1)

<element functor="ACT" type="oblig">

<externallink type="PB role" rolesetref="ride.01" n="0" frequency="6"/>

</element>

<element functor="PAT" type="oblig">

<externallink type="PB role" rolesetref="ride.01" n="1" frequency="9"/>

</element>

run (ev-w2806f8)

<element functor="ACT" type="oblig">

<externallink type="PB role" rolesetref="run.01" n="1" frequency="8"/>

</element>

<element functor="PAT" type="non-oblig">

<externallink type="PB role" rolesetref="run.01" n="4" frequency="1"/>

</element>

Mapping for class vec00026

Mover

arg-0 26 (19 for drive + 1 for pedal + 6 for ride)

Area

Vehicle

arg-1 17 (8 for drive + 9 for ride)

- we have no mapping for Area – functors DIR1, DIR2 and DIR3 (drive and pedal) have no mapping to PropBank arguments

- classmember “run (EngVallex-ID-ev-w2806f8)” has no mapping to the PropBank – so we do not include its mapping

3/ SynSemClass – synsemclass\_ces\_cms.xml r.15196 from <https://svn.ms.mff.cuni.cz/trac/SynSemClass/browser/data/anotace/BEST>

4/ pdt\_vallex\_umr\_mapping.txt and frames\_czech\_chars.json from https://github.com/ufal/UMR/tree/main/valency-frames-cs-verbs/

=> pdt\_pb\_mapping.xlsx

In this file is one row for each frame and one row for each of elements from its valency frame.

In the first column is the UMR id for the frame, in the second column is PDT-Vallex id (with the link to the TEITOK - <https://lindat.mff.cuni.cz/services/teitok/pdtc10/>) or functor and its form from PDT-Vallex. In the third column is collected mapping for the funcotr and values in the following columns – there are two columns for the each occurrence of the frame in the SynSemClass.In the first of them is class ID (with the link to the SynSemClass4.0 (<https://lindat.mff.cuni.cz/services/SynSemClass40/> - this is an older version of the data than we use but only one for easy browsing) and Functor→Role mapping for the frame in this class, in the second is the SynSemClass roles with mapping to the PropBank arguments and Functor→Role with mapping to the PropBank arguments (these values are collected to the column C)

Example for the verb frame “anulovat (v-w70f1)” (UMR frame “anulovat-001”)

|  |  |  |
| --- | --- | --- |
| "anulovat-001" | anulovat (v-w70f1) |  |
|  | ACT: 1 | ACT->ARG0/122,ARG1/14 |
|  | PAT: 4 | PAT->ARG1/222,ARG2/1 |
|  |  |  |

|  |  |
| --- | --- |
| vec00198 | Agent(ARG0/108,ARG1/14), Abandoned(ARG1/199,ARG2/1) |
| ACT->Agent | ACT->Agent(ARG0/108,ARG1/14) |
| PAT->Abandoned | PAT->Abandoned(ARG1/199,ARG2/1) |

|  |  |
| --- | --- |
| vec01482 | Cause(ARG0/14), Affected(ARG1/23) |
| ACT->Cause | ACT->Cause(ARG0/14) |
| PAT->Affected | PAT->Affected(ARG1/23) |

ACT is mapping to the ARG0 122x (108x from vec00198 ACT→Agent + 14x from vec01482 ACT→Cause) and to the ARG1 14x (from vec00198 ACT→Agent)

PAT is mapping to the ARG1 222x (199x from vec00198 PAT→Abandoned, 23x from vec01482 PAT→Affected) and to the ARG2 1x (from vec00198 PAT→Abandoned)