

Recenzja pracy mgr. inż. Zbigniewa Kalety  
„Regułowy algorytm automatycznego  
rozstrzygania wieloznaczności leksykalnej  
na granicy części mowy  
w tekstach języka polskiego”

Filip Graliński

13 października 2022

## 1 Wstęp

Celem niniejszej recenzji jest stwierdzenie, czy rozprawa doktorska mgr. inż. Zbigniewa Kalety „Regułowy algorytm automatycznego rozstrzygania wieloznaczności leksykalnej na granicy części mowy w tekstach języka polskiego” spełnia wymagania Ustawy o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki (Dz. U. 2003 Nr 65, poz. 595, ze zm.).

## 2 Ocena formalnej strony rozprawy

### 2.1 Ocena układu pracy

Układ recenzowanej pracy oceniam pozytywnie. W rozdziale 2 autor zakreśla tło zagadnienia wieloznaczności leksykalnej oraz opisuje aktualny stan badań. W kolejnych dwóch rozdziałach mgr. inż. Kaleta przedstawia swoje wyniki,

w rozdziale 2 — algorytm automatycznego rozstrzygnięcia wieloznaczności na granicy części mowy, zaś w rozdziale 3 — metodę wykrywania par aspektowych czasownika. Rozdział 2 przechodzi płynnie w kolejny rozdział dzięki podrozdziałowi 3.7, w którym autor opisuje wykorzystanie odpowiedników aspektowych w algorytmie rozstrzygnięcia wieloznaczności. Pracę kończy podsumowanie.

Być może lepiej byłoby oddzielić opis korpusów i zbiorów danych do osobnego rozdziału, z drugiej strony treści te są klarownie wydzielone w stosownych podrozdziałach (2.2 i 3.1.2).

Punkt 2.3.1 jest zbyt długi, powinien zostać podzielony na mniejsze części (w szczególności omówienie prac nad WSD dla języka polskiego, zaczynające się na s. 39, powinno zostać wydzielone do osobnego podpunktu).

Praca zaopatrzona jest w kilka wartościowych dodatkowych, spis treści, bibliografię, spisy rysunków i tabel. Brakuje niestety indeksu, jego częściowym substytutem jest „Indeks skrótów”, trzeba przyznać realnie bardzo pomocny przy lekturze pracy.

## **2.2 Ocena stosowanego aparatu naukowego**

Autor stosuje standardowy w pracach z przetwarzania języka naturalnego i lingwistyki aparat naukowy, działając w paradygmacie regułowym sztucznej inteligencji (czy szerzej: w oparciu o wiedzę) ze znaczącymi elementami wykorzystania efektów uczenia nienadzorowanego (zastosowanie, jak to nazywa autor, osadzeń słów) i nadzorowanego (punkt 4.5.1).

Odnosnie do automatycznego rozstrzygnięcia wieloznaczności na granicy słów (rozdział 3, w szczególności podrozdział 3.4) słabością pracy, w zakresie stosowanego aparatu naukowego, jest nieodwoływanie się do środków udostępnianych przez programowanie w logice i podobne narzędzia — łatwiejsze byłoby zgeneralizowanie kroków algorytmu i procedury rozstrzygnięcia hipotez.

## 2.3 Ocena spójności terminologicznej

Praca jest spójna terminologicznie. Niektóre z proponowanych tłumaczeń angielskich terminów mogą być kontrowersyjne; na przykład uważam, że trafniejszym tłumaczeniem terminu *embedding* jest *zanurzenie* zamiast proponowanego przez autora *osadzenia*.

Trzeba jednak przy tym przyznać, że autor używa wybranych przez siebie terminów konsekwentnie.

Zdarzają się też autorowi pomyłki terminologicznej w niekluczowych dla pracy obszarach, np. *automatyczne tłumaczenie* zamiast *tłumaczenie automatyczne* czy lepiej *tłumaczenie maszynowe*, s. 30, *analizy sentymentu* zamiast *analizy wydźwięku*, przyp. 43, *algorytmy uczone* (raczej *algorytmy oparte na uczeniu maszynowym*), s. 37.

Należy dodać, że autor posługuje się sprawnie nie tylko terminologią informatyczną, lecz również językoznawczą, choć zdarzają się nieścisłości, np. czasownik w wypowiedzi na s. 15 („[...] użytego w niej czasownika.”) jest zbyt wąskim pojęciem (nie są uwzględnione choćby predykatywy). Trochę zaskakujące, że omawiając wieloznaczność wewnątrzparadygmatyczną (s. 16), autor nie przywołuje pojęcia synkretyzmu.

W każdym razie kluczowe pojęcia są w pracy zdefiniowane we właściwy sposób.

## 2.4 Ocena stosowanego języka

Praca jest napisana klarownym i zrozumiałym językiem. Autor nie komplikuje niepotrzebnie języka, ale z drugiej strony unika zbyt wielu uproszczeń.

W obliczu tej ogólnie pozytywnej oceny jakości języka tym bardziej rażą dość liczne błędy i usterki ortograficzne, interpunkcyjne i stylistyczne, czasami seryjne. Wymienię je tutaj:

- niepoprawne lub niestandardowe zapisy skrótów i nazw: pisownia skrótu *wg* z kropką, *word2vec* (zamiast poprawnego *Word2vec*), *formacie djvu* zamiast *formacie DjVu*, s. 50, *miara f1* (w literaturze raczej: *miara F1*),

- niekonsekwentne użycie kursywy w funkcji metajęzykowej (np. *atako-  
wać*, s. 50, *czytać*, s. 104),
- brak interpunkcji w wyliczeniach na s. 50 i 96,
- niefortunne literówki: *oko* zamiast *okno* (s. 11),
- liczne błędy interpunkcyjne (szczególnie rażące w pracy dotyczącej w dużym stopniu składni...):
  - brak przecinków przed zaimkami, np. *jaka* (s. 20), *co* (przyp. 3), *jak* (s. 48 i 52), *czy* (przyp. 111),
  - brak przecinka przed *to* w dwóch następujących po sobie zdaniach (s. 16, „Jeżeli identyczność...”),
  - niekiedy zbędne przecinki, np. przed *tw.* na s. 16, przed *do* w przyp. 42, przed *a* na s. 67,
- niezręczności stylistyczne, np. „Więcej na temat problemu [...] znajduje się [...]”, s. 19, „Możemy się spotkać z taką sytuacją, [...], na podstawie kontekstu szerszego niż zdaniowy”, s. 61, „algorytmów dedykowanych” (zamiast „przeznaczonych”), s. 116.

Wyliczone usterki obniżają ocenę formalnej strony pracy, w tym zakresie pozostaje ona jednak akceptowalna.

## 3 Ocena uzyskanych wyników

### 3.1 Uwagi ogólne

Jak już wspomniałem, badania autora umocowane są raczej w paradygmacie regułowym sztucznej inteligencji z pewnymi znaczącymi (zwłaszcza w rozdziale 4) elementami uczenia maszynowego. Z jednej strony może budzić to krytykę pracy, jeśli chodzi o użytkowy aspekt opisywanych metod w zakresie liczbowych wyników na takim czy innym konkretnym zbiorze testowym. Z

tego punktu widzenia brakuje w pracy odniesienia do czysto statystycznych metod *state of the art* opartych na pretrenowanych sieciach Transformer (choć autor wspomina je w punkcie 2.3.1, ale czyni to mimochodem, nie jest to centralny punkt odniesienia dla pracy).

Z drugiej strony, nie można dyskwalifikować obranego przez autora kierunku badań. Dalej zachowuje on swoje znaczenie i można wskazać kilka powodów, dla których uzyskane wyniki są wartościowe. Po pierwsze jak słusznie zauważa autor (punkt 2.3.3), metody statystyczne są bardzo czułe na użyte zbiory uczące, metody regułowe mogą stanowić tutaj niejako kontrapunkt, punkt odniesienia o na ogół niższych wynikach, ale gwarantujący pewien poziom. Po drugie metody regułowe są interpretowalne w przeciwieństwie do bardziej zaawansowanych rozwiązań statystycznych mających często charakter nieprzeniknionej czarnej skrzynki. W związku z tym, po trzecie, łatwiej przy tych metodach o wnioski o charakterze poznawczym, w wypadku przedmiotowej pracy — wnioski związane z językoznawstwem.

Po czwarte istotnym elementem recenzowanych danych jest eksploracyjna analiza danych w zakresie rozpatrywanych korpusów i zbiorów danych. Praca z obszaru sztucznej inteligencji ograniczająca się tylko do niezinterpretowanych wyników ewaluacji byłaby uboga, autor nie popełnia tego błędu. Ważne poznawczo jest np. stwierdzenie przez autora znaczących różnic w wynikach algorytmów rozstrzygania wieloznaczności (punkt 3.5.3).

W końcu w zakresie metodologii ewaluacji (co jest do pewnego stopnia wspólne zarówno dla metod opartych na uczeniu maszynowym, jak i metod opartych na wiedzy) autor przeprowadza badania zgodnie z zasadami sztuki (zob. np. podrozdział 3.5).

## 3.2 Ocena rozdziału 2

Przedmiotem rozdziału 2 jest głównie omówienie tła i stanu badań w zakresie algorytmów rozstrzygania wieloznaczności leksykalnej. Autorskim wkładem jest dokładna analiza statystyczna zjawiska wieloznaczności leksykalnej zarówno w słownikach, jak i tekstach. Jest to jeden z mocniejszych punk-

tów recenzowanej pracy. Brakowało mi jedynie analizy form z teoretyczną niejednoznacznością, tj. takich, dla których jedna z form jest wysoce nieprawdopodobna (typu *albo*, *musi*, później też w podrozdziale 3.3 brak wyjaśnienia, dlaczego *do* nie jest traktowane jako — teoretycznie możliwa — forma rzeczownika). Form tych nie znajdziemy w, skądinąd interesujących, zestawieniach przygotowanych przez autora i nie jest jasne, czy wynika to z umyślnego zastosowania jakiejś procedury czy z jakiegoś innego powodu.

### 3.3 Ocena rozdziału 3

W rozdziale 3 mgr inż. Zbigniew Kaleta opisuje swoje autorskie rozwiązanie problemu ujednoznaczniania słów, rozwiązanie oparte na schematach syntaktycznych. Rozwiązanie stanowi istotny wkład w dziedzinę, opracowana metoda może stanowić uzupełnienie dla metod czysto statystycznych. Autor zachowuje dobrą równowagę między skutecznością algorytmu, a jego szczegółowością. Rozczarowuje natomiast brak odniesień do bardziej systematycznych technik, np. programowania w logice czy metod typu *beam search*, zwłaszcza w dyskusji rozstrzygania (pod)hipotez (s. 53) czy przetwarzania równoległego (s. 83).

Uwagi szczegółowe:

- w dyskusji założeń (podrozdział 3.1) obok imiesłowów autor powinien również omówić odsłowniki,
- z kolei zanegowane czasowniki (podrozdział 3.4.2) mogłyby być obsługiwane na poziomie tokenizacji,
- niejasne, jak założenie (dość kontrowersyjne zresztą) o szyku zdania z punktu 3.4.11 ma się do przykładu z punktu 3.4.9 (o innym szyku).

### 3.4 Ocena rozdziału 4

Przedmiotem rozdziału 4 jest automatyczne wykrywanie par aspektowych dla języka polskiego. Jest to problem bardziej szczegółowy niż podejmowany

przez autora w rozdziale 3, nowość i istotność uzyskanych wyników oceniam jednak tutaj wyżej, tym bardziej że opracowane przez autora algorytmy mogą mieć zastosowanie dla innych języków słowiańskich, również dla tych o mniejszych dostępnych zasobach. Autor podjął problem w sposób wyczerpujący, uwzględniając zarówno metody oparte na wiedzy, jak i wykorzystujące efekty uczenia maszynowego (osadzenia/zanurzenia). Zresztą analiza zjawiska par aspektowych w języku polskim dokonana przez autora jest bardzo wnikliwa.

Analiza uzyskanych wyników jest wyczerpująca i klarowna (szczególnie interesujący jest tutaj rysunek 4.13).

## 4 Wiedza kandydata

Mgr inż. Kaleta wykazał się w rozprawie ogólną i szczegółową wiedzą teoretyczną w przedmiotowej dyscyplinie naukowej. Szczególną wiedzą i wnikliwością wykazał się autor w zakresie algorytmów rozstrzygania wieloznaczności i tagowania morfosyntaktycznego (podrozdział 2.3); swoją drogą w kolejnych częściach pracy autor mógłby częściej nawiązywać do tej wiedzy (np. jak się ma rozwiązanie zaproponowane w rozdziale 3 do wyboru na podstawie preferencji lub restrykcji omówionego na s. 36).

Badania przeprowadzone przez autora zdecydowanie potwierdzają umiejętność samodzielnego prowadzenia pracy naukowej.

Po przeczytaniu rozprawy z całym przekonaniem stwierdzam, że kandydat ma ogólną wiedzę w dyscyplinie informatyka techniczna i telekomunikacja.

## 5 Podsumowanie recenzji

Rozprawa nie jest pozbawiona słabości, dla mnie największą jest brak szerszego odniesienia do sieci Transformer. Mimo wszystko badania mgr. inż. Zbigniewa Kalety stanowią istotny wkład w dziedzinę komputerowego przetwarzania języka naturalnego, autorowi udało się obronić stawiane przez sie-

bie tezy; ostatecznie pracę oceniam pozytywnie.

Stwierdzam, że recenzowana rozprawa spełnia wymagania stawiane pracom doktorskim. Z przekonaniem rekomenduję dopuszczenie Doktoranta do kolejnych etapów przewodu.

Filip Graliński