

Wytyczne KPWr

Sytuacje (atrybuty i relacje)

Spis treści:

Spis treści:	1
Założenia	3
Komponenty relacji	4
Tradycyjne wykładniki predykcji	4
Złożone znaki predykcji	5
Relacje między sytuacjami a klasyfikacja sytuacji	6
Semantyczno-składniowa hierarchia sytuacji	8
EVENT_ACTION	9
EVENT_STATE	9
REPORTING (anotacja EVENT_REPORTING)	9
PERCEPTION (anotacja EVENT_PERCEPTION)	10
ASPECTUAL (anotacja EVENT_ASPECTUAL)	10
I_ACTION (anotacja EVENT_I_ACTION)	11
I_STATE (anotacja EVENT_I_STATE)	11
LIGHT_PREDICATE	12
Relacje między sytuacjami	12
SLINK	12
ALINK	14

Sposób anotacji	14
Szczegółowe rozstrzygnięcia	15
Atrybuty sytuacji	16
Atrybuty MAKEINSTANCE według TimeML	16
Czasowniki w formie osobowej	16
Rzeczowniki i przymiotniki	17
Atrybuty sytuacji	18
Ogólność (generality)	18
Biegunowość (polarity)	18
Negacja zdaniowa	18
Negacja wyrazowa	19
Modalność (modality)	19
Modalność intencjonalna	20
Modalność epistemiczna	21
Modalność deontyczna	21
Czas (tense)	22
Aspekt (aspect)	22
Kandydaci do atrybutów	22
Referencje	23

Założenia

★ Definicje

- Sytuacja – akcja (zdarzenie, czynność, proces, działanie, wypadek, akt) lub stan (a także ich zbiory) rozpatrywane na poziomie pozatekstowym.
- Językowy wykładnik sytuacji – fragment tekstu denotujący pewną sytuację. Dla uproszczenia znakowania językowe wykładniki sytuacji reprezentowane są przez głowę frazy będącej wykładnikiem (por. Jak znakujemy).
- Relacja - zależność pomiędzy dwiema sytuacjami

★ Znakujemy różne formy wyrażania sytuacji, nie tylko czasowniki (zob. punkt Co znakujemy).

★ Wyróżniamy siedem klas sytuacji: ACTION, STATE, REPORTING, PERCEPTION, ASPECTUAL, I_ACTION, I_STATE (zob. punkt Klasy zdarzeń). Specyfikacja prezentuje ich skrócone opisy szczegółowo przedstawione w oddzielnej specyfikacji

★ Wyróżniamy dwie klasy relacji: SLINK i ALINK, które mają swoje podtypy

★ Zawsze znakujemy pojedynczy token. W przypadku fraz najczęściej będzie to głowa (zob. Jak znakujemy)

★ Relacje są wprowadzane pomiędzy pojedynczymi tokenami

Komponenty relacji

W tej części rozważymy, na które składniki zdania zwrócić uwagę, przystępując do znakowania sytuacji. Głównym celem będzie odpowiedź na pytanie, które z nich znakujemy.

Tradycyjne wykładniki predykcji

Jodłowski (31-33) wskazuje na następujące wykładniki predykcji:

1. czasowniki
 - 1.1. osobowa forma czasownika
 - 1.2. formy na *-no*, *-to*
 - 1.3. czasownik niefleksyjny nieosobowy (*trzeba*, *można...*)
 - 1.4. czasownik niefleksyjny momentalny (*ciach*, *lu...*)
2. deklinacyjna forma orzecznika (*Śląsk mistrzem Polski*)
3. krótka forma przymiotnika w formie orzecznika (forma rzeczownikowa: *zdrow*, *wart...*)
4. inne formy orzecznika w pozycji po podmiocie
5. wyraz *to* jako łącznik
6. rzeczowniki *żał*, *wstyd*, *strach*, *pora* itp.
7. przysłówki, zaimki przysłówne i modulanty sugerujące możliwość zastosowania czasownika jako ich substratu (np. *Na mnie już nie licz. Trudno, musisz wybierać*)
8. kontekst, sytuacja lub charakter informacji sugerujące domyślność predykcji (np. *Sam jestem. Żona na nieszporych*)
9. akcent lub intonacja
10. pauza (*Felek już student, Elżunia - panna sklepowa*)

Złożone znaki predykcji

Specyfikacja TimeMLowa dla języka angielskiego uwzględnia złożone sposoby orzekania. Niezadowolające są jednak: przejrzystość klasyfikacji (z tego względu w niniejszym opracowaniu proponujemy własną jej interpretację) oraz metody znakowania, którym brak systematyki.

Interesującą klasyfikację złożonych predykatów przedstawia Ewa Jędrzejko (2011, 34-37). W skrócie można ją przedstawić następująco:

1. (J1)
standardowe predykaty nominalne z czasownikami o funkcji łącznika: *być, stać się...*
2. (J2)
predykaty modalne z czasownikami: *móc, musieć, należy...*
3. (J3)
predykaty fazowo-aspektualne z czasownikami: *zacząć, trwać, przestać...*
4. (J4)
frazy z współznaczącymi czasownikami generycznymi (synsemantycznymi): *robić, czuć, mieć, doświadczać*, element werbalny nie niesie dodatkowych znaczeń, poza tymi wyrażonymi przez element nominalny, a całość jest właściwie synonimiczna z pełnym czasownikiem (*robić pranie = prać; odczuwać zaniepokojenie = niepokoić się; doznać cierpienia = cierpieć*)
5. (J5)
peryfrastyczne predykaty o charakterze metaforycznym - pełnoznanące, konkretne czasowniki są użyte w tych konstrukcjach metaforyczne, są motywowane przez konceptualne metafory (*brać udział, podnieść głos*)
6. (J6)
idiomy (*pleść, co ślina na język przyniesie, zadzierać nosa*)

Zasadnicza różnica pomiędzy 4 a 5 typem polega na tym, że w przypadku 4 czasowniki nie wprowadzają dodatkowej zawartości modyfikującej wartość predykatu nominalnego, zmieniają jedynie perspektywę funkcjonalną, pomagają zaliczyć sytuację do określonej klasy (akcja, stan...) i wprowadzają gramatyczne właściwości czasownika (czas, aspekt, osobę...). Frazy te mają charakter endocentryczny.

W odniesieniu do predykatów złożonych przyjmujemy następujące ustalenia (symbole odnoszą się do klasyfikacji Ewy Jędrzejko):

- J1. Nic nie znakujemy (poza wyjątkiem dotyczącym przymiotników z kategorii I_State (por. przykład N). Łącznik jako czasownik o gramatycznej funkcji (słowo posiłkowe) również nie podlega znakowaniu.

- "Wiesław w 2015 został strażakiem"

J2. Znakujemy osobną klasą czasownik modalny, osobną czasownik z nim połączony

J3. Znakujemy osobną klasą (ASPECTUAL) czasownik fazowy, osobną czasownik z nim połączony

J4. Znakujemy osobną klasą czasownik synsemantyczny (ang. *light predicate*, *light verb*) - LIGHT_PREDICATE; osobną zaś klasą nominalizację:

- "Józek zrobił_[LIGHT_PREDICATE] pranie_[ACTION] wczoraj rano" (zrobił pranie = wyprał)
- "Demonstracje_[ACTION] trwały_[LIGHT_PREDICATE] dwa tygodnie" ?? (demonstracje trwały = demonstrowali)
- "Zbigniew W. dokonał_[LIGHT_PREDICATE] morderstwa_[ACTION] porwanego mężczyzny" (dokonał morderstwa = zamordował)
- "Zbigniew W. dokonał_[LIGHT_PREDICATE] egzekucji_[ACTION] na porwanym mężczyźnie"
- "Demonstracje_[ACTION] miały_[LIGHT_PREDICATE] miejsce"

J5. Znakujemy tylko czasownik. Chunker powinien rozpoznać, że jest to jednostka wielowyrazowa.

[Zasada 1] Czasownik pojawiający się w tej klasie nie należy do kategorii *light predicate* tylko do kategorii, do której zaliczylibyśmy całą frazę. Nie rozpatrujemy tego czasownika w oderwaniu od innych elementów złożonego predykatu.

- "Franek [dał_[ACTION] ciała]_{VP} na wczorajszym egzaminie."
- "Franek [dał_[REPORTING] znać]_{VP} o wczorajszym egzaminie."

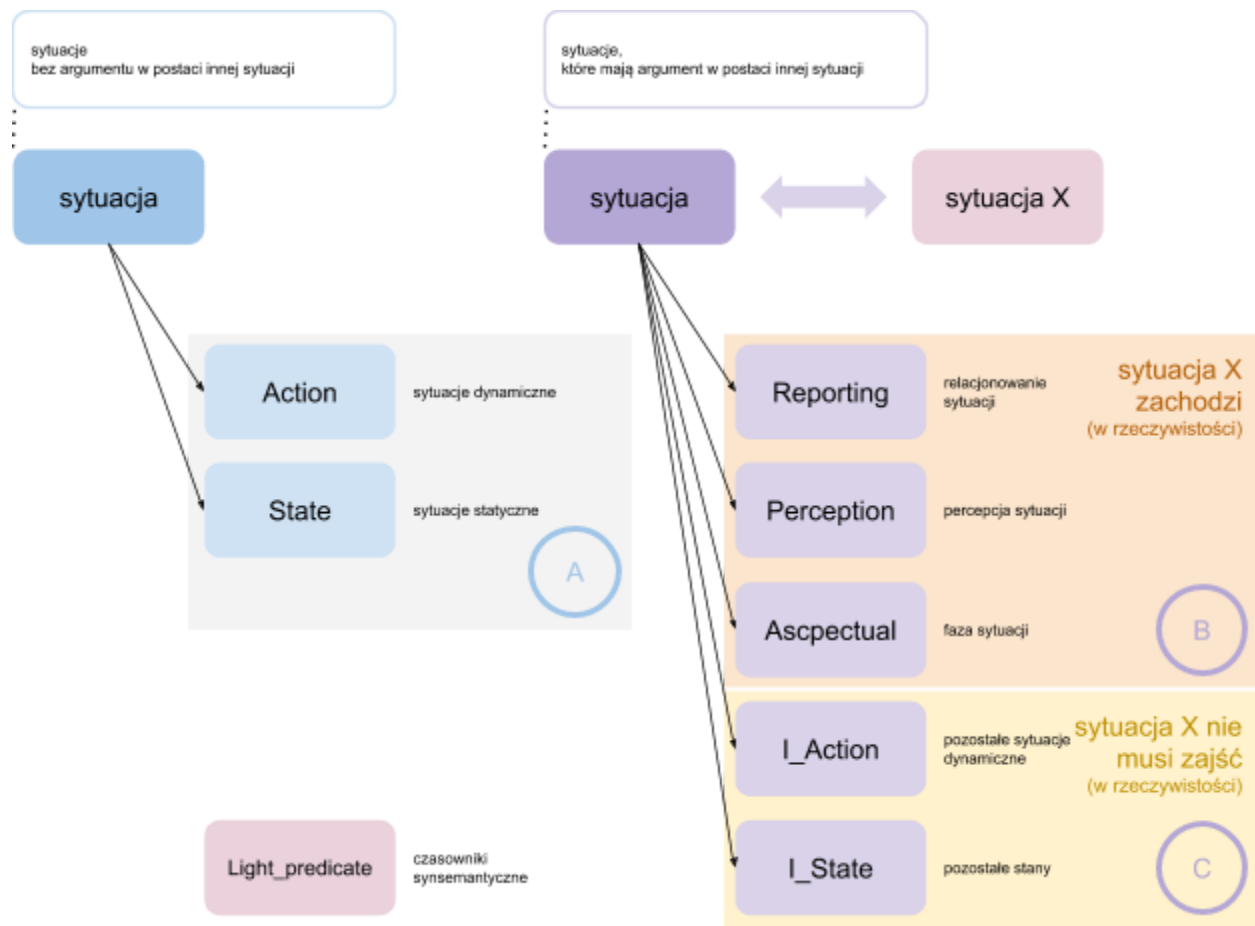
J6. Znakujemy tylko czasownik
Stosujemy Zasadę 1.

Relacje między sytuacjami a klasyfikacja sytuacji

Podział sytuacji w specyfikacji TimeML tylko częściowo pokrywa się z klasyczną klasyfikacją czasowników opartą na przykład na schemacie Vendlerowskim czy klasycznej już koncepcji Laskowskiego (1998, 152-156). Został zaznaczony właściwie tylko podział na sytuacje dynamiczne i statyczne (ACTION vs. State oraz I_Action vs. I_State). Może to wynikać z faktu, że opisy sytuacji przyjmowane w tym standardzie to nie tylko formy czasowników. Najważniejszym jednak powodem wydaje się nastawienie na określony cel. Proponowana typologia ma służyć przede wszystkim wychwytywaniu relacji między sytuacjami. Daje się więc zauważyć (niewskazany *explicite* w specyfikacji TimeML) podział sytuacji na dwie grupy:

- sytuacje, które nie mają argumentu w postaci innej sytuacji
- sytuacje, które takiego argumentu wymagają (lub taki argument dopuszczają).

Uproszczony schemat wyglądałby następująco:



Rys. 1 Typologia sytuacji

Podsumowując, wyróżniamy dwa główne kryteria podziału sytuacji:

Przebieg w czasie:

- statyczne,
- dynamiczne,

Dopuszczalny argument:

- sytuacje, których argumentem może być inna sytuacja,

- sytuacje, których argumentem nie może być inna sytuacja.

Pełna hierarchia kategorii została przedstawiona w punkcie Semantyczno-składniowa hierarchia sytuacji.

Sytuacje z grupy B i C (Reporting, Perception, Aspectual, I_Action, I_State) oznaczamy także wtedy, gdy ten argument nie został wyrażony w tekście w formie, która podlega anotacji. A więc wprowadzamy taką anotację zarówno w przypadku zdania:

- Marek powiedział_[REPORTING] nam dużo o górach,

jak i zdania:

- Marek nam powiedział_[REPORTING], że idzie_[ACTION] już do domu.

dla którego zostaną wprowadzone dwie anotacje (Reporting i Action) ze względu na to, że raportowaniu (*powiedział*) podlega wyrażona w zdaniu akcja (*idzie*). Należy przy tym pamiętać, że nie wykraczamy w analizie poza granice zdania.

Nieco inny przykład, lecz rodzący podobne konsekwencje, to zdanie:

- Według projektu miał_[I_STATE] to być sześciomiejscowy metalowy dolnopłat pasażerski z silnikiem Pratt-Whitney Wasp o mocy 420 KM.

Występują w nim dwa powiązane czasowniki (*miał* i *być*). Argument w postaci czasownika “być” nie podlega znakowaniu ze względu na swoją gramatyczną funkcję, lecz zostanie wprowadzone oznaczenie (I_State) dla czasownika “mieć”, gdyż kontekst wskazuje na powiązanie z inną sytuacją.

Ale:

- Według projektu miał_[I_STATE] powstać_[ACTION] sześciomiejscowy metalowy dolnopłat pasażerski z silnikiem Pratt-Whitney Wasp o mocy 420 KM.

Do kategorii I_Action i I_State zaliczamy wykładniki na podstawie kontekstu.

Semantyczno-składniowa hierarchia sytuacji

1. pełnoznaczące [brak nazwy]
 - 1.1. statyczne [brak nazwy]
 - 1.1.1. bez argumentu zdarzeniowego [STATE]
 - 1.1.2. z argumentem zdarzeniowym (nie musi wystąpić w tekście) [I_STATE]
 - 1.2. dynamiczne [brak nazwy]
 - 1.2.1. bez argumentu zdarzeniowego [ACTION]
 - 1.2.2. z argumentem zdarzeniowym (nie musi wystąpić w tekście) [brak nazwy]

- 1.2.2.1. [REPORTING]
- 1.2.2.2. [PERCEPTION]
- 1.2.2.3. [ASPECTUAL]
- 1.2.2.4. [L_ACTION]

2. niepełnoznaczne (wymagają uzupełnienia w postaci nominalizacji) [LIGHT_PREDICATE]

EVENT_ACTION

[\[next class\]](#) [\[index\]](#)

Sytuacje dynamiczne, które nie wiążą się z argumentem w postaci innej sytuacji (zob. REPORTING, PERCEPTION, ASPECTUAL, L_ACTION). Chodzi tu o pewien potencjał wykładnika sytuacji tzn. w ramie walencyjnej wyrażenia nazywającego daną sytuację nie znajduje się jako argument inna sytuacja. Konsekwencją takiego podejścia jest Założenie pomocnicze 1.

EVENT_STATE

[\[next class\]](#) [\[first class\]](#) [\[index\]](#)

Stany to sytuacje polegające na tym, że ktoś/coś jest kimś/czymś, ktoś/coś jest jakimś [Apresjan 2000, 47-48]. Jako EVENT_STATE znakujemy stany niewyróżnione jako L_STATE. \

REPORTING (anotacja EVENT_REPORTING)

[\[next class\]](#) [\[first class\]](#) [\[index\]](#)

Akcja osoby lub instytucji polegająca na informowaniu o jakiejś sytuacji, relacjonowaniu jej, np. *mówić, raportować, opowiadać, wyjaśniać, przedstawiać,...*

Reguły:

Co ważne, pojawienie się kategorii "Reporting" powinno oznaczać zaistnienie jakiejś innej (relacjonowanej) sytuacji.

Przykłady:

- "Perkusista wyjaśnił_[REPORTING] iż odejście_[ACTION] z zespołu "jest dla mnie pozytywnym ruchem naprzód (...)".
- "W dniu 14 października 1938 roku von Blücher raportował_[REPORTING] do szefa wydziału MSZ w Berlinie, Gründherna, iż władze szwedzkie zamierzają fortyfikować_[ACTION] Wyspy Alandzkie obawiając się ataku niemieckiego na ten archipelag."

PERCEPTION (anotacja EVENT_PERCEPTION)

[\[next class\]](#) [\[first class\]](#) [\[index\]](#)

Sytuacje polegające na fizycznej percepcji innych sytuacji, np. *zobaczyć, usłyszeć*. Jest to druga kategoria (po Reporting), która powinno oznaczać zajście jakiejś innej sytuacji.

- Świadkowie powiedzieli policji, że widzieli_[PERCEPTION] uciekającego_[ACTION] człowieka.
- Słyszeliśmy_[PERCEPTION] tysiące małych eksplozji_[ACTION] tam w dole.

Tego typu sytuacje są powiązane z rolą świadka zdarzenia i wiąże się z nimi swego rodzaju bierność subiekta.

ASPECTUAL (anotacja EVENT_ASPECTUAL)

[\[next class\]](#) [\[first class\]](#) [\[index\]](#)

Zdarzenie określające zmianę fazy trwania innej sytuacji. Fazy odnoszą się do sytuacji, które rzeczywiście zaszły/zachodzą, należy więc pod tym względem tej samej podgrupy co Reporting i Perception.

Fazy sytuacji:

1. Rozpoczęcie, np. *zacząć, inicjować*
 - "Rząd rozpoczął_[ASPECTUAL] prace nad nową ustawą."
 - "W drugiej połowie 1930 roku na zamówienie_[ACTION] Ministerstwa Komunikacji inż. Stanisław Prauss rozpoczął_[ASPECTUAL] prace_[ACTION] projektowe nad sześciomiejscowym samolotem pasażerskim."
2. Wznowienie, np. *restartować, wznowić*
 - "Rząd wznowił_[ASPECTUAL] prace_[ACTION] nad ustawą."
3. Przerwanie, np. *przestać, odstąpić, odwołać, zwolnić, zrezygnować*
4. Zakończenie, np. *skończyć*.
5. Kontynuacja, np. *kontynuować, trwać*

I_ACTION (anotacja EVENT_I_ACTION)

[\[next class\]](#) [\[first class\]](#) [\[index\]](#)

Akcja podejmowana w związku z inną sytuacją. Argumentem akcji intensjonalnej jest inne zdarzenie, które powinno (ale nie musi) być wprost wyrażone w tekście. Intensjonalność w logice jest rozumiana jako "własność połączeń zdaniowych (zdań złożonych) polegająca na tym, że ich wartość prawdziwościowa (tj. prawda lub fałsz) nie wynika wyłącznie z prawdy lub fałszu zdań składowych, lecz także z innych czynników, takich jak np. sposób połączenia i kolejność występowania". W naszym ujęciu intensjonalność jest własnością sytuacji polegającą na tym, że jej prawdziwość nie jest uzależniona od prawdziwości (rozumianej jako zajście) sytuacji będącej argumentem. Oczywiście przy przestrzeganiu praw regulujących komunikację językową (chodzi tu zwłaszcza o Grice'owską maksymę jakości).

1. (próba wykonania akcji) *próbować...*
2. (opóźnienie) *przekładać, opóźniać...*
3. (uniknięcie) *zapobiegać, unikać...*
4. (dyrektywy) *prosić, namawiać, rozkazywać...*
5. (obietnice) *obietcywać, zapewnić, oferować, zdecydować, zgodzić się*
6. (przysięgi) *przysięgać*
7. (propozycje, asercje) *twierdzić, sugerować*

I_STATE (anotacja EVENT_I_STATE)

[\[next class\]](#) [\[first class\]](#) [\[index\]](#)

Zwroty wprowadzające możliwe sytuacje lub stany. Argumentem I_State nie muszą więc być inne stany (State).

Sytuacje oznaczane tagiem I_State mają charakter stanów odnoszących się do alternatywnych lub możliwych światów.

1. (myślenie) *wierzyć, myśleć, wyobrażać sobie...*
2. (wola) *chcieć, pragnąć, lubić...*
3. (oczekiwania) *oczekiwać, planować...*
4. (emocje i uczucia) *bać się, martwić się...*
5. (potrzeby) *potrzebować, wymagać...*
6. (gotowość) *być gotowym, być przygotowanym...*
7. (potencja) *móc, nie móc...*

Sytuacje zaliczane do kategorii I_State spełniają również wymagania stawiane przed sytuacjami z kategorii State.

LIGHT_PREDICATE

[\[first class\]](#) [\[index\]](#)

Do tej klasy należą czasowniki o charakterze pomocniczym (ang. *light predicate*), a więc współznaczące czasowniki generyczne (synsemantyczne), o znaczeniu bardzo ogólnym, np. *robić, czuć, mieć, doświadczać*. Zawsze stanowią one element werbalny predykatów złożonych z *light predicate* i nominalizacji, który nie niesie dodatkowych znaczeń poza tymi wyrażonymi przez element nominalny (por. Złożone znaki predykacji), a służy przede wszystkim do wtórnej werbalizacji.

Relacje między sytuacjami

Źródłem relacji między sytuacjami są przede wszystkim sytuacje należące do kategorii Perception, Reporting, Aspectual, I_Action i I_State. Istnieją jednak od tego wyjątki dotyczące zdań okolicznikowych celu i warunkowych (por. poniżej).

Obiektem docelowym relacji jest tzw. Sytuacja X. Jeden wykładnik może mieć kilka obiektów docelowych, np. *mogę leżeć i jeść arbuzy - mogę → leżeć; mogę → jeść*. Relacja może być poprowadzona tylko między składnikami jednego wypowiedzenia. Podział na wypowiedzenia został dokonany automatycznie (prezentuje go opcja `Display every sentence separately`). Każdy napotkany błąd w segmentacji zgłaszamy koordynatorowi (nie wprowadzamy relacji).

SLINK

Relacja SLINK łączy dwa zdarzenia, z których jedno (nadrzędne) ma charakter:

1. Modalny

Relacja jest przywołana przez zdarzenia wprowadzające odniesienie do możliwych światów, np.:

Jan obiecał Marysi, że *kupi* jej buty
Marysia chciała *kupić* sobie biżuterię

(obiecał → kupi)
(chciała → kupić)

2. Faktywny

Niektóre czasowniki wprowadzają presupozycję wystąpienia (lub niewystąpienia) połączonych z nimi zdarzeń, np.:

Jan zapomniał *kupić* żonie butów

(zapomniał → kupić)

Jan zapobiegł *rozwodowi* kolegi

(zapobiegł → rozwodowi)

3. Poświadczający

Poświadczenie może mieć charakter pozytywny bądź negatywny. Typowe jest dla zdarzeń z klasy REPORTING i PERCEPTION, np.:

Jan widział, jak żona *kupowała* sobie buty
Marysia opowiadała o *ślubie* koleżanki

(widział → kupowała)

(opowiadała → ślubie)

Relacje typu SLINK mogą mieć różną naturę.

Relacja SLINK może być oparta na leksyce. Jej wyzwalaczami są sytuacje z kategorii L_ACTION, L_STATE, PERCEPTION i REPORTING, które są zazwyczaj predykatami, których argumentem jest zdanie, czasownik stanowiący człon podrzędny predykatu złożonego lub fraza rzeczownikowa, której głową jest nominalizacja czasownika. SLINK łączy takie sytuacje z ich argumentami.

W przypadku REPORTING i PERCEPTION musi zostać wprowadzony SLINK.

Przykłady:

10 minut później zobaczyłem, jak inny samolot *trafił* w WTC

To *brzmiało*, jak rakietą, powiedział Jose.

Na podobnej zasadzie wprowadzany jest link pomiędzy L_ACTION lub L_STATE a podrzędnym względem nich zdarzeniem.

Przykłady:

Profesor próbował *wytłumaczyć* zawichości metafizyki.

Chcieliśmy *uczestniczyć* w pielgrzymce, ale niestety nie mieliśmy namiotu.

Podrzędne zdarzenia narzucają następujące typy relacji SLINK:

- PERCEPCJA – evidential (poświadczający pozytywnie) neg_evidential (poświadczający negatywnie)
- L_ACTION, L_STATE – modal (modalny), factive (faktyczny), counter_factive (niefaktyczny)
- REPORTING może być wiązany z podrzędnym zdarzeniem relacjami SLINK dowolnego typu

Relacja SLINK może być również oparta na strukturze składniowej wypowiedzeń złożonych. W zdaniach okolicznikowych celu, SLINK łączy czasownik ze zdania głównego z czasownikiem ze zdania podrzędnego, np.

*Komisja Środowiska musi **przyjąć** regulacje, by zapewnić ludziom bezpieczeństwo od skażeń radioaktywnych.*

Relacja SLINK pojawia się również w zdaniach warunkowych i łączy zdarzenia ze zdania poprzedzającego z tym, który występuje w zdaniu następnym, np.

*15 grudnia obiecał **wrócić** do negocjacji, jeśli firma Zalewskiego spełni następujące warunki.*

Jest to swego rodzaju wyjątek - w przypadku takich wypowiedzeń relacją SLINK mogą zostać połączone sytuacje, z których żadna nie należy do kategorii I_ACTION, I_STATE, PERCEPTION czy REPORTING.

Znakowaniu mogą podlegać relacje, które nie zachodzą bezpośrednio między składnikami zdania. Do tego typu powiązań należą:

- A. strona bierna
- B. czasy złożone

W obu przypadkach nie podlegają znakowania słowa posiłkowe. Z tego względu możemy poprowadzić relację z pominięciem słowa posiłkowego, np.

- A. *może być magazynowana - może → magazynowana*
- B. *będzie mogła być magazynowana - mogła → magazynowana*

Innym przypadkiem są frazy z Light Predicatem. **TODO**

ALINK

ALINK to Aspectual Link i reprezentuje relacje pomiędzy czasownikami należącymi do klasy ASPECTUAL (a więc czasownikami fazowymi) a ich argumentami zdarzeniowymi. Tego typu relacje obserwujemy w przypadku:

1. Inicjacji

Jan zaczął *pisać*

2. Zatrzymania (i przerwania)

Jan przestał *gadać*

3. Kontynuacji

Jan kontynuował *drzemkę*

4. Wznowienie

Jan wznowił *negocjacje*

Powyższe przykłady obrazują jednocześnie typy relacji ALINK.

Sposób anotacji

1. Logujemy się do Inforexa i wybieramy korpus
2. Włączamy warstwę anotacji TimeML, a w niej podwarstwę EVENT
3. Znajdujemy w tekście znajdujemy następujące sytuacje:
 1. ASPECTUAL
 2. REPORTING, PERCEPTION, I_ACTION, I_STATE
4. Znakujemy relacje między nimi, a ich argumentami (innymi sytuacjami, do których się odnoszą)
 1. jeśli znajdziemy ASPECTUAL to klikamy ją, w polu Annotation relations klikamy Add relation i z rozwijanej listy wybieramy ALINK, a potem zaznaczamy w tekście sytuację, z którą się łączy, np. zaczął → pracować
 2. jeśli znajdziemy REPORTING, PERCEPTION, I_ACTION lub I_STATE to klikamy ją, w polu Annotation relations klikamy Add relation i z rozwijanej listy wybieramy SLINK, a potem zaznaczamy w tekście sytuację, z którą się łączy, np. chciał → pracować+
5. Po przejrzaniu dokumentu ustawiamy flagę ev_rel na "gotowy"

Szczegółowe rozstrzygnięcia

- Wkrótce zaczęto o nią podejrzewać Stalina. (101210)
- I_Action [podejrzewać] powinno łączyć się poprzez koreferencję z [nią] czyli [operacją] z poprzedniego zdania, ale jeszcze nie ma możliwości, żeby tego typu relację oznaczyć;
- Uprawianej, dodajmy, nie tylko przez młodych mężczyzn, ale również przez mocno posiwiałych, jak i przez młode kobiety. (101508)
- reporting [dodajmy] łączy się z [uprawianej]
- (...) a także firmy zajmujące się szkoleniem i certyfikacją profesjonalistów z firmami, które potrzebują tego rodzaju kadry. (101542)
- tu nie powinno się pojawić oznaczenie I_state [potrzebują], gdyż potrzeba dotyczy

obiekty, a nie sytuacji

- A postanowienie to kiedy wchodzi w wykonanie? / – Od chwili ogłoszenia. (101575)
 - nie wprowadzamy relacji pomiędzy reporting [ogłoszenia], a [postanowienie], gdyż przekraczamy granicę zdania
- – Przyjacielu Kerabanie... zaczął Van Mitten chcąc go odciągnąć. (101575)
 - w domyśle [zaczął] mówić, ale nie wprowadzamy relacji ze względu na elipsę
- - Dajże mi pokój, odrzekł gniewnie. (101575)
 - reporting [odrzekł] łączy się z action [dajże]
- Polacy mogą być w swoim kraju bezpieczni. (101583)
 - [być] jest nieznakowane, więc tutaj I_State się z niczym nie łączy
- (...) projekty ustaw, które mogą być od razu procedowane (101583)
 - [mogą] z [procedowane], gdyż [być] jako słowo posiłkowe nie jest znakowane, a imiesłów otrzyma atrybut "bierności"
- Chciałbym zadać jedno krótkie pytanie (101592)
 - I_State [chciałbym] łączy się bezpośrednio [zadać] - w przypadku konstrukcji z nominalizacjami funkcję bezpośredniego argumentu pełnią light predicate'y

Do specyfikacji dołączono anotowane przykłady z korpusu KPWr.

Atrybuty sytuacji

Atrybuty MAKEINSTANCE według TimeML

Event Instance ID number (eiid). Atrybut obligatoryjny. Każde wystąpienie sytuacji powinno mieć przypisany unikatowy numer ID. Zostanie on automatycznie przydzielony za każdym razem, gdy zostanie utworzony tag <MAKEINSTANCE>.

Event ID number (eventID). Atrybut obligatoryjny. Każde wystąpienie sytuacji jest powiązane z pewną sytuacją wyrażoną w tekście. Numer tej sytuacji zostanie automatycznie przydzielony za każdym razem, gdy zostanie utworzony tag <MAKEINSTANCE>.

czas (tense), aspekt (aspect), część mowy (pos). Atrybuty obligatoryjne. Zostaną one wprowadzone w etapie wstępnego przetwarzania, jednak należy je ręcznie poprawić lub uzupełnić w razie potrzeby. Poniżej znajdują się wskazówki do oznaczania czasu, aspektu i części mowy. Atrybuty czasu i aspektu wychwytyją standardowe rozróżnienia kategorii gramatycznych obecnych we frazach czasownikowych. Czas może przyjmować wartości: TERAŻNIEJSZY, PRZESZŁY, PRZYSZŁY, NONE. Aspekt może przyjmować wartości: DOKONANY, NIEDOKONANY, NONE. Atrybut części mowy wychwytyje rozróżnienia kategorii składniowych, do których należą językowe wykładniki sytuacji. Ten atrybut może przyjmować wartości: CZASOWNIK, RZECZOWNIK, PRZYMIOTNIK

Czasowniki w formie osobowej

Wszystkie czasowniki w formie osobowej będą miały część mowy = „CZASOWNIK”. Czas i aspekt należy znakować tak, jak w poniższych przykładach.

Strona czynna

czas="TERAŻNIEJSZY"

grupa czasowników aspekt=
uczy NIEDOKONANY

czas="PRZESZŁY"

grupa czasowników aspekt=
uczył NIEDOKONANY
nauczył DOKONANY

czas="PRZYSZŁY"

grupa czasowników aspekt=
będzie uczył NIEDOKONANY

Strona bierna

czas=TERAŹNIEJSZY

grupa czasowników

jest uczoney

aspekt=NIEDOKONANY

czas=PRZESZŁY

grupa czasowników

był uczoney

aspekt=NIEDOKONANY

został nauczoney

aspekt=DOKONANY

czas=PRZYSZŁY

grupa czasowników

będzie uczoney

aspekt=NIEDOKONANY

Rzeczowniki i przymiotniki

Atrybuty rzeczowników i przymiotników są podobne do atrybutów czasownikowych. Ze względu na swoje właściwości gramatyczne przymiotniki i rzeczowniki, które są wykładnikami sytuacji, nie będą miały wskazanego atrybutu czasu. Wyjątek stanowią te przymiotniki i rzeczowniki, będące częścią predykatów złożonych, których głową jest słowo posiłkowe "być" ("zostać", "stać się"...). W takim przypadku atrybut czasu przyjmuje wartość zgadzającą się z czasem gramatycznym słowa posiłkowego, np.

Pranie bardzo mnie nudziło

Pranie

czas=NONE

To jest jawne wypowiedzenie wojny

Wypowiedzenie

czas=TERAŹNIEJSZY

Atrybuty sytuacji

Ogólność (generality)

Niektóre sytuacje odznaczają się znaczną czasową rozciągłością i nie wiążą się z żadnym konkretnym momentem (odcinkiem) czasu. Gramatyka mówi o aktualnym i nieaktualnym użyciu czasów gramatycznych (Laskowski: 173), ale również w odniesieniu do form nominalnych, którym nie można przypisać wartości czasu, można stosować tego typu rozróżnienie denotowanych sytuacji.

W ramach tego atrybutu wprowadzamy dwie wartości: *generic* ('GEN') i *specific* ('SPEC'). Wartość *generic* odnosi się do sytuacji ogólnych tj. praw, zwyczajów, reguł, zasad etc. W ujęciu Laskowskiego (173) będą to sytuacje mające charakter:

- zwyczajowy
np. *Wiesiu dużo jeździ na rowerze, Andrzej kąpie się w soboty*
- wszechczasowy
np. *Ziemia krąży wokół słońca, Rejestr wyborców służy do sporządzania spisów wyborców*

Jest oczywiste, że tylko zdarzenia odbywające się w teraźniejszości i odbywające się w przeszłości są realnymi zdarzeniami podlegającymi falsyfikacji. Tej faktywności nie należy jednak mylić z ogólnością, bowiem również przewidywania mogą dotyczyć umiejscowionych w czasie sytuacji. Tak należy potraktować również sytuacje, do których, pomimo przewidywań, nie doszło. Tak więc wszystkie sytuacje konkretne, które można umieścić w czasie (bez względu na czas zajścia), powinny uzyskać wartość *specific*.

Sytuacje, które mają charakter *specific*, można przypisać do punktu w czasie, odcinka czasu lub pewnego określonego w czasie cyklu (np. *co dwa tygodnie, codziennie...*). Kiedy jednak występowanie sytuacji w czasie jest opisane w sposób mniej precyzyjny, przesuwa się ona w kierunku *generic* (np. *zawsze, często...*).

Istnieją gatunki tekstów, które zwiększają prawdopodobieństwo występowania sytuacji *generic* (np. przysłowia, ustawy...). Nie znaczy to jednak, że w takich tekstach nie mogą pojawić się sytuacje o charakterze *specific*.

Biegunowość (polarity)

W ramach tego atrybutu zostaną wprowadzone dwie przeciwległe wartości: *afirmacja* ('POS') i *negacja* ('NEG'). Negacja jest "składnikiem semantycznej struktury wyrażenia zdaniowego, za pomocą którego mówiący odrzuca wydany uprzednio (lub domyślny) sąd" (EJO: 390). Negacja może pełnić samodzielnie funkcję rematu w zdaniu, którego tematem jest cały wydany sąd. Są to zdania typu: *nieprawda, że...; jest fałszem to, że...; Zaprzeczam, jakobym...* Takiej negacji w tym momencie nie uwzględniamy. Na tym etapie wprowadzamy jako atrybut jedynie negację wewnętrzną, którą zawierają zdania przeczące. Inny podział negacji ma

charakter formalnosemantyczny i nie ma pokrywa się z podziałem na negację rematyczną i wewnętrzną. Jemu poświęcimy więcej miejsca poniżej.

Negacja zdaniowa

Na ogół za negację zdaniową rozumie się partykułę przeczącą przy orzeczeniu. Jest to podstawowa rozpatrywana przez nas forma negowania zdarzeń. Kategoria negatywności dzieli zbiór konstrukcji werbalnych na dwa podzbiory: bezpośrednio (a) lub niebezpośrednio (b) zaprzeczone i niezaprzeczone (c) (Saloni et al.: 157). Różnica pomiędzy konstrukcjami zaprzeczonymi bezpośrednio a konstrukcjami z pośrednim przeczeniem wiąże się z bezpośrednim lub pośrednim występowaniem partykuły "nie" przed formą wyrazową czasownikową. Ta druga możliwość jest realizowana w predykatkach złożonych, których oba elementy podlegają znakowaniu (J2, J3 i J4). W typowych sytuacjach w ich przypadku bezpośrednio negowany jest tylko jeden (nadrzędny składniowo) element, np.

- A. nie→[chcę]_{I_STATE} [iść]_{ACTION} do domu
- B. nie→[skończył]_{ASPECTUAL} [budować]_{ACTION} domu
- C. nie→[wykonał]_{LIGHT_PREDICATE} [skoku]_{ACTION}

Należy zauważyć, że zaprzeczenie sytuacji należącej do klasy I_STATE, I_ACTION bądź ASPECTUAL nie wiąże się z zaprzeczeniem pojawienia się sytuacji będącej argumentem (prawdziwość frazy A nie oznacza, że osoba mówiąca nie *szła* do domu, podobnie jak prawdziwość zdania B nie oznacza, że osoba mówiąca nie *budowała* domu). Wprowadzając atrybuty nie analizujemy jednak stanu faktycznego sytuacji pośrednio zaprzeczonych. Przypisywanie atrybutu biegunowości wiąże się z oceną, czy mamy do czynienia z bezpośrednią negacją, czy nie, tak więc powyższe przykłady powinny zostać następująco opisane:

- A. [chcę] polarity = 'NEG' [iść] polarity = 'POS'
- B. [skończył] polarity = 'NEG' [budować] polarity = 'POS'
- C. [wykonał] polarity = 'NEG' [skoku] polarity = 'POS'

Jeśli bezpośrednio zostało zanegowane słowo posiłkowe (które nie podlega znakowaniu w ramach EVENT), przypisujemy wartość atrybutu polarity = 'NEG' wykładnikom zanegowanym niebezpośrednio, np.

- D. nie→[będę]_{NIE ZNAKOWANE} [budować]_{ACTION} domu

- D. [budować] polarity = NEG

Negacja wyrazowa

Negacja wyrazowa w naszym przypadku dotyczy negowania: imiesłówów, gerundiów oraz (w sytuacjach wyjątkowych) przymiotników czy rzeczowników. *Może ona być wyrażona nie tylko za pomocą partykuły "nie", ale również innych wyrazów, np. "bez" (bez wdawania się w szczegóły...), "żaden" (żadnego głosowania...).*

Negacja ma charakter zewnętrzny, co znaczy, że wprowadzamy wartość *neg* dla wykładników sytuacji, których negacja została wyrażona innym słowem. Zaprzeczone rzeczowniki (nie) zgodnie z konwencją piszemy łącznie, mimo że jest zewnętrzne) powinny otrzymać wartość *neg* (łączna pisownia nie "uwewnętrznia" negacji), np. niemożność. Nie dotyczy to gerundiów i im nie przypisujemy atrybutu *neg*, gdyż zaprzeczone gerundia otrzymują taki tag morfosyntaktyczny. Uwaga! "Nienawidzić" to nie jest zanegowana sytuacja, nie ma sytuacji *"nawidzić".

Modalność (modality)

Modalność w szerokim rozumieniu to ustosunkowanie się mówiącego do treści *dictum* (EJO: 371).

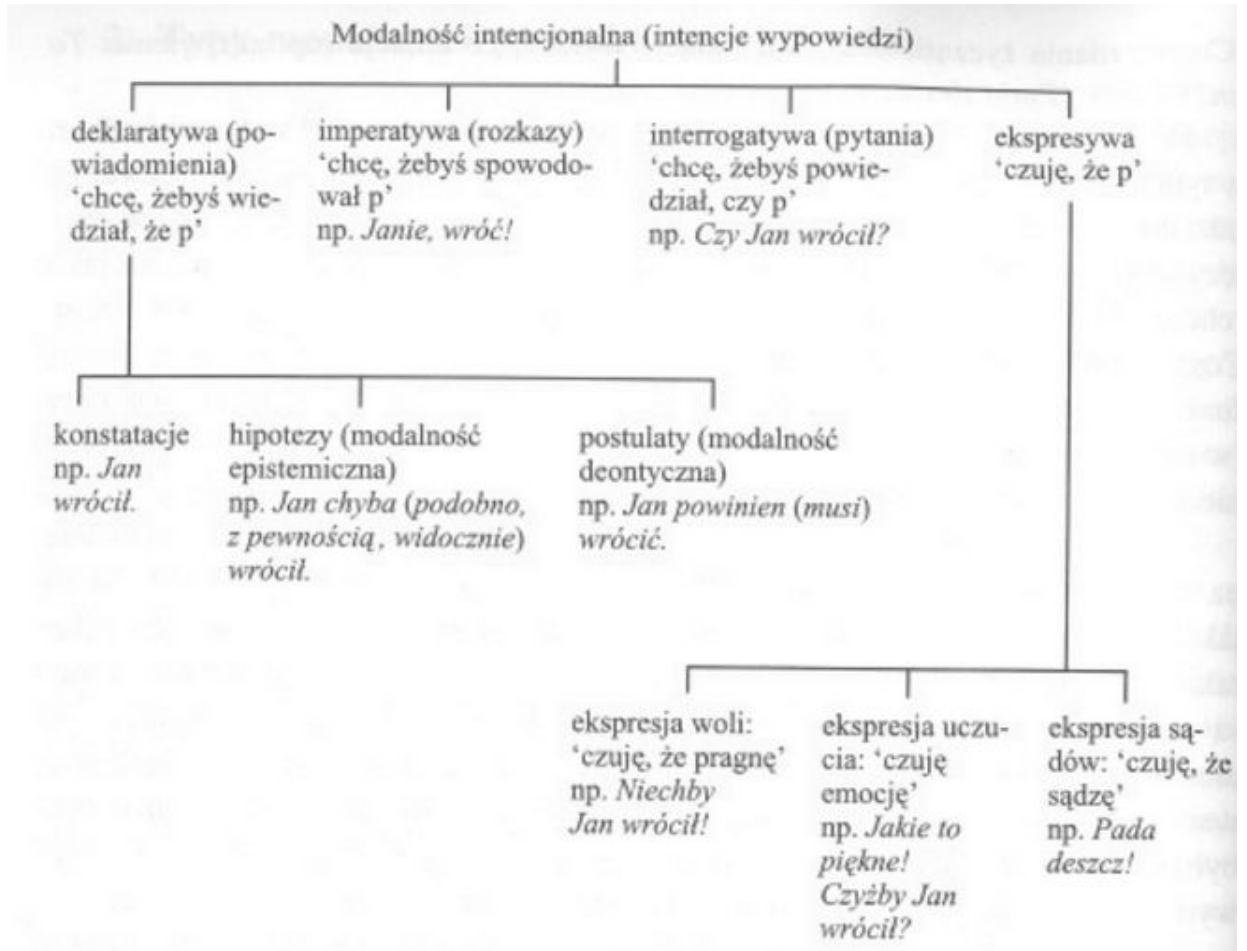
Środkami gramatycznymi reprezentującymi modalność, które mogą zostać automatycznie zinterpretowane, są tryby czasownika (tryb rozkazujący jest wykładnikiem ustosunkowania nakazującego; tryb przypuszczający jest morfologicznym wskaźnikiem nacechowanej modalności epistemicznej).

"Subiektywny stosunek mówiącego do tego, co jest treścią jego wypowiedzi (jak np. niepewność, powątpiewanie, przypuszczenie)" może być wyrażany również inaczej. W tekstach pisanych nie są dostępne środki prozodyczne. Interesujące dla nas będą za to środki leksykalne. Niektóre z czasowników wyrażają modalność, np. *musieć, chcieć, prosić, nakazywać...* i są oznaczane jako jedna z kategorii sytuacji (zwykle *I_State* lub *I_Action*) i nie podlegają znakowaniu jako "modality". Dodatkowym więc zadaniem będzie wskazanie w tekście słów, które nie zostały oznaczone jako sytuacja, a wyrażają modalność, np. *chyba, może, na pewno, niech, by..., moim zdaniem...* Takie fragmenty tekstu uzyskują anotację "modality". W dalszym etapie zostaną połączone z odpowiednią sytuacją w ramach określania atrybutów. Należy zwracać uwagę na to, czy rozpatrywany wykładnik modalności łączy się z wykładnikiem sytuacji (tylko wtedy podlega znakowaniu), czy jest określnikiem innego składnika (wtedy nie znakujemy), np.

Najprawdopodobniej zagrają z Realem - TAK (najprawdopodobniej → zagrają)

Zagrali najprawdopodobniej najslabszy mecz - NIE (najprawdopodobniej → najslabszy)

Modalność intencjonalna



(rys. za Grzegorzyczkowa R., Wprowadzenie do semantyki językoznawczej, Warszawa 2001, s. 126)

Główne typy modalności intencjonalnej:

1. deklaratywa
 - a. oznajmienie
 - konstatacja
 - hipoteza
 - *na pewno,*
 - *niewątpliwie,*
 - *z całą pewnością*
 - zdanie postulatywne
 - b. zdanie deklaratywne
2. imperatywa
 - a. rozkaz
 - b. zakaz
 - c. pozwolenie
 - d. prośba

- e. rada
 - f. ostrzeżenie
 - g. groźba
3. interrogatywa
- a. pytanie właściwe
 - o rozstrzygnięcie
 - Czy...?
 - uzupełniające
 - Kto...?, Kiedy...? etc.
 - sytuacje graniczne: pytania egzaminacyjne, hipotetyczne, deliberatywne
4. ekspresywa
- a. ekspresja pragnienia (zdania życzące bez angażowania odbiorcy)
 - *niechby, oby...*
 - b. ekspresja emocji (np. zachwyt, zdziwienie, irytacja),
 - c. ekspresja sądu (może także przyjąć formę oznajmienia lub pytania)

Modalność epistemiczna (prawdziwościowa)

Wypowiedzi tego typu mają charakter zdań hipotetycznych (hipotez), a ich przedmiotem jest „sąd nadawcy na temat prawdopodobieństwa zajścia zdarzenia komunikowanego przez składnik propozycjonalny” (Grzegorzczkova 2001: 132), „stan przekonaniowy nadawcy, jego przypuszczenie, pewność, wnioskowanie” (Grzegorzczkova 2001: 132).

1. Przykładowe wykładniki przekonania (nieczasownikowe):
 - *na pewno, z pewnością, niewątpliwie;*
 - *pewno, zapewne, chyba, pewnie, prawdopodobnie;*
 - *może.*
2. Przykładowe wykładniki prawdopodobieństwa (nieczasownikowe):
 - *podobno, widocznie;*
 - *jakoby, rzekomo.*
3. Partykuły o charakterze replikującym (będące reakcją na czyjąś wypowiedź):
 - *rzeczywiście, oczywiście.*
4. Czasownikowe wykładniki modalności epistemicznej (czasownikowe - nieznakowane jako "modality"¹):
 - *Przypuszczam, że..., Sądzę, że..., Jestem przekonany, że...*

Modalność deontyczna (wolicjonalno-oceniająca)

Ten typ modalności właściwy jest wypowiedziom informującym o „postawie wolitywno-oceniającej nadawcy”, funkcjonującym jako „zalecenia (dyrektywy), a nie informacje o istniejącym stanie rzeczy” (Grzegorzczkova 2001: 134). Wypowiedzi tego typu „mają więc za swój przedmiot bezpośredni przekonania aksjologiczne nadawcy [...] lub jego postawę

¹ Powinny zostać oznakowane jako odpowiednia kategoria sytuacji. Poniższe przykłady należą do I_State

wolitywną, intencję życzeniową wobec komunikowanego stanu rzeczy” (Grzegorzczkowska 2001: 134).

Przykładowe wykładniki modalności deontycznej

- *Koniecznie (np. Koniecznie idź...)*
-

Przykładowe wykładniki modalności deontycznej (czasownikowe - nieznakowane jako “modality”):

- *Trzeba..., Należy...*
- *Nie wolno..., Wolno...*

Z punktu widzenia TimeML interesujący jest szczególny rodzaj modalności. Wskazują na to następujące stwierdzenia z pracy *Annotating and Recognizing Event Modality in Text*:

Events in natural language discourse can be characterized along a veridicality axis that ranges from truly factual to counter-factual. Between these two extremes, there is a complex spectrum of modal types, expressing: Degrees of possibility, Belief, Evidentiality, Expectation, Attempting, Command.

Inferences derived from events that have not happened or have only possibly happened, are different from those derivable from events that are judged as factual in nature.

SSPs [Situation Selecting Predicates] are interesting because part of their lexical semantics is projected as modality information onto the event denoted by its argument (underlined in examples) by syntactic means. The event denoted by the argument is then marked as: Not totally certain, Certain according to a source, Factual, Counterfactual, Possibility.

Clausal and sentential adverbial modifiers may express similar modal information:

a. Possibility: probably, perhaps;

b. Frequency: usually, always.

Między innymi w związku z powyższym przyjmujemy, że typem modalności, który NIE jest dla nas interesujący z perspektywy ostatecznej funkcjonalności systemu jest modalność związana z wyrażaniem emocji

Czas (tense)

Atrybut sytuacji, jakim jest czas, zostanie rozpoznany automatycznie. Formy nominalne (poza gerundiami, które mogą wchodzić w paradygmat fleksyjny czasownika) w ramach tego atrybutu uzyskują wartość ‘NONE’.

Aspekt (aspect)

Atrybut sytuacji, jakim jest aspekt, zostanie rozpoznany automatycznie. Formy nominalne (poza gerundiami, które mogą wchodzić w paradygmat fleksyjny czasownika) w ramach tego atrybutu uzyskują wartość ‘NONE’.

Sposób anotacji

Kandydaci do atrybutów

1. *active* vs. *passive*
2. aktualność vs. nieaktualność
generality – specific or generic,
3. pozytywny vs. negatywny
polarity – affirmative or negative,
4. czy sytuacja X zaszła czy nie
za trudne na tym poziomie; część informacji przekazują relacje
5. zmiennostanowość vs. niezmiennostanowość
6. aspekt
automatycznie - tagger
7. czas
automatycznie - tagger
8. tryb (modalność) - tagger + funkcjonalność LEMa
9. modalność

Referencje

1. (EJO) *Encyklopedia językoznawstwa ogólnego* (1999), red. Kazimierz Polański, Wrocław: ZNiO
2. R. Grzegorczykowa (2001), *Wprowadzenie do semantyki językoznawczej*, Warszawa: Wydawnictwo Naukowe PWN
3. E. Jędrzejko (1993), *Nominalizacje w systemie i w tekstach współczesnej polszczyzny*, Katowice: Uniwersytet Śląski
4. E. Jędrzejko (2011), *The problematics of describing periphrastic predication. Between word and image*, "Studies in Polish Linguistics", vol. VI, <http://www.wuj.pl/UserFiles/File/SPL%206/3-SPL-Vol-6.pdf>
5. S. Jodłowski (1976), *Podstawy polskiej składni*, Warszawa: PWN
6. Kotsyba, N. (2007). Semantics of light verbs in lexicographical presentation. In: *Proceedings of the international conference 'Modern Problems of Lexicography' held in Grodno, Belarus, 19-21 October 2007*. Retrieved from: http://domeczek.pl/~natko/papers/nk_GrodnoX07.pdf
7. R. Laskowski (1998), *Kategorie morfologiczne języka polskiego - charakterystyka funkcjonalna*, [w:] *Gramatyka współczesnego języka polskiego. Morfologia*, cz. 1, Grzegorczykowa R., Laskowski R., Wróbel H. (red.), Warszawa: PWN.
8. K. Polański (red.) (1999), *Encyklopedia językoznawstwa ogólnego*, Wrocław: Zakład Narodowy imienia Ossolińskich
9. Z. Saloni, M Świdziński (1998), *Składnia współczesnego języka polskiego*, wyd. IV, Warszawa: PWN
10. Roser Saurí, Jessica Littman, Bob Knippen, Robert Gaizauskas, Andrea Setzer, James Pustejovsky (2006), *TimeML Annotation Guidelines Version 1.2.1*, http://timeml.org/site/publications/timeMLdocs/annguide_1.2.1.pdf
11. Z. Topolińska (1984), *Składnia grupy imiennej* [w:] *Gramatyka współczesnego języka polskiego*, red. Z Topolińska, Warszawa, s. 301-384