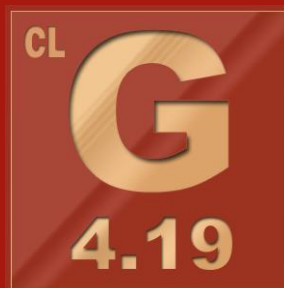




Politechnika Wrocławska

Wydobywanie wiedzy lingwistycznej z korpusów tekstu na potrzeby konstrukcji polskiej podstawowej technologii językowej

Maciej Piasecki, Bartosz Broda



Grupa Naukowa G4.19

Instytut Informatyki PWr.

nlp.pwr.wroc.pl

Grupa Naukowa G4.19

- Obszary działania:
 - inżynieria języka naturalnego
 - technologia językowa
 - lingwistyka informatyczna
 - semantyka formalna, lingwistyka i leksykografia
- Główny cel
 - budowa podstawowej technologii językowej dla języka polskiego dostępnej na licencjach otwartych
 - szerokie współdziałanie w tym celu
- Koordynator projektu CLARIN w Polsce
 - zastosowania inżynierii języka naturalnego w badaniach w dziedzinie nauk humanistycznych i społecznych

Oferta G4.19: morfologia

- **MACA** (Adam Radziszewski *i inni*)
 - = *Morphological Analysis Converter and Aggregator*
 - konfigurowalny moduł analizy morfologicznej i konwersji między tagsetami
 - podział na zdania i segmenty, wsparcie dla standardu SRX
 - łączy kilka dostępnych analizatorów morfologicznych, m.in. Morfeusz SGJP/SIAT
 - GNU LGPL 3.0

www.nlp.pwr.wroc.pl/narzedzia-i-zasoby

Oferta G4.19: ujednoznacznianie morfosyntaktyczne

- **TaKIPI** (2006) (Maciej Piasecki *i inni*)
 - pierwszy szeroko stosowany tager dla języka polskiego
 - działa w tagsecie Korpusu IPI PAN
 - łączy niewielki zbiór reguł ręczny z trzema poziomami około 200 drzew decyzyjnych
- **WCCL** (Adam Radziszewski *i inni*)
 - = *Wrocław Corpus Constraint Language*
 - formalizm wyrażen funkcyjnych wartościowanych na tekście oznakowanym morfo-syntaktycznie
 - reguły i ograniczenia: tagowanie, płytka analiza składniowa i semantyczna
- **WMBT** i **WCRFT** (Adam Radziszewski *i inni*)
 - tagery oparte na uczeniu pamięciowym i na technice warunkowych pól losowych (CRF)
 - wysoka trafność

Oferta G4.19: płytka analiza składniowa

- **Disaster** (Adam Radziszewski *i inni*)
 - = DISAmbiguator and STatistical chunkER
 - płytka analiza składniowa (*chunking*)
 - środowisko do testów różnych klasyfikatorów
- **IOBBER** (Adam Radziszewski *i inni*)
 - wersja wdrożeniowa, CRF
 - GNU LGPL 3.0
 - domyślna konfiguracja rozpoznaje frazy rzeczownikowe, czasownikowe i przymiotnikowe

Oferta G4.19: płytka analiza składniowa

Gdy już zdecydujemy się na optymalny dla naszego biznesu model auta,



kolejnym krokiem będzie wybór odpowiedniego napędu.



LEGENDA

VP

NP / PP

uzgodnione

Oferta G4.19: anotowane korpusy

- **KPWr** (Marek Maziarz *i inni*)
 - = Polish Corpus of Wrocław University of Technology
 - = pol. *Korpus Języka Polskiego Politechniki Wrocławskiej*
 - korpus dokumentów języka pisanego i mówionego na licencji Creative Commons
 - 14 kategorii tekstów (różne dziedziny, gatunki)
 - częściowa anotacja ręczna: metadane, płytka struktura składniowa, nazwy własne, powiązania anaforyczne, relacje semantyczne, ujednoznacznione znaczenia leksykalne dla wystąpień 80 lematów
- **WNG** (Michał Marcińczuk *i inni*)
 - korpus 797 wiadomości ekonomicznych
 - częściowa anotacja ręczna: nazwy własne



Oferta G4.19: KPWr - przykład z Inforex

Corpora

Download

NER (56nam)

Annotations

Events

Relations

Sense

Users

Użytkownik: Michał Marcińczuk

Opcje: wyloguj

Inforex

↓ Corpora » KPWr » wikipedia » Toronto Dominion Centre

Settings Documents Annotation map Relations Tests Statistics Word frequency list Add document

Flags: [Clean] [Tokens] [Names] [Names Rel] [Chunks] [Chunk Rel] [Chunk Heads] [WSD] [Anaphora] Options: [delete]

(0) <- pierwszy, -100, -10, < poprzedni | 1 z 1349 | następny > , +10, +100, ostatni > (1348)

View Document

View Content

Edit Metadata

View Anaphora

Edit Content

History of changes

Bootstrapping

Annotator

WSD

Edit Anaphora

TEI

Document content:

1 Toronto Dominion Centre

Toronto Dominion Centre - kompleks handlowo-kulturalny w kanadyjskim mieście Toronto, w Financial District.

Składa się z 3 czarnych budynków, zaprojektowanych przez architekta Ludwiga Mies van der Rohe.

Budynki tworzą odgrodzony od ulic dziedziniec, na którym Joe Fafard ustawił 6 odpoczywających krów z brązu.

Pomiędzy budynkami stoi także wielkie krzesło.

W południe odbywają się koncerty jazzowe.

W kompleksie znajduje się jedna z najważniejszych galerii sztuki Inuitów - Toronto Dominion Gallery of Inuit Art.

Toronto Dominion Centre

Toronto Dominion Centre - kompleks handlowo-kulturalny w kanadyjskim mieście Toronto, w Financial District.

Składa się z 3 czarnych budynków, zaprojektowanych przez architekta Ludwiga Mies van der Rohe.

Budynki tworzą odgrodzony od ulic dziedziniec, na którym Joe Fafard ustawił 6 odpoczywających krów z brązu.

Pomiędzy budynkami stoi także wielkie krzesło.

W południe odbywają się koncerty jazzowe.

W kompleksie znajduje się jedna z najważniejszych galerii sztuki Inuitów - Toronto Dominion Gallery of Inuit Art.

View configuration

Layer	Display			Show
	None	Left	Right	
Anaphora	☒	☐	☐	☐
Chunking	☐	☐	☐	☒
Proper names	☐	☐	☐	☒

Relation sets

☐ Syntactic relations

☐ Semantic relations

☐ Anaphora

☐ Semantic relations (rare)

☐ Semantic relations (anaphora)

Apply

☒ Display every sentence separately

Annotation list

Relation list

Oferta G4.19: zarządzanie i edycja korpusu

- **Inforex** (Michał Marcińczuk *i inni*)
 - system do zarządzania i edycji korpusów
 - wiele poziomów anotacji
 - anotacja ciągów oraz relacji, np. anafory
 - automatyczna: segmentacja i analiza morfosyntaktyczna dla języka polskiego

Oferta G4.19: płytka, statystyczna analiza semantyczna

- **SuperMatrix** (Bartosz Broda *i inni*)
 - kompleksowy system do wydobywania powiązań znaczeniowych z wielkich korpusów
- **Fextor** (G4.19)
 - narzędzie służące do ekstrakcji cech ze zbiorów tekstów
 - wskaźnik: token, anotacja lub para anotacji
- **WordnetWeaver** (Maciej Piasecki *i inni*)
 - system do półautomatycznego rozszerzania leksykalnej sieci semantycznej w oparciu o analizę korpusu

Oferta G4.19: zasoby leksykalne

- **NELexicon** (Michał Marcińczuk *i inni*)
 - gazetteer nazw własnych,
 - ponad 1,4 miliona unikalnych nazw własnych przypisanych do kategorii
 - ponad 1,37 miliona unikalnych napisów
- **Lista frekwencyjna** (Bartosz Broda *i inni*)
 - wydobyta za pomocą SuperMatrix
 - ze zbiorczego korpusu liczącego ponad 1,8 miliarda tokenów (m.in. *Korpus IPI PAN*, *Korpus Rzeczpospolitej*, *Wikipedia*)
- **Listy dystrybucyjnego podobieństwa semantycznego** (Maciej Piasecki, Bartosz Broda *i inni*)
 - oparte na analizie zbiorczego korpusu za pomocą SuperMatrix
 - rzeczowniki, czasowniki - po kilkadziesiąt tysięcy najczęstszych
 - dla każdego lista 20 najbardziej podobnych



Słownosieć (plWordNet) 2.0



Słownosieć 2.0 (na 2012) (Marek Maziarz,
Maciej Piasecki *i inni*)

plwordnet.pwr.wroc.pl

- leksykalna sieć semantyczna inspirowana Princeton WordNet
- części: rzeczowniki, czasowniki, przymiotniki
- stan obecny:
151 000 jednostek leksykalnych, 110 000 synsetów 100 000 lematów, 425 000 instancji relacji semantycznych
- N N516 068637 finansowany przez **MNiSW**
- POIG.01.01.02-14-013/09 **UE** oraz **MNiSW**
- SyNaT projekt strategiczny - **NCBiR**



NARODOWA
STRATEGIA SPÓJNOŚCI

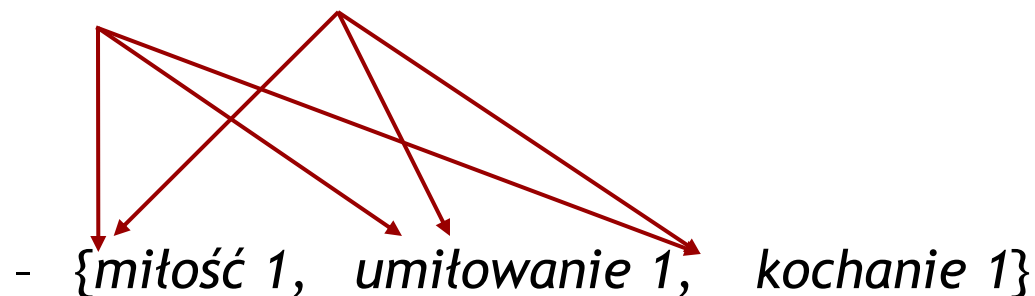
UNIA EUROPEJSKA
EUROPEJSKI FUNDUSZ
ROZWOJU REGIONALNEGO



Słownosieć: podstawowe założenia

- Rola synsetu

- grupuje jednostki leksykalne o wspólnych relacjach hiperonimii i holo/meronimii
- jednostki zawarte w synsecie są uznawane za synonimy
- jest rodzajem skrótu notacyjnego, np.
- {*afekt 1*, *uczucie 2*} —hiperonim→



- Relacje konstytutywne

- podstawa konstrukcji synsetu

- Dodatkowe rozróżnienia: rejestr stylistyczny, aspekt

Słownosieć: leksykalne relacje semantyczne

- Relacje pomiędzy synsetami
 - 21 relacji głównych, wiele podtypów
 - np. hiperonimia/hiponimia, holonimia/meronima, bliskoznaczność, kauzacja, presupozycja, określnik, wartość cechy, gradacyjność, itd..
- Relacje pomiędzy jednostkami leksykalnymi
 - 20 relacji, wiele podtypów
 - wiele relacji motywowanych derywacyjnie
 - np. antonimia, charakteryzowanie, predyspozycyjność, zawieranie roli, rola, nacechowanie, itd.

Oferta G4.19: technologia dwujęzyczna

- **Słownosieć 2.0 rzutowana na WordNet 3.0** (Ewa Rudnicka *i inni*)
 - powiązania pomiędzy synsetami
 - dobór par na zasadzie podobieństwa grafów wordnetowych relacji
 - hierarchia relacji międzyjęzykowych
 - stan obecny: około 23 000 synsetów zrzutowanych
 - licencja: otwarta dla wszystkich zastosowań
- **RL-button** (Paweł Kędzia i Kondrad Przybycień)
 - narzędzie do automatycznego rzutowania wordnetu na inny wordnet
 - reimplementacja algorytmu (Daudé, Padró i Rigau, 2003)
 - zastosowane na wielką skalę przy rzutowaniu Słownosieci

[illegible]

Oferta G4.19: ujednoznacznianie znaczeń leksykalnych

- **LexCSD** (Bartosz Broda *i inni*)
 - = Lexicographer-Controlled Semi-automatic Sense Disambiguation
 - system do półautomatycznego ujednoznaczniania znaczeń leksykalnych słów
 - automatyczna identyfikacja znaczeń
 - wydobywanie charakterystycznych przykładów użycia dla znaczenia

Oferta G4.19: rozpoznawanie nazw własnych

- **Liner2** (Michał Marcińczuk *i inni*)
 - narzędzie do rozpoznawania i znakowania nazw własnych w tekście
 - rozpoznanie lub rozpoznanie i klasyfikacja do: 5 lub 56 kategorii nazw własnych
- **NamEnSim** (Jan Kocoń *i inni*)
 - = Named Entity Similarity Function
 - funkcja podobieństwa nazw własnych
 - narzędzie do wyznaczania podobieństwa par słów, wykorzystujące heterogeniczną funkcję podobieństwa nazw własnych



Oferta G4.19: rozpoznawanie nazw własnych NER-WS

Corpora
Sense

Download

NER (56nam)

Activities

Annotations

Events

Relations

Użytkownik: Michał Marcińczuk
Opcje: [wyloguj](#)

Inforex

Automatic proper names recognition — under development

1 CRF-based model for recognition of 56 categories of proper names in Polish texts.
The description of the base model can be found in [Rich Set of Features for Proper Name Recognition in Polish Texts](#).

2 Usługa NER jest hostowana na serwerach Poznańskie Centrum Superkomputerowo-Sieciowe w ramach współpracy przy realizacji projektu SYNAT finansowanego przez Narodowe Centrum Badań i Rozwoju (numer grantu SP/I/1/77063/10).

Text to analyze:

Rzecznik Izby Celnej w Warszawie Piotr Tałaj, pytany czy podjęto w Warszawie kontrole punktów Totolotka, podkreślił, że w ramach działań Izby "prowadzone są również kontrole w podmiotach urządzających i prowadzących zakłady wzajemne". Zastrzegł, że nie może ujawniać informacji na temat toczących się kontroli.

Zgodnie z rozporządzeniem ministra finansów ustalenie, czy Totolotek działa, nie wymagało wszczynania kontroli. Każda firma organizująca zakłady wzajemne jest zobowiązana do przekazywania właściwym Izbowi Celnym "Miesięcznej informacji o sumie wpłat, sumie wypłaconych lub wydanych wygranych (...)" we wszystkich swoich punktach.

Stelmachowska, pytana dlaczego po odebraniu zezwoleń w marcu, ministerstwo nie podało decyzji do publicznej wiadomości, wyjaśniła, że „jest to rozpatrywane”. Zaznaczyła jednak, że firmę chronią tajemnice postępowań. „Prawnie musi być wszystko sprawdzone, by nikogo nie skrzywdzić” - dodała.

Text after analysis:

Rzecznik Izby Celnej w Warszawie: Piotr Tałaj

pytany czy podjęto w Warszawie: kontrole punktów Totolotka, podkreślił, że w ramach działań Izby "prowadzone są również kontrole w podmiotach urządzających i prowadzących zakłady wzajemne". Zastrzegł, że nie może ujawniać informacji na temat toczących się kontroli.

Zgodnie z rozporządzeniem ministra finansów ustalenie, czy Totolotek: działa, nie wymagało wszczynania kontroli. Każda firma organizująca zakłady wzajemne jest zobowiązana do przekazywania właściwym Izbowi Celnym "Miesięcznej informacji o sumie wpłat, sumie wypłaconych lub wydanych wygranych (...)" we wszystkich swoich punktach.

Stelmachowska, pytana dlaczego po odebraniu zezwoleń w marcu, ministerstwo nie podało decyzji do publicznej wiadomości, wyjaśniła, że „jest to rozpatrywane”. Zaznaczyła jednak, że firmę chronią tajemnice postępowań. „Prawnie musi być wszystko sprawdzone, by nikogo nie skrzywdzić” - dodała.

List of recognized names:

- CITY_NAME
 - Warszawie
 - Warszawie
 - Prawnie
- COMPANY_NAME
 - Totolotek
- INSTITUTION_NAME
 - Rzecznik Izby Celnej
- ORGANIZATION_NAME
 - Izbowi Celnym
- PERSON_FIRST_NAME
 - Piotr
- PERSON_LAST_NAME
 - Tałaj
 - Stelmachowska
- PERSON_NAME
 - Piotr Tałaj

Sample texts: [informacja prasowa](#), [powołanie prezesa](#), [przystąpienie do ESPI](#).

[Recognize names »](#)

Processed in 3 sec(s)

G4.19 - prace w toku

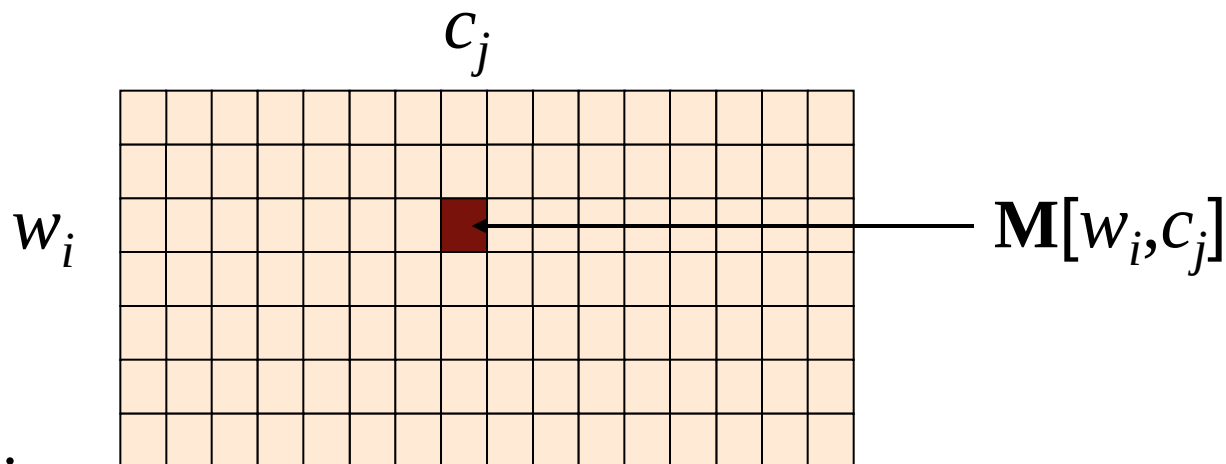
- Rozpoznawanie relacji
 - składniowych
 - semantycznych
 - pomiędzy fragmentami tekstu
- Rozpoznawanie sytuacji jako struktur relacji w tekście
- Pogłębiana analiza semantyczna
- System odpowiedzi na pytania w języku polskim
 - pytania proste różnych typów
 - pytania kontekstowe
- Prototyp węzła sieciowego technologii językowych opartego na zbiorze usług sieciowych

Wydobywanie wiedzy lingwistycznej: leksykalna sieć semantyczna

- Podobieństwo użycia wyrażen językowych
 - podobieństwo znaczenia
- (paradygmat dystrybucyjny)
- Wyznaczniki relacji leksykalnych
 - konstrukcje językowe sygnalizujące konkretną relację
- (paradygmat wzorców)
- Klasyfikacja par lematów w oparciu o metody maszynowego uczenia się
 - różnorodne cechy, w tym oparte na wzorcach i dystrybucji korpusowej

Miara powiązania znaczeniowego

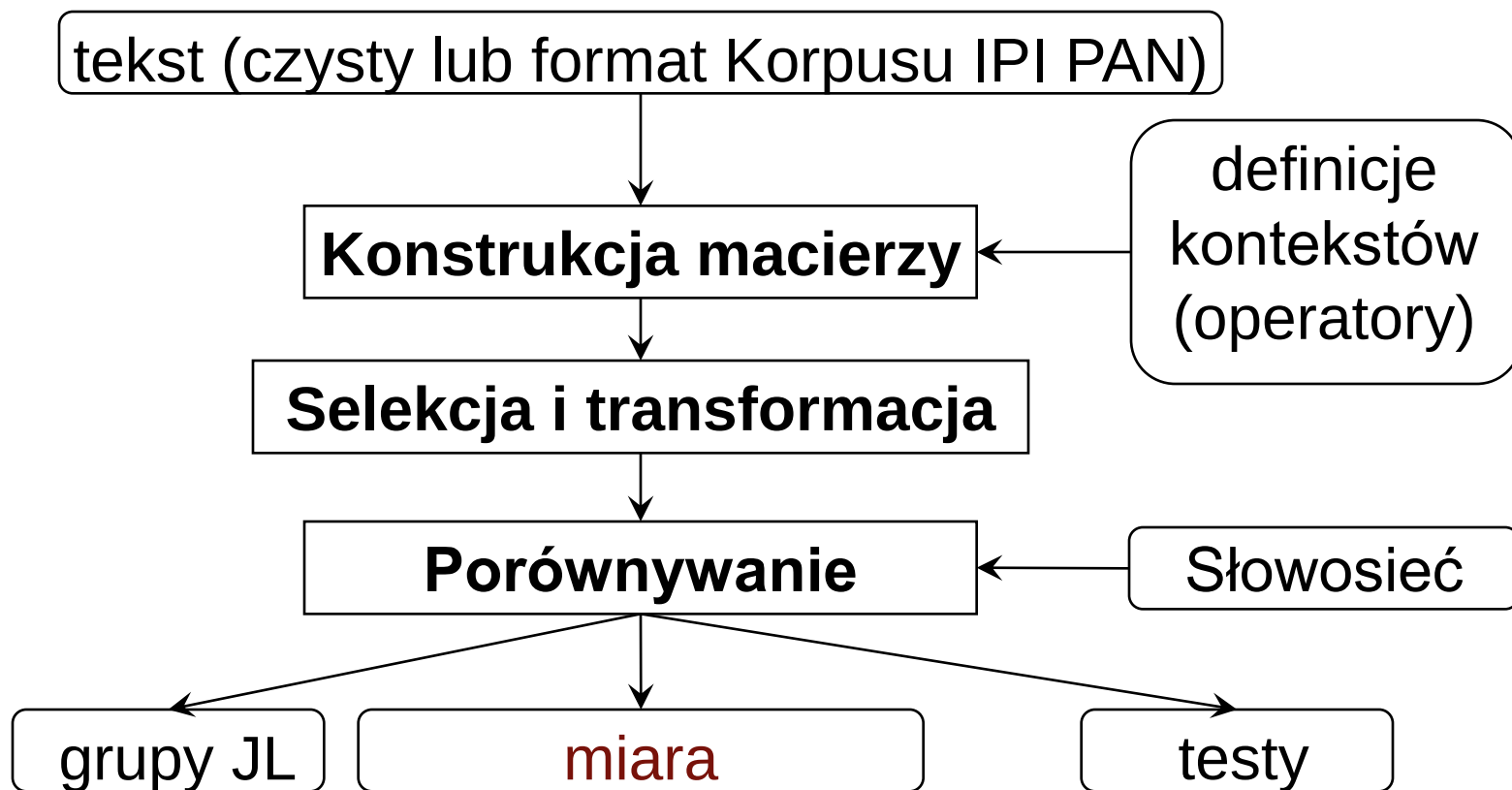
- Wzorzec dystrybucji → konteksty użycia
- Macierz koincydencji



- gdzie:

- w_i - lemat
- c_j - cecha powiązana z określonym kontekstem
- $\mathbf{M}[w_i, c_j]$ - częstość wystąpienia w określonym kontekście

Miara powiązania znaczeniowego: schemat wydobywania (*SuperMatrix*)



$$MPZ: L \times L \rightarrow R$$

Miara powiązania znaczeniowego: przykład

- **subst:wachlarz**

- | | | | |
|------------|--------------------|-------------|-----------------|
| • 0.150451 | subst:spektrum | • 0.107077 | subst:wielość |
| • 0.131427 | subst:gama | • 0.104133 | subst:zasłona |
| • 0.128613 | subst:parasolka | • 0.104081 | subst:peleryna |
| | | • 0.102885 | subst:kapelusz |
| • 0.123844 | subst:paleta | • 0.0984388 | subst:parasol |
| • 0.112755 | | • 0.0980375 | |
| | subst:różnorodność | | subst:tarnowian |
| • 0.112449 | subst:suknia | • 0.0973776 | subst:płaszcz |
| • 0.111019 | subst:wstążka | • 0.0970713 | subst:dywan |
| | | • 0.0965161 | subst:szereg |
| • 0.110588 | subst:szal | • 0.0963453 | subst:szata |
| • 0.110178 | subst:koronka | | |

WordnetWeaver i algorytm dołączania poprzez regiony aktywacji (ADRA)

- Założenia
 - celem jest rozszerzanie wordnetu, a nie konstrukcja od podstaw
 - jednolita reprezentacja źródeł wiedzy
- Źródła wiedzy: $K_1, \dots, K_s$
 - $K_i = \{ \langle l_n, l_j, w \rangle :$
 - l_n - nowy lemat, nieopisany w wordnecie,
 - l_j - lemat z wordnetu
 - w - waga lokalna - dla pary }
 - $waga(K_i) \in (0, 1]$ - dla źródła danych
 - np. oparta na ręcznie ocenionej dokładności K_i
- Trójka ze źródła danych wskazuje na region wordnetu, a nie na konkretny synset.

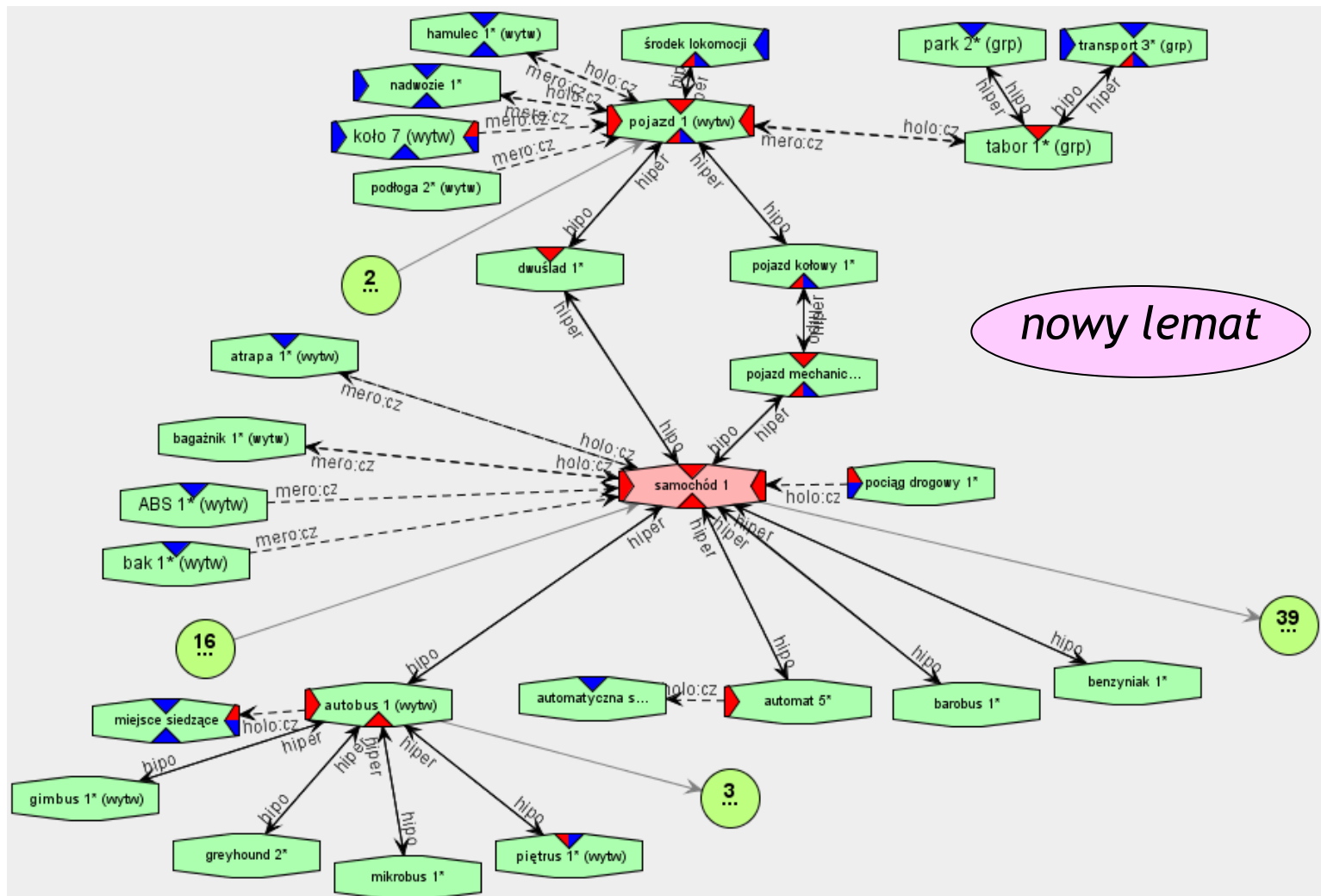
WordnetWeaver i algorytm dołączania poprzez regiony aktywacji

- Łuki wordnetu jako połączenia wzdłuż których powielane jest wsparcie
- Transmitancja
 - funkcja: $f_T(\text{relacja}, \text{wsparcie}) = \text{zmodyfikowane wsparcie}$
 - zdolność łuku do przenoszenia wsparcia
- Impedancja
 - $f_I(\text{relacja_wej}, \text{relacja_wyj}, \text{wsparcie}) = \text{zmodyfikowane wsparcie}$
 - opisuje do jakiego stopnia pośrednie wsparcie może być powielane przez połączenie łuków
 - np. holonimia->meronimia =
 - *samochód*-holonim-*szyba przednia*-meronim:substancja-*szkło*
 - pośrednie wsparcie może być powielone od *samochodu* do *szkła*

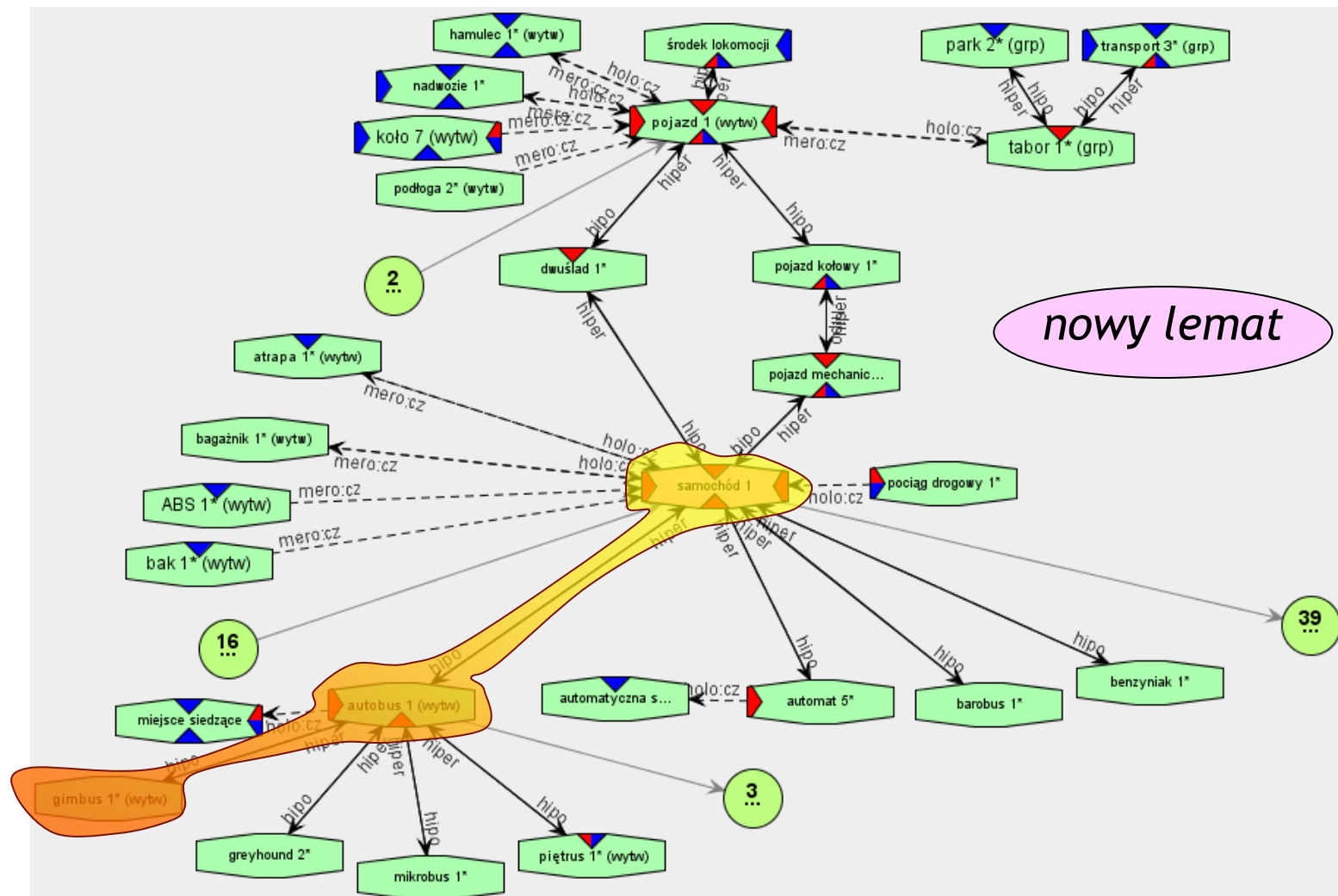
Algorytm dołączania poprzez regiony aktywacji: główne kroki

- Dla nowego lematu x i każdego lematu $y \in \text{SłowoSieci}$
- Gromadzenie wsparcia lokalnego
 - ze źródeł wiedzy w celu oceny powiązania x z lematami wordnetu
 - semantyczne dopasowanie = schemat głosowania ważonego: lokalne i/lub globalne wagi ze źródeł wiedzy
- Powielanie semantycznego dopasowania zgodnie ze strukturą wordnetu
 - bezpośrednie wsparcie dla synsetu S pochodzące z lematów synsetu
 - pośrednie wsparcie dla S : z synsetów połączonych z S ścieżkami w grafie wordnetu
 - własności strukturalne: *transmitancja* i *impedancja*
- Rozpoznawanie regionów dołączania x
 - regiony aktywacji: podgrafy wordnetowe z semantycznym dopasowaniem powyżej pewnego poziomu
 - regiony dołączania - z najwyższym dopasowaniem, wybrane jako sugestie dla lingwistów

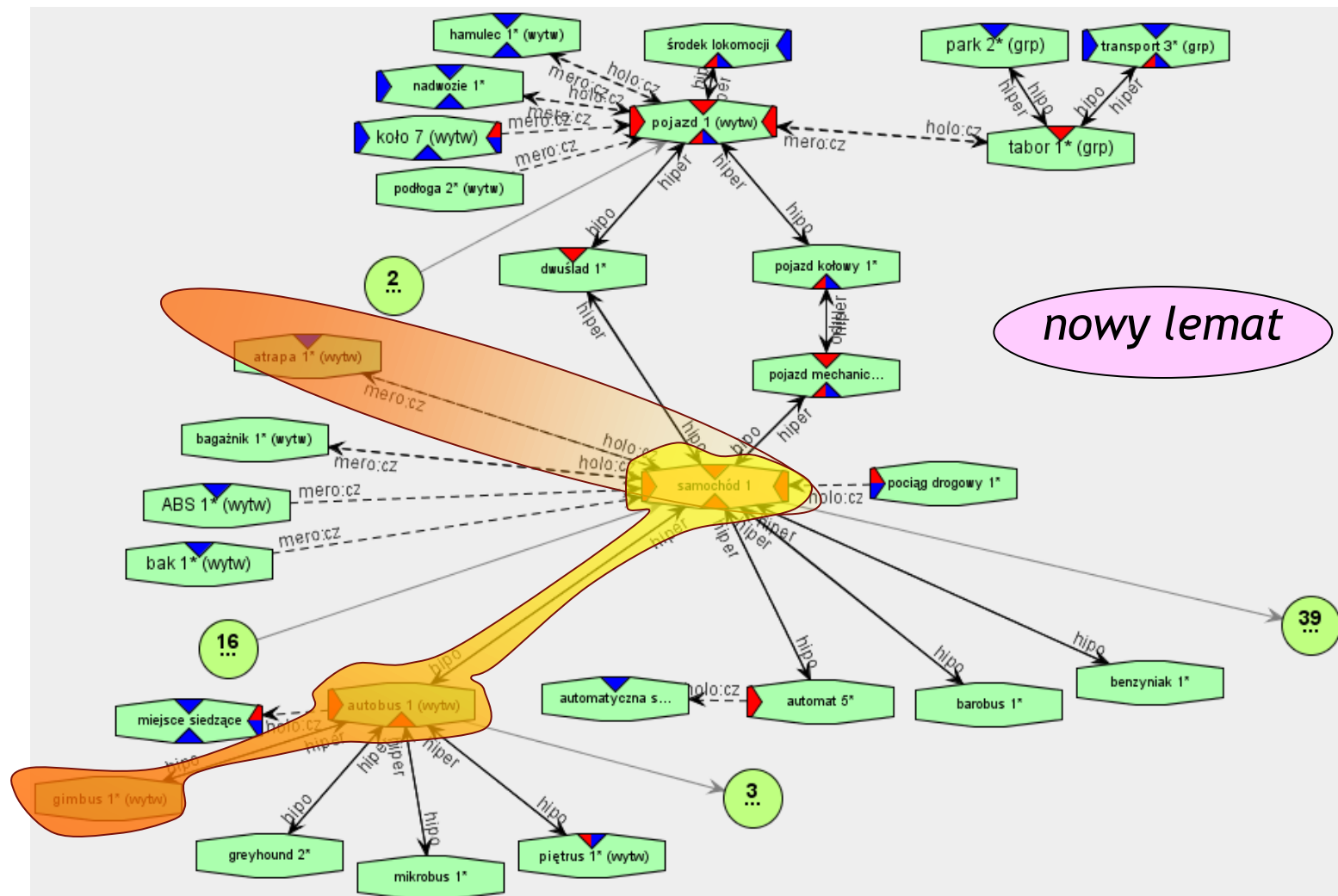
WordnetWeaver: stan początkowy



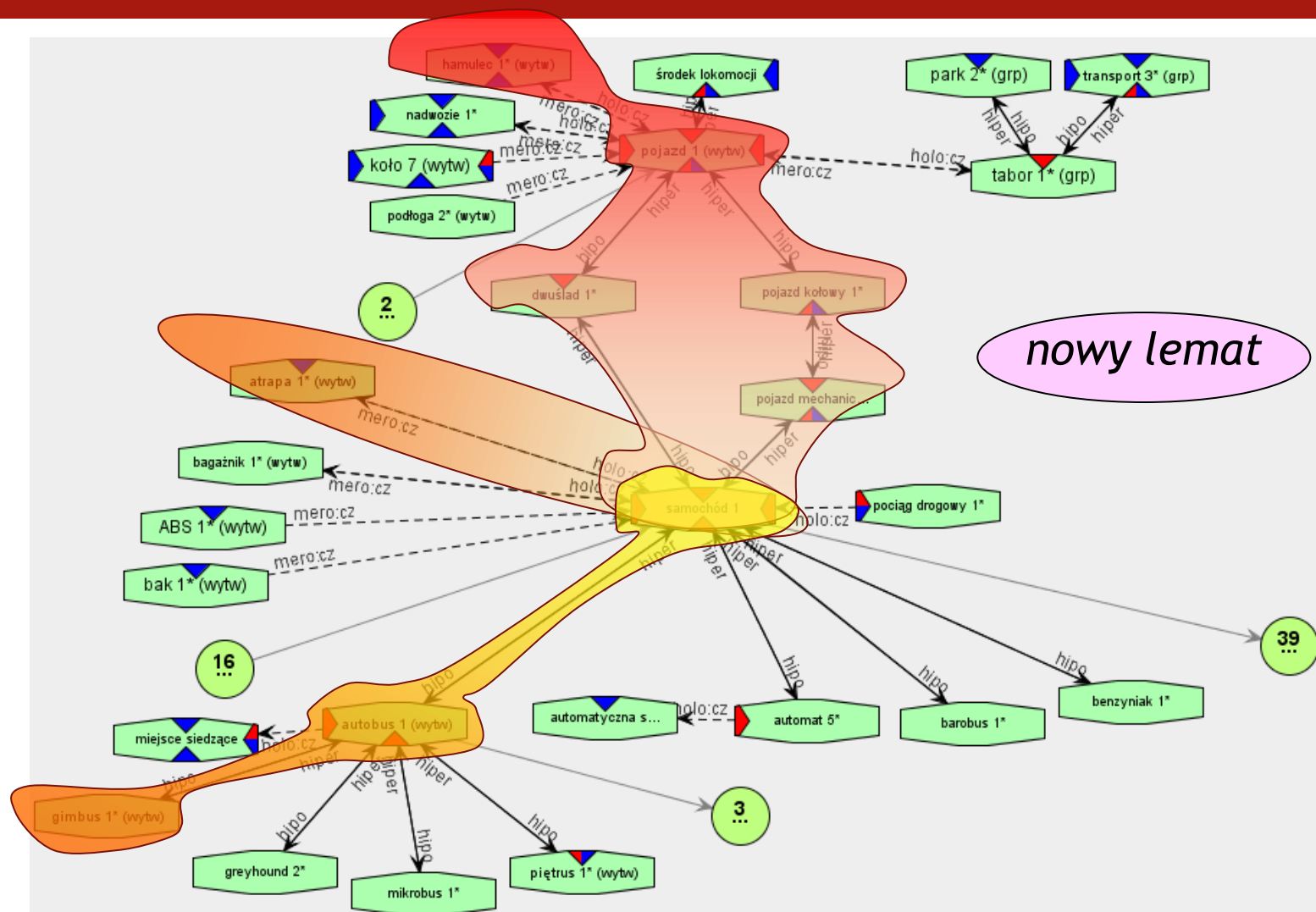
WordnetWeaver: gromadzenie wsparcia i powielanie semantycznego dopasowania



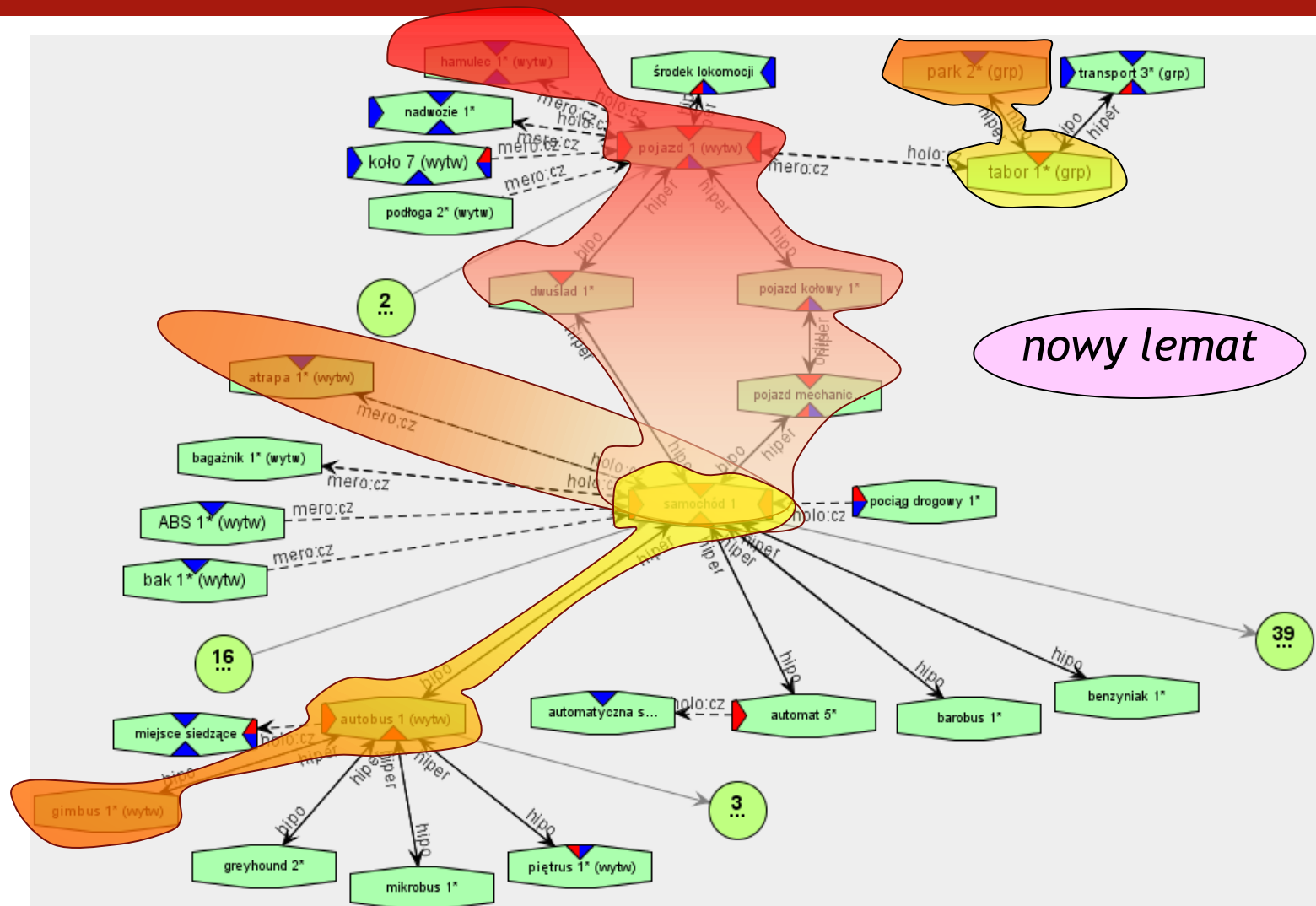
WordnetWeaver: gromadzenie wsparcia i powielanie semantycznego dopasowania



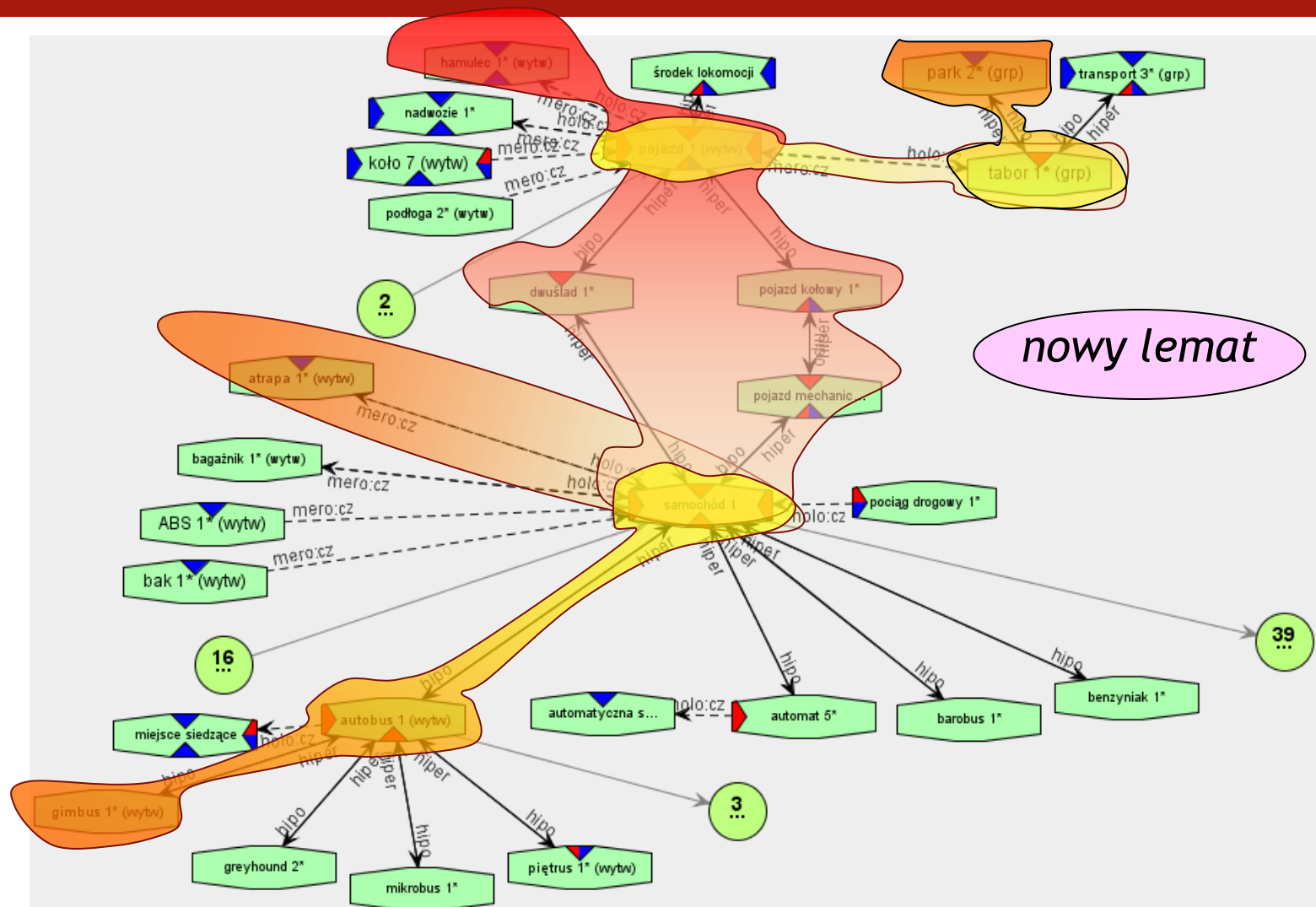
WordnetWeaver: gromadzenie wsparcia i powielanie semantycznego dopasowania



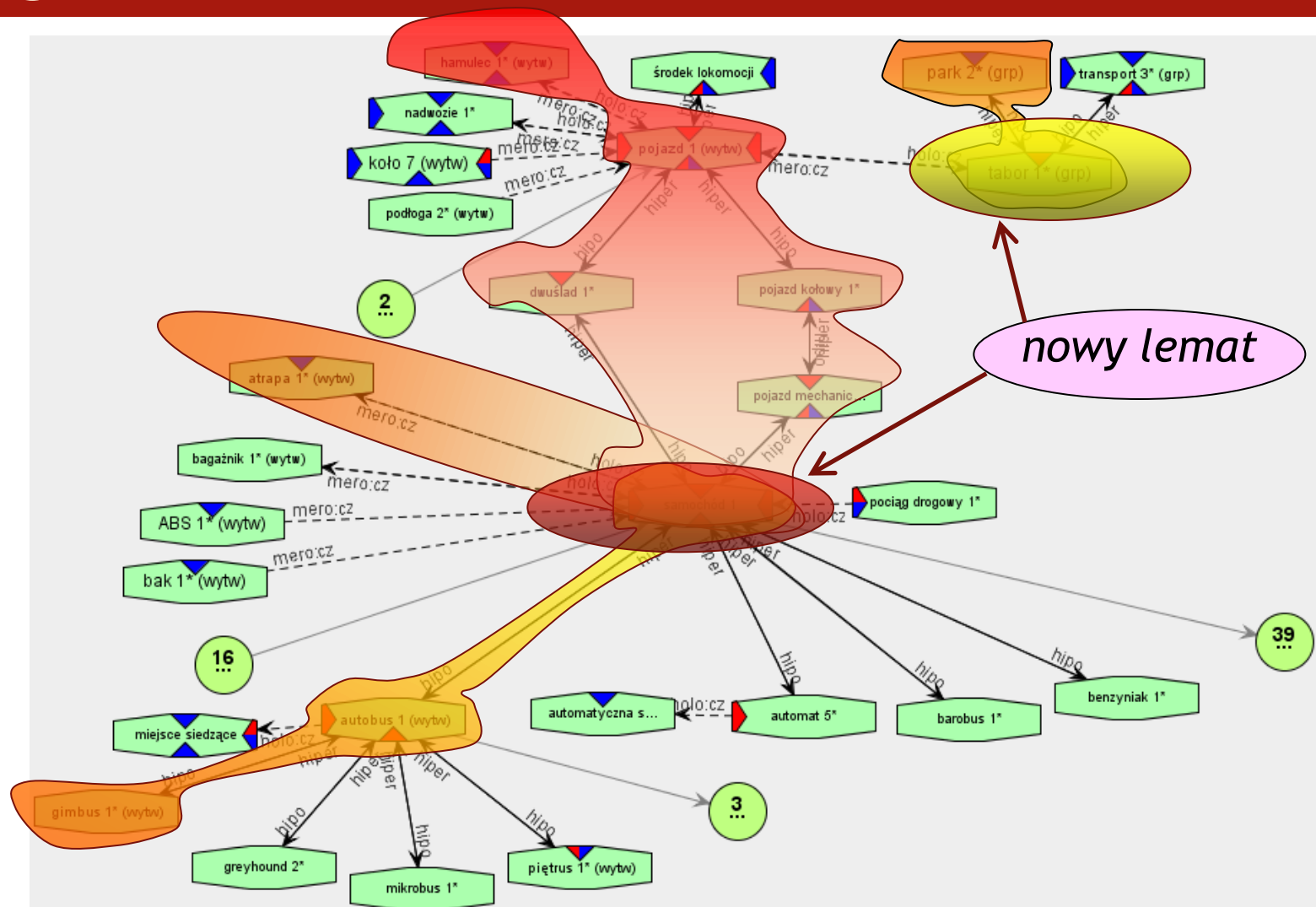
WordnetWeaver: gromadzenie wsparcia i powielanie semantycznego dopasowania



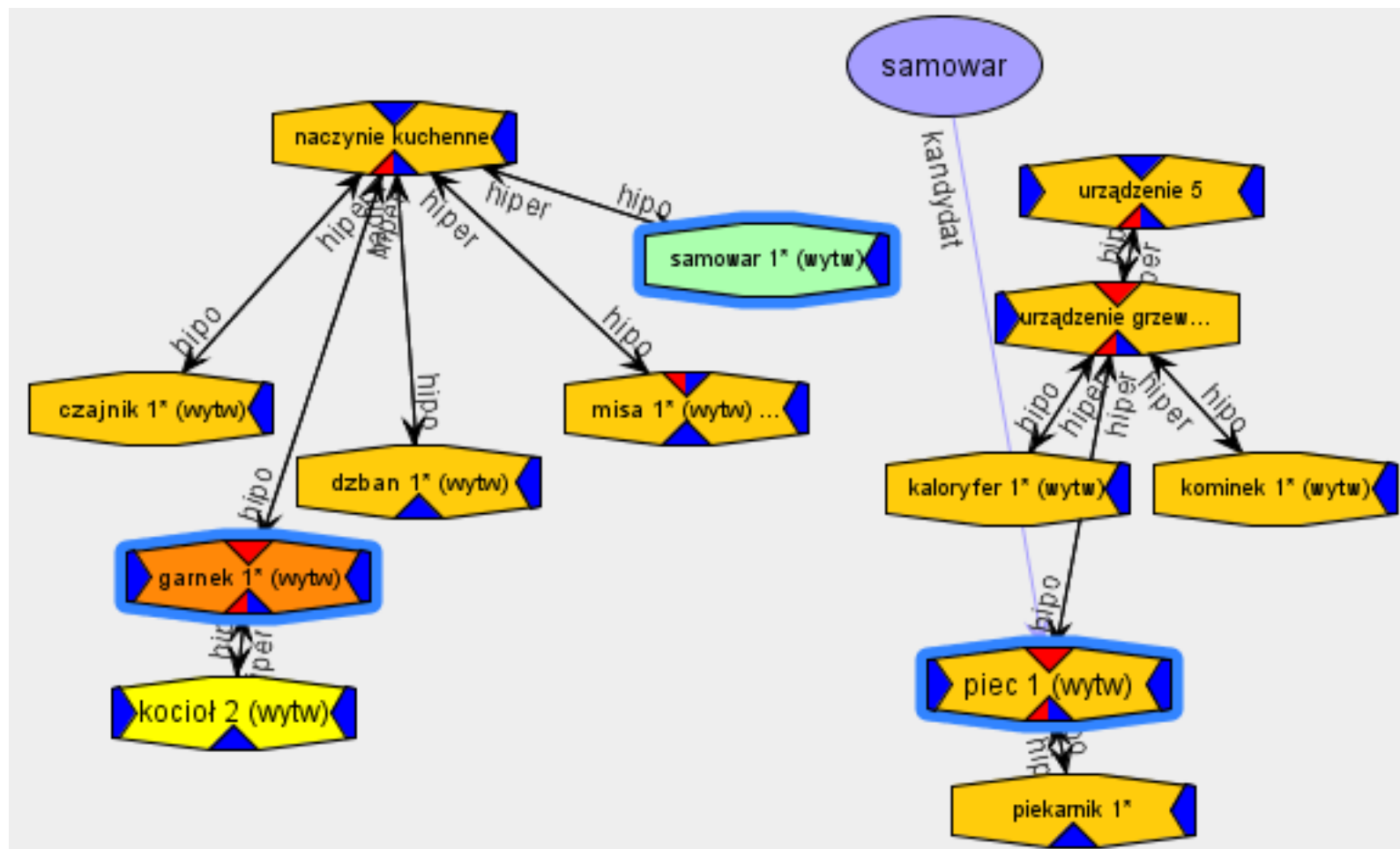
WordnetWeaver: gromadzenie wsparcia i powielanie semantycznego dopasowania



WordnetWeaver: rozpoznawanie regionów dołączania



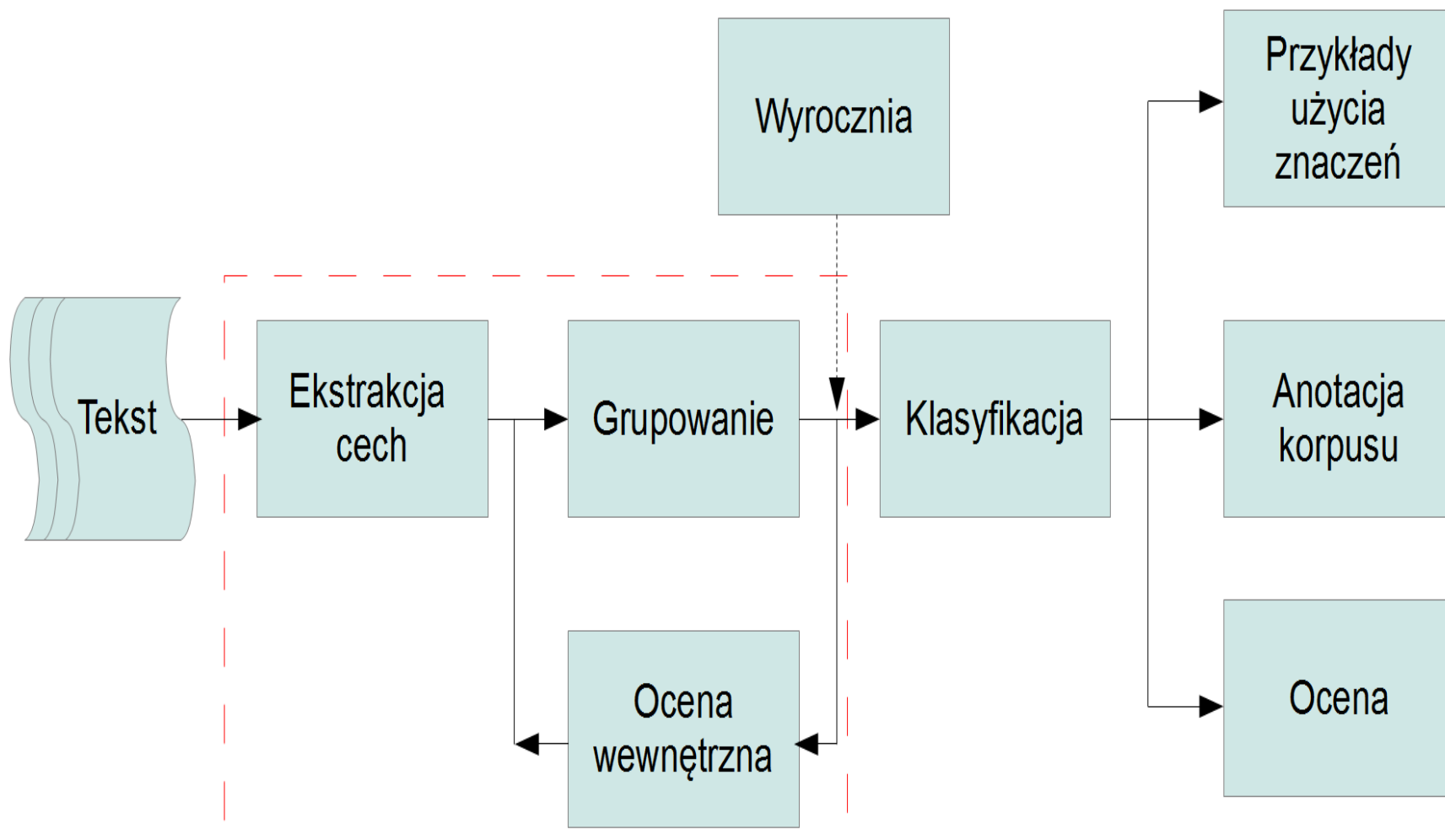
WordnetWeaver: przykład



LexCSD

- = Lexicographer-Controlled Semi-automatic word sense Disambiguation
- Podejście zostało oparte na sposobie pracy leksykografów
 - Zbieranie przesłanek z korpusów tekstów w postaci przykładów użycia danego słowa
 - Intuicyjny podział zebranych danych na grupy
 - Analiza grup w poszukiwaniu wspólnych właściwości przykładów użycia słowa w grupach
 - Sformułowanie definicji dla znaczeń
- Zalety metody
 - Minimalny nadzór (brak nadzoru)
 - Automatyczne wyznaczanie liczby znaczeń
 - Automatyczny wybór metody grupowania dla słów
 - Reprezentatywne przykłady użycia
 - Klasyfikatory do ujednoznaczniania niewidzianego wcześniej tekstu

LexCSD - schemat metody



Przykład - kleić

- W domach jednorodzinnych drzwi wewnętrzne najczęściej wykonane są z drewna litego lub **klejonego**.
- Oczy mu się **kleiły**; zaczął szukać po omacku butelki.
- Rozmowa się zbytnio nie **kleiła**. Nic dziwnego.
- Z meczu na mecz gramy coraz lepiej. Gra się nam **klei**, zgrywamy się, stwarzamy dużo sytuacji bramkowych.



LexCSD - zastosowanie w WordnetLoom

WordnetLoom 1.48 (Centralna - Robocza) NEKST

Plik Edycja Perspektywa Inne Okno Pomoc

Relacje jednostek Relacje synsetów Edycja synsetów Wizualizacja w...

Jednostki Kandydaci

Część mowy: Czasownik
Numer paczki: 116 (491)
Przelicz

Wyszukaj: Szukaj

Kandydaci:
kołatać się
najść
nasunąć się
przeminać
rozwiązać się
szczepać
skojarzyć się
spełniać
świtać
tać się
ulecieć
uolnić
uolnić się
wywietrzeć
zagościć
zaświtać
złagodnieć

Podgląd

Synset

Jednostki w synsencie:
ucieć 1 (ruch)
ujść 1* (ruch)
wziąć nogi za pas 1* (ruch)
zbiec 2* (ruch)

Komentarz:
brak danych

Właściciel: Aleksandra Pawlikowska

Relacje jednostki

ucieć 1 (ruch)
Od
aspektowość czysta DK-NDK
uciekać 1 (ruch)
Do
aspektowość czysta NDK-DK
uciekać 1 (ruch)

Przykłady Schowek

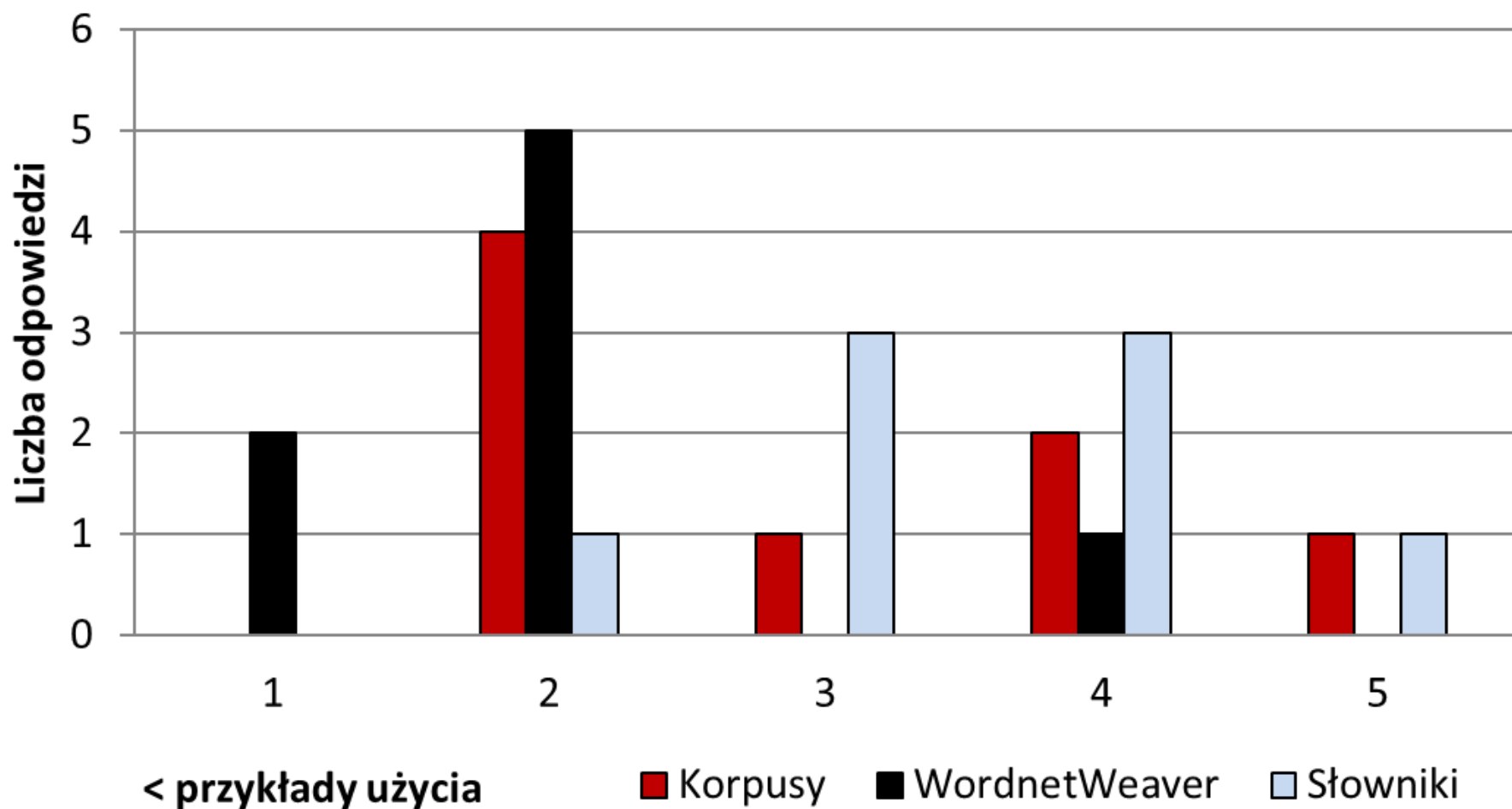
by z nią porozmawiać i zwierzyć się ze swych szalonych myśli. Gdy tylko zaczął mówić, rzekła: „Wiem, co chcesz powiedzieć. Już dawno temu też **naszły** mnie takie myśli, ale nie chciałam się nimi z tobą dzielić, bo wiem, że nie wierzysz w miłość. Miłość chyba istnieje, tylko jest czymś mniej nie codziennie człowiek dostarcza sobie kulturalnej stymulacji więc dziś będzie o czym innym, mam nadzieję że nie stracę przez to moich regularnych czytelników o ile tacy istnieją. **Naszła** mnie dziś myśl że w sieci funkcjonuje pod kilkoma nickami - co więcej nie mają one wiele wspólnego z tym kim jestem. W jednym portalu społecznościowym mam nick po młodego. Ku temu rzekł pan Bojanowski. Wątpię ja, aby kto nalaż się taki. Dołożył pan Dersniak: Mężczyzna trudno, a białogłowa żadną miarą się nie **najdzie**. Na to pan Myszkowski powiedział: Wdy w. m., panie Dersniaku, pana Bojanowskiego wydać niechcesz, abyś nie miał białyngłowam podnosków dawać, a one tego schematu? — Jest to po prostu pierwsza rzecz, która przyszła mi na myśl — powiedział Spence. — Być może na przyjęciu był ktoś, kogo nagle **naszła** przemożna chęć morderstwa. Być może zrobił już kiedyś coś takiego, a może tylko miał na to ochotę. Chcę przez to powiedzieć, że być może w przeszłości Dobrze - powiedział szybko. - Zgadza się, że Carewe'a trzeba uciszyć, ale jej nie. Zabierzmy... zabierzmy ją ze sobą. - Co cię **naszło**. Manny? - Przecież ona jest w ciąży - wykrztusił Pleeth. - I co z tego? - spytał Barenholm, lekko marszcząc czoło. - Nie jesteś...
- Właśnie. Dlatego zlekceważyłem go po przebudzeniu. Wiedziałem, że przez krótki czas posługiwałem się Mocą, ale pomyślałem, że po prostu **naszedł** mnie ten sen i że wszystko potem mi się przyśniło. No więc w tym śnie, w ten dziwny sposób właściwy snom, nagle zjawił się obok mnie Młotek jak smagnięcie bicia i coś przeorało powietrze tuż przy jego prawym uchu. Pierwszy raz do mnie strzelają - pomyślał. - Dziewięć lat w policji, dopóki mnie nie **naszli** i nie zniszczyli wszystkiego (czterech krawężników, czterech w cywilu i jeden z wywiadu), i nigdy do mnie nikt nie strzelał, dopiero dziś. Znowu eksplozja O takiej długości, podgrzeję je w ciepłej wodzie... Nimber jęknął tylko raz, gdy Lawier rozplaszczył mu ramiona tak, aby oboczyk wygiął się i kość **naszła** z po-wrotem na miejsce. Potem zamknął oczy i wydawał się me-dytować, podczas gdy Lawier robił, co mógł, aby zmniejszyć opuchlinę i unieruchomić mu ramię, zabezpieczając knie, dzieliły pojedyncze oparcia. Te po moich bokach były zajęte przez ręce kolegów. Obaj trzymali dłonie dziwnie wykrzywione ku górze - gotowe na przyjęcie mojej, gdyby **naszła** mnie taka

Liczba elementów: 17

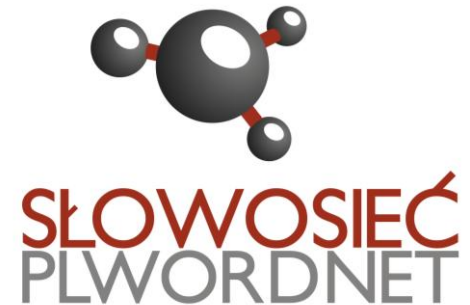
Cache włączony, zapytania: (47%), czas życia: 2 min., zapytań: 14, wielkość: 0 tyś. | Ogólne zużycie pamięci: 71 MB



Ankiety



Wydobywanie wiedzy lingwistycznej z korpusów tekstu ...



- Dziękuję bardzo za uwagę!

nlp.pwr.wroc.pl

plwordnet.pwr.wroc.pl