

Ontologie i przestrzenie wiedzy

Metody klasyfikacji i reprezentacji wiedzy

Przemysław Szymoniak

29 maja 2026

Spis treści

1	Wprowadzenie	3
2	Ontologia	3
2.1	Częściowy porządek	3
2.2	Ontologia	4
2.3	Drzewo ukorzenione	4
2.4	Typy ontologii	4
2.5	Równoważność ontologii	5
2.6	Operacje na ontologii	6
2.7	Złożenia operacji	6
3	Przestrzeń wiedzy	6
3.1	Wiedza	6
3.2	Zanurzenie	7
3.3	Reprezentacja wektorowa	7
3.4	Warunek zanurzenia	7
3.5	Dziedziczenie	8
3.6	Przestrzeń wiedzy	8
3.7	Minimalna ontologia zanurzenia	8
3.8	Uogólnienie a warunek zanurzenia	9
3.9	Struktury porządkowe w przestrzeni wiedzy	9
4	Algebra przestrzeni wiedzy	10
4.1	Klasyfikacja	10
4.2	Algebra na wspólnej ontologii	10
4.3	Algebra na różnych ontologiach	11
4.4	Złożenie ontologii	11
5	Operatory klasyfikacyjne	12
5.1	Operatory syntaktyczne	13
5.2	Domknięcie hierarchiczne	14
5.3	Reguła klasyfikacyjna	14
5.4	Rozszerzenia	14
5.5	Kategorie przestrzeni wiedzy	15

1 Wprowadzenie

W wielu obszarach działalności pojawia się problem reprezentacji informacji w sposób umożliwiający ich jednoznaczną interpretację oraz systematyczną analizę. Surowe dane, takie jak opisy zdarzeń, dokumenty czy obserwacje, nie posiadają same w sobie struktury pozwalającej na ich porównywanie i przetwarzanie w sposób uporządkowany.

W niniejszej pracy przyjmujemy podejście, w którym informacja jest interpretowana poprzez odniesienie do uporządkowanego zbioru pojęć. Strukturę tę nazywamy ontologią, rozumianą nie w sensie filozoficznym ani systemowym, lecz jako formalny zbiór częściowo uporządkowany (P, S) , gdzie relacja S opisuje zależności hierarchiczne pomiędzy pojęciami.

Na tej podstawie wprowadzamy pojęcie przestrzeni wiedzy, która w niniejszej pracy oznacza zbiór reprezentacji informacji zgodnych z daną ontologią. Każda informacja zostaje przypisana do podzbioru pojęć ontologii, a następnie reprezentowana w postaci struktury odzwierciedlającej zależności pomiędzy tymi pojęciami.

Tak zdefiniowana przestrzeń wiedzy umożliwia analizę informacji z wykorzystaniem narzędzi matematycznych, wynikających ze struktury porządkowej ontologii. W szczególności możliwe staje się porównywanie informacji, ich agregacja oraz badanie zależności pomiędzy różnymi zbiorami wiedzy.

Celem pracy jest formalizacja tego podejścia oraz analiza własności otrzymanych struktur. W kolejnych częściach wprowadzamy definicję ontologii, a następnie pokazujemy, w jaki sposób informacje mogą być osadzone w tej strukturze, tworząc przestrzeń wiedzy. Rozważamy również operacje na ontologiach oraz wynikające z nich przekształcenia przestrzeni wiedzy.

Przedstawione podejście prowadzi do ujęcia, w którym różnorodne systemy reprezentacji informacji mogą być analizowane w ramach wspólnej struktury formalnej, niezależnie od ich dziedziny zastosowania.

2 Ontologia

2.1 Częściowy porządek

Definicja 2.1 (Częściowy porządek). *Niech P będzie zbiorem, a $S \subseteq P \times P$ relacją binarną na P . Relacja S nazywa się częściowym porządkiem, jeżeli dla wszystkich $x, y, z \in P$ zachodzi:*

1. (refleksyjność) $\forall x \in P \ (xSx)$,
2. (antysymetria) $\forall x, y \in P \ ((xSy \wedge ySx) \Rightarrow x = y)$,
3. (tranzytywność) $\forall x, y, z \in P \ ((xSy \wedge ySz) \Rightarrow xSz)$.

Strukturę (P, S) nazywamy zbiorem częściowo uporządkowanym (poset).

2.2 Ontologia

Definicja 2.2 (Ontologia). Niech (P, S) będzie zbiorem częściowo uporządkowanym. Strukturę $\mathcal{O} = (P, S)$ nazywamy ontologią, jeżeli:

- elementy P interpretujemy jako pojęcia,
- relację S interpretujemy jako relację semantyczną.

Relacja S opisuje zależności pomiędzy pojęciami, takie jak relacja ogólności, przynależności lub zależności.

2.3 Drzewo ukorzenione

Definicja 2.3 (Drzewo ukorzenione). Niech (P, S) będzie zbiorem częściowo uporządkowanym. Mówimy, że (P, S) jest drzewem ukorzenionym, jeżeli:

1. (korzeń) istnieje element $r \in P$ taki, że

$$\forall x \in P \ rSx,$$

2. (unikalny poprzednik) dla każdego $x \in P \setminus \{r\}$ istnieje dokładnie jeden element $y \in P$ taki, że

$$ySx \wedge \neg \exists z \in P \ (ySz \wedge zSx).$$

2.4 Typy ontologii

Struktura ontologii zależy od interpretacji relacji semantycznej S . W szczególności można wyróżnić dwa podstawowe typy ontologii.

Definicja 2.4 (Ontologia klasyfikacyjna). *Ontologię $\mathcal{O} = (P, S)$ nazywamy ontologią klasyfikacyjną, jeżeli (P, S) jest drzewem ukorzenionym, a relacja S wyraża zależność typu „jest typem”, „należy do” lub „jest przypadkiem”.*

W takiej ontologii każdy element posiada dokładnie jednego bezpośredniego poprzednika, co zapewnia jednoznaczność klasyfikacji.

Definicja 2.5 (Ontologia zależnościowa). *Ontologię $\mathcal{O} = (P, S)$ nazywamy ontologią zależnościową, jeżeli (P, S) jest zbiorem częściowo uporządkowanym, a relacja S wyraża zależność typu „zależy od”, „wymaga” lub „jest warunkiem”.*

Strukturę taką można otrzymać z drzewa ukorzenionego poprzez usunięcie warunku unikalnego poprzednika, co dopuszcza istnienie wielu bezpośrednich poprzedników dla danego elementu.

Ontologia klasyfikacyjna stanowi szczególny przypadek ontologii zależnościowej, w którym każdy element posiada dokładnie jednego poprzednika.

W dalszej części pracy, o ile nie zaznaczono inaczej, przez ontologię będziemy rozumieć ontologię klasyfikacyjną.

2.5 Równoważność ontologii

Definicja 2.6 (Równoważność strukturalna). *Niech $\mathcal{O}_1 = (P_1, S_1)$ oraz $\mathcal{O}_2 = (P_2, S_2)$ będą ontologiami. Mówimy, że są strukturalnie równoważne, jeżeli istnieje bijekcja*

$$f : P_1 \rightarrow P_2$$

taka, że

$$xS_1y \Leftrightarrow f(x)S_2f(y).$$

Innymi słowy, struktury (P, S) są izomorficzne jako posety.

Definicja 2.7 (Równoważność semantyczna). *Ontologie $\mathcal{O}_1 = (P_1, S_1)$ i $\mathcal{O}_2 = (P_2, S_2)$ nazywamy semantycznie równoważnymi, jeżeli*

$$P_1 = P_2.$$

Definicja 2.8 (Równoważność ontologii). *Ontologie nazywamy równoważnymi, jeżeli są jednocześnie strukturalnie i semantycznie równoważne.*

2.6 Operacje na ontologii

Definicja 2.9 (Uszczegółowienie). *Niech $\mathcal{O} = (P, S)$ będzie ontologią oraz $t \in P$. Definiujemy operację*

$$U(\mathcal{O}, t, p) = (P \cup \{p\}, S \cup \{(t, p)\}),$$

gdzie $p \notin P$.

Definicja 2.10 (Uogólnienie). *Niech $\mathcal{O}' = U(\mathcal{O}, t, p)$. Definiujemy operację*

$$G(\mathcal{O}', p) = (P, S),$$

polegającą na usunięciu p oraz relacji (t, p) .

2.7 Złożenia operacji

Definicja 2.11 (Złożenie uszczegółowień). *Niech \mathcal{O}_0 będzie ontologią początkową. Definiujemy*

$$\mathcal{O}_k = U^k(\mathcal{O}_0).$$

Definicja 2.12 (Złożenie uogólnień). *Dla ontologii \mathcal{O}_k zachodzi*

$$\mathcal{O}_0 = G^k(\mathcal{O}_k).$$

Lemat 2.13 (Odwracalność). *Dla każdej ontologii \mathcal{O} zachodzi*

$$G^k(U^k(\mathcal{O})) = \mathcal{O}.$$

3 Przestrzeń wiedzy

3.1 Wiedza

Definicja 3.1 (Wiedza). *Wiedzą nazywamy zbiór*

$$\mathcal{W} \subseteq \Sigma^*,$$

gdzie Σ^* oznacza zbiór wszystkich możliwych ciągów znaków. Elementy \mathcal{W} interpretujemy jako nieustrukturyzowane informacje, takie jak opisy, obserwacje lub komunikaty.

3.2 Zanurzenie

Definicja 3.2 (Zanurzenie). Niech $\mathcal{W} \subseteq \Sigma^*$ będzie zbiorem wiedzy, a $\mathcal{O} = (P, S)$ ontologią. Zanurzeniem nazywamy odwzorowanie

$$\iota : \mathcal{W} \rightarrow 2^P,$$

które każdej informacji $w \in \mathcal{W}$ przypisuje zbiór pojęć $\iota(w) \subseteq P$.

W ten sposób nieustrukturyzowany zbiór informacji \mathcal{W} zostaje powiązany z uporządkowanym zbiorem pojęć P , co umożliwia jego dalszą formalną analizę.

3.3 Reprezentacja wektorowa

Definicja 3.3 (Reprezentacja wektorowa). Niech $\mathcal{O} = (P, S)$ będzie ontologią oraz $P = \{p_1, \dots, p_n\}$. Każdemu $w \in \mathcal{W}$ przypisujemy wektor

$$v(w) \in \{0, 1\}^n, \quad v_i(w) = \begin{cases} 1 & \text{gdy } p_i \in \iota(w), \\ 0 & \text{w przeciwnym razie.} \end{cases}$$

Zbiór wszystkich takich wektorów oznaczamy przez

$$\mathcal{V} = \{v(w) : w \in \mathcal{W}\} \subseteq \{0, 1\}^n.$$

Każdy wektor $v(w)$ reprezentuje podzbiór pojęć ontologii, a zbiór \mathcal{V} opisuje strukturę zgromadzonej wiedzy.

3.4 Warunek zanurzenia

Definicja 3.4 (Warunek zanurzenia). Mówimy, że przestrzeń \mathcal{V} jest zanurzona w ontologii \mathcal{O} , jeżeli dla każdego $v(w) \in \mathcal{V}$ zachodzi

$$\text{supp}(v(w)) \subseteq P,$$

gdzie

$$\text{supp}(v(w)) = \{p \in P : v_p(w) = 1\}.$$

Innymi słowy, wszystkie elementy wiedzy są reprezentowane wyłącznie przy użyciu pojęć należących do ontologii.

3.5 Dziedziczenie

Definicja 3.5 (Dziedziczenie). *Mówimy, że \mathcal{V} spełnia warunek dziedziczenia, jeżeli dla każdego $v(w) \in \mathcal{V}$ oraz $p \in P$ zachodzi:*

$$v_p(w) = 1 \Rightarrow v_q(w) = 1 \quad \text{dla wszystkich } q \in \text{Anc}(p),$$

gdzie

$$\text{Anc}(p) = \{q \in P : qSp\}$$

oznacza zbiór wszystkich przodków elementu p .

Warunek dziedziczenia oznacza, że przypisanie informacji do pojęcia szczegółowego implikuje przypisanie jej do wszystkich jego uogólnień.

W przypadku ontologii zależnościowych zbiór $\text{Anc}(p)$ może zawierać wiele elementów nieporównywalnych względem relacji S , co oznacza, że dziedziczenie ma charakter wieloźródłowy.

3.6 Przestrzeń wiedzy

Definicja 3.6 (Przestrzeń wiedzy). *Niech $\mathcal{O} = (P, S)$ będzie ontologią. Przestrzenią wiedzy na \mathcal{O} nazywamy parę*

$$\mathfrak{W} = (\mathcal{O}, \mathcal{V}),$$

gdzie $\mathcal{V} \subseteq \{0, 1\}^{|P|}$ powstaje poprzez zanurzenie \mathcal{W} w \mathcal{O} .

Jeżeli dodatkowo \mathcal{V} spełnia warunek dziedziczenia, to \mathfrak{W} nazywamy *hierarchiczną przestrzenią wiedzy*.

3.7 Minimalna ontologia zanurzenia

Definicja 3.7 (Minimalna ontologia zanurzenia). *Niech $\mathfrak{W} = (\mathcal{O}, \mathcal{V})$, gdzie $\mathcal{O} = (P, S)$. Minimalną ontologią zanurzenia nazywamy najmniejszy zbiór $P_m \subseteq P$ spełniający:*

1. $\text{supp}(v(w)) \subseteq P_m$ dla każdego $v(w) \in \mathcal{V}$,
2. jeżeli $p \in P_m$ oraz qSp , to $q \in P_m$ (domknięcie względem przodków).

Przez \mathcal{O}_m oznaczamy ontologię ograniczoną do zbioru P_m .

Innymi słowy, \mathcal{O}_m jest najmniejszą częścią ontologii, która zawiera wszystkie użyte pojęcia oraz ich uogólnienia.

3.8 Uogólnienie a warunek zanurzenia

Odwrótnością uszczegółowienia jest *uogólnienie*. W szczególności może ono polegać na utożsamieniu zbioru dzieci $C(t) = \{c_1, \dots, c_k\}$ z ich rodzicem t .

Niech $\mathcal{O}' = (P', S')$ będzie ontologią powstałą po takim uogólnieniu.

Lemat 3.8 (Uogólnienie łamie warunek zanurzenia). *Jeżeli istnieje $v(w) \in \mathcal{V}$ takie, że $c_i \in \text{supp}(v(w))$ dla pewnego i , to $v(w)$ nie spełnia warunku zanurzenia względem \mathcal{O}' .*

Dowód. Po uogólnieniu elementy c_1, \dots, c_k nie należą do P' . Zatem $\text{supp}(v(w)) \not\subseteq P'$, co przeczy warunkowi zanurzenia. \square

Aby zachować spójność, definiujemy transformację

$$\pi : \{0, 1\}^{|P|} \rightarrow \{0, 1\}^{|P'|}$$

daną wzorem

$$\pi(v)_t = \max\{v_{c_1}, \dots, v_{c_k}, v_t\}, \quad \pi(v)_p = v_p \quad \text{dla } p \notin C(t) \cup \{t\}.$$

Nowa przestrzeń $\mathcal{V}' = \pi(\mathcal{V})$ spełnia warunek zanurzenia względem \mathcal{O}' , kosztem utraty informacji o rozróżnialności elementów c_i .

3.9 Struktury porządkowe w przestrzeni wiedzy

Definicja 3.9 (Ideał). *Niech (P, S) będzie ontologią. Podzbiór $I \subseteq P$ nazywamy ideałem, jeżeli:*

1. (domknięcie w dół) jeżeli $p \in I$ oraz qSp , to $q \in I$,
2. (skierowanie) dla dowolnych $p, q \in I$ istnieje $r \in I$ takie, że pSr oraz qSr .

Każdy wektor $v(w)$ spełniający warunek dziedziczenia wyznacza ideał w ontologii.

Definicja 3.10 (Filtr). *Podzbiór $F \subseteq P$ nazywamy filtrem, jeżeli:*

1. (domknięcie w górę) jeżeli $p \in F$ oraz pSq , to $q \in F$,
2. (skierowanie w dół) dla dowolnych $p, q \in F$ istnieje $r \in F$ takie, że rSp oraz rSq .

Definicja 3.11 (Krata). *Zbiór częściowo uporządkowany (P, S) nazywamy kratą, jeżeli dla dowolnych $p, q \in P$ istnieją:*

- kres dolny (infimum) p, q ,
- kres górny (supremum) p, q .

W szczególności zbiór wszystkich ideałów ontologii, uporządkowany przez inkluzję, tworzy kratę.

4 Algebra przestrzeni wiedzy

4.1 Klasyfikacja

Możemy wyróżnić dwie zasadnicze klasy algebr przestrzeni wiedzy:

1. **Algebra na wspólnej ontologii** — wszystkie przestrzenie wiedzy odnoszą się do tej samej ontologii \mathcal{O} . Operacje wykonywane są na zbiorach wektorów \mathcal{V} , przy stałej strukturze ontologicznej.
2. **Algebra na różnych ontologiach** — przestrzenie wiedzy zbudowane są na ontologiach $\mathcal{O}_1, \mathcal{O}_2, \dots$, które mogą być równoważne strukturalnie, semantycznie lub nie spełniać żadnego z warunków równoważności. W takim przypadku operacje wymagają wcześniejszego ustalenia odwzorowań pomiędzy ontologiami.

4.2 Algebra na wspólnej ontologii

Definicja 4.1 (Algebra przestrzeni wiedzy na wspólnej ontologii). *Niech $\mathcal{O} = (P, S)$ będzie ustaloną ontologią. Przez algebrę przestrzeni wiedzy nad \mathcal{O} rozumiemy strukturę*

$$\mathfrak{A}(\mathcal{O}) = (\mathcal{K}, \cup, \cap, \setminus, \complement),$$

gdzie:

- $\mathcal{K} = \{(\mathcal{O}, \mathcal{V}) : \mathcal{V} \subseteq \{0, 1\}^{|P|}\}$,
- operacje $\cup, \cap, \setminus, \complement$ działają na składnikach \mathcal{V} przy ustalonej ontologii \mathcal{O} :

$$(\mathcal{O}, \mathcal{V}_1) \cup (\mathcal{O}, \mathcal{V}_2) = (\mathcal{O}, \mathcal{V}_1 \cup \mathcal{V}_2),$$

$$(\mathcal{O}, \mathcal{V}_1) \cap (\mathcal{O}, \mathcal{V}_2) = (\mathcal{O}, \mathcal{V}_1 \cap \mathcal{V}_2),$$

$$(\mathcal{O}, \mathcal{V}_1) \setminus (\mathcal{O}, \mathcal{V}_2) = (\mathcal{O}, \mathcal{V}_1 \setminus \mathcal{V}_2),$$

$$\complement(\mathcal{O}, \mathcal{V}) = (\mathcal{O}, \{0, 1\}^{|P|} \setminus \mathcal{V}).$$

Lemat 4.2. *Struktura $\mathfrak{A}(\mathcal{O})$ jest kratą względem operacji \cup oraz \cap .*

Dowód. Operacje \cup i \cap są przemienne, łączne i rozdzielne oraz posiadają elementy neutralne, co wynika bezpośrednio z własności operacji zbiorowych. \square

Należy zauważyć, że operacja dopełnienia \complement nie zachowuje ogólnie warunku dziedziczenia, a zatem może prowadzić poza klasę hierarchicznych przestrzeni wiedzy.

4.3 Algebra na różnych ontologiach

W przypadku, gdy przestrzenie wiedzy zbudowane są na różnych ontologiach, można wyróżnić następujące sytuacje:

1. **Ontologie równoważne strukturalnie i semantycznie** Istnieje bijekcja $f : P_1 \rightarrow P_2$ zachowująca relację S , co pozwala przenosić przestrzenie wiedzy pomiędzy ontologiami. Operacje algebraiczne są w tym sensie izomorficzne.
2. **Ontologie równoważne semantycznie, lecz nie strukturalnie** Zbiory pojęć są identyczne ($P_1 = P_2$), lecz relacje S_1 i S_2 różnią się. Można wówczas operować na wspólnej przestrzeni $\{0, 1\}^{|P|}$, przy czym własności dziedziczenia mogą się różnić.
3. **Ontologie nierównoważne semantycznie** Zbiory pojęć są różne. W takim przypadku operacje wymagają sprowadzenia przestrzeni do wspólnej dziedziny:
 - przez ograniczenie do części wspólnej $P_1 \cap P_2$,
 - przez rozszerzenie o brakujące pojęcia (z wartościami zerowymi),
 - przez transformacje zanurzenia zgodne z relacją S .

4.4 Złożenie ontologii

Definicja 4.3 (Złożenie ontologii). Niech $\mathcal{O}_1 = (P_1, S_1), \dots, \mathcal{O}_n = (P_n, S_n)$ będą ontologiami. Niech r_1, \dots, r_n będą elementami minimalnymi w odpowiednich ontologiach.

Złożeniem ontologii nazywamy strukturę

$$\mathcal{O}^* = (P^*, S^*),$$

gdzie:

$$P^* = \bigcup_{i=1}^n P_i \cup \{r\}, \quad S^* = \bigcup_{i=1}^n S_i \cup \{(r, r_i) : i = 1, \dots, n\},$$

a r jest nowym elementem minimalnym.

Lemat 4.4. *Struktura \mathcal{O}^* jest ontologią, tzn. zbiorem częściowo uporządkowanym.*

Dowód. Relacja S^* zachowuje refleksywność, antysymetrię i tranzytywność, ponieważ każda z relacji S_i jest częściowym porządkiem, a dodanie elementu r jako nowego elementu minimalnego nie wprowadza sprzeczności z tymi własnościami. \square

Złożenie ontologii pozwala sprowadzić wiele niezależnych ontologii do jednej wspólnej struktury, w której można definiować operacje algebraiczne na przestrzeniach wiedzy.

5 Operatory klasyfikacyjne

Najprostszą metodą zanurzania przestrzeni wiedzy w ontologii jest *ręczne przypisywanie* pojęć. Dla każdej informacji $w \in \mathcal{W}$ użytkownik wybiera zbiór węzłów $\iota(w) \subseteq P$, tworząc w ten sposób wektor $v(w) \in \{0, 1\}^{|P|}$. Metoda ta jest formalnie poprawna, lecz w praktyce czasochłonna i podatna na niespójności wynikające z interpretacji użytkowników.

W celu automatyzacji procesu klasyfikacji wprowadzamy pojęcie *operatora klasyfikującego*, który analizuje treść opisu $x \in \Sigma^*$ i odwzorowuje go na pojęcie ontologii.

Podstawą konstrukcji takiego operatora jest teoria języka i pojęć, a w szczególności ujęcie semiotyczne, które wyróżnia trzy poziomy analizy:

- **syntaktyczny** – dotyczący form znakowych, aliasów i tokenów,
- **semantyczny** – dotyczący przypisywania znaczeń,
- **pragmatyczny** – dotyczący użycia symboli w kontekście.

W pierwszym przybliżeniu skupiamy się na poziomie syntaktycznym.

Definicja 5.1 (Alias). *Każdemu pojęciu $p \in P$ przypisany jest skończony zbiór*

$$A(p) \subseteq \Sigma^*.$$

Element $a \in A(p)$ nazywamy aliasem pojęcia p . Alias stanowi formę powierzchniową, umożliwiającą powiązanie tekstu z pojęciem ontologii.

5.1 Operatory syntaktyczne

Definicja 5.2 (Operator syntaktyczny). *Operatorem syntaktycznym nazywamy funkcję*

$$O_{\text{syn}} : \Sigma^* \rightarrow [0, 1]^{|P|},$$

która każdemu opisowi $x \in \Sigma^*$ przypisuje wektor dopasowania

$$v_{O_{\text{syn}}}(x) = (O_{\text{syn}}(x)_1, \dots, O_{\text{syn}}(x)_{|P|}),$$

gdzie współrzędna $O_{\text{syn}}(x)_i$ określa stopień dopasowania tekstu x do pojęcia p_i na podstawie aliasów $A(p_i)$.

Operator nie musi spełniać warunku dziedziczenia.

Najprostszym przypadkiem jest operator binarny:

$$O_{\text{syn}}(x)_i = \begin{cases} 1 & \text{jeśli istnieje } a \in A(p_i) \text{ takie, że } a \sqsubseteq x, \\ 0 & \text{w przeciwnym razie.} \end{cases}$$

Definicja 5.3 (Operator dyskretny). *Operatorem dyskretnym nazywamy operator syntaktyczny, dla którego*

$$v_{O_{\text{disc},i}}(x) = \max\{w(a, x) : a \in A(p_i) \cup \{p_i\}, a \sqsubseteq x\}.$$

Funkcja wagowa $w(a, x)$ może uwzględniać:

- długość aliasu,
- liczbę wystąpień,
- ograniczenia górne.

Definicja 5.4 (Operator leksykalny). *Operatorem leksykalnym nazywamy operator syntaktyczny O_{lex} , który wykorzystuje funkcję podobieństwa*

$$\text{sim}_{\text{lex}}(s, a) \in [0, 1],$$

gdzie s jest fragmentem tekstu x , $a \in A(p)$.

Niech $\theta_{\text{lex}} \in (0, 1)$.

Wektor wynikowy ma postać:

$$v_{O_{\text{lex},i}}(x) = \max\{w_{\text{lex}}(a, x) : a \in A(p_i), \text{sim}_{\text{lex}}(s, a) \geq \theta_{\text{lex}}\}.$$

5.2 Domknięcie hierarchiczne

Definicja 5.5 (Domknięcie hierarchiczne). *Niech $v \in [0, 1]^{|P|}$. Definiujemy operator*

$$\text{cl}(v)_q = \max\{v_p : p \in P, qSp\}.$$

Operator cl zapewnia spełnienie warunku dziedziczenia, tj. propaguje wartości na wszystkich przodków w ontologii.

5.3 Reguła klasyfikacyjna

Definicja 5.6 (Reguła klasyfikacyjna). *Niech \mathcal{F} będzie zbiorem operatorów, a $\lambda_O \geq 0$, $\sum_{O \in \mathcal{F}} \lambda_O = 1$.*

Definiujemy wektor

$$v(x) = \sum_{O \in \mathcal{F}} \lambda_O v_O(x).$$

Ostateczny wynik klasyfikacji:

$$\hat{v}(x) = \text{cl}(v(x)).$$

Reguła klasyfikacyjna łączy wyniki wielu operatorów i zapewnia zgodność z ontologią poprzez domknięcie hierarchiczne.

5.4 Rozszerzenia

Model można rozszerzyć o:

- operatory semantyczne (embedding),
- operatory pragmatyczne (kontekst).

5.5 Kategorie przestrzeni wiedzy

Na potrzeby dalszych badań wyróżniamy różne warianty przestrzeni wiedzy, które mogą być ze sobą porównywane:

- **Przestrzeń referencyjna** \mathcal{V}_{ref} – powstaje na podstawie wpisów tworzonych przez użytkowników, wspieranych przez operatorów klasyfikacyjnych i regułę klasyfikacyjną. Jest to główny punkt odniesienia, gdyż łączy wyniki algorytmu z decyzjami człowieka, zapewniając praktyczną wiarygodność systemu.
- **Przestrzeń LLM** \mathcal{V}_{llm} – generowana automatycznie przez model językowy, który przypisuje węzły ontologii na podstawie treści wpisów. Może pełnić rolę filtra redukującego *szum* wynikający z nieprecyzyjnych lub pobieżnych opisów operatorów (np. spowodowanych pośpiechem lub rutyną). Dzięki temu \mathcal{V}_{llm} może stanowić bardziej spójny punkt odniesienia semantycznego.
- **Przestrzenie parametryczne** \mathcal{V}_{θ} – powstają w wyniku modyfikacji konfiguracji operatorów klasyfikacyjnych (np. progów dopasowania, wag operatorów, liczby prezentowanych dopasowań k). Pozwalają one *dostrajać* działanie systemu, tak aby lepiej odzwierciedlał przestrzeń referencyjną albo alternatywnie przestrzeń \mathcal{V}_{llm} .

Porównywanie różnych przestrzeni wiedzy umożliwia identyfikację wariantów najbardziej stabilnych i zgodnych z oceną ekspercką, a także wskazanie potencjalnych ulepszeń w konstrukcji operatorów oraz reguły klasyfikacyjnej.