

ДЕКОМПОЗИЦІЯ ОНТОЛОГІЙ ТА ГРАФІВ ЗНАНЬ ЧЕРЕЗ ВІДНОШЕННЯ НУЛЬОВОГО ПОРЯДКУ

Галушка О.В.¹ [ORCID], Шинкаренко В.І.² [ORCID]

¹Український державний університет науки і технологій, аспірант, Україна

²Український державний університет науки і технологій, д.т.н, професор, Україна

Анотація. Графи знань та онтології оперують структурно гетерогенними висловлюваннями – бінарними триплетами, n -арними реіфікованими відношеннями, вкладеними фактами – що ускладнює їх обробку нейромережевими моделями. У роботі запропоновано декомпозицію довільного онтологічного висловлювання до єдиного примітиву – анонімного спрямованого зв'язку нульового порядку (r_0). Предикат переноситься з ребра у вузол, і граф стає однорідним: усі елементи є вузлами, з'єднаними єдиним типом ребра. Декомпозиція спирається на два принципи: ролеву нейтральність концептів, за якою розрізнення «концепт» і «відношення» є контекстуальним, та рекурсивний порядок висловлювання, що формалізує структурну складність. Показано, що r_0 -декомпозиція виконує для графів знань роль, аналогічну токенизації для текстів: перетворює гетерогенну структуру на однорідну, придатну для нейромережевої обробки.

Ключові слова. Граф знань, онтологія, реіфікація, декомпозиція логічних висловлювань, представлення знань, нейронні мережі.

Вступ. Графи знань оперують структурно різномірними формами представлення фактів. Найпростішою є бінарний триплет виду «Суб'єкт-Предикат-Об'єкт» [1], проте реальні предметні області часто потребують складніших конструкцій: n -арних відношень із кваліфікаторами, вкладених фактів. Для їх формалізації розроблено низку підходів – реіфікацію RDF [2], формат RDF-star [3], гіперреляційні моделі [4, 5] – кожен з яких орієнтований на певний тип структури. Модель, створена для одного класу структур, як правило, не переноситься на інший без суттєвих змін.

Це контрастує з обробкою природної мови, де довільне висловлювання зводиться до однорідної послідовності слів, і саме однорідність представлення робить текст придатним для обробки нейромережевими моделями [6]. Для графів знань аналогічного уніфікованого представлення досі не запропоновано.

У цій роботі ми вводимо таке представлення шляхом декомпозиції довільного онтологічного висловлювання до єдиного анонімного спрямованого примітиву нульового порядку (r_0) та формалізуємо структурну складність висловлювань через рекурсивну метрику порядку.

Основний матеріал. Ми визначаємо анонімний спрямований зв'язок нульового порядку r_0 як примітив, що не несе власної семантики й лише фіксує спрямований зв'язок між двома вузлами. Бінарний триплет розкладається у два зв'язки типу r_0 :

$$R(S, O) \rightarrow r_0(S, R), r_0(R, O) \quad (1)$$

Семантика при цьому переноситься з ребра у вузол: предикат relation перестає бути міткою ребра й стає повноцінним вузлом графа нарівні з іншими концептами. В результаті отримуємо однорідний граф, у якому всі елементи онтології – і сутності, і відношення – є вузлами, з'єднаними єдиним типом ребра r_0 .

N-арне відношення обробляється аналогічно через реіфікацію. Тернарне відношення $R(a, b, c)$ спочатку реіфікується у вузол R^* , пов'язаний із кожним аргументом відповідною роллю, а потім кожен із цих зв'язків декомпозиується до пари r_0 . Таким чином, незалежно від початкової арності, результатом завжди є однорідний r_0 -граф.

Декомпозиція передбачає, що предикат стає вузлом. Це можливо завдяки принципу ролевої нейтральності: розрізнення між «концептом» та «відношенням» є контекстуальним, а не онтологічним. Концепт, що виступає предикатом в одному висловлюванні, може бути аргументом в іншому. Наприклад, предикат «працює» у триплеті `works_at(Petro, University)` стає аргументом у висловлюванні `believes(Mary, works_at(Petro, University))`. Ролева нейтральність забезпечує необмежену вкладеність: реіфікований концепт автоматично може брати участь у висловлюваннях вищого рівня без додаткових механізмів.

Для формалізації структурної складності ми вводимо рекурсивну цілочислову метрику – порядок висловлювання. Атомарний концепт, що не є

висловлюванням, має порядок 0. Порядок висловлювання φ з аргументами a_1, a_2, \dots, a_n визначається як:

$$\rho(\varphi) = \delta + \max(\rho(a_1), \rho(a_2), \dots, \rho(a_n)) \quad (2)$$

де $\delta = 1$ для бінарних відношень (арність 2) та $\delta = 1$ для n-арних (арність > 2), оскільки n-арне відношення потребує додаткового кроку реіфікації при декомпозиції.

Розглянемо приклади різних порядків:

- `works_at(Petro, University)`: бінарне, атомарні аргументи, $\rho = 1$;
- `between(Kyiv, Dnipro, Lviv)`: тернарне, потребує реіфікації, $\rho = 2$;
- `believes(Mary, works_at(Petro, University))`: бінарне, але другий аргумент є висловлюванням порядку 1, тому $\rho = 2$.

Порядок є властивістю саме висловлювання, а не предиката: один і той самий предикат може брати участь у висловлюваннях різних порядків залежно від складності аргументів. Це принципово відрізняє запропоновану метрику від класичного поділу відношень на «бінарні» та «n-арні», який характеризує предикат, а не конкретне висловлювання.

Подібно до того, як токенізація зводить природномовне висловлювання до послідовності однотипних одиниць [6], r_0 -декомпозиція зводить онтологічне висловлювання довільного порядку до однорідного графа з єдиним типом ребра. Висловлювання будь-якого порядку представлені однаковими структурними елементами, тому обробляються однаково незалежно від вихідної складності.

Висновки. У роботі запропоновано декомпозицію довільного онтологічного висловлювання до єдиного анонімного спрямованого примітиву r_0 та формалізовано структурну складність висловлювань через рекурсивну метрику порядку. Проведений нами огляд літератури показав, що існуючі підходи до вкладених фактів або обмежуються фіксованою глибиною [4, 5], або допускають довільну глибину без формальної метрики. Деякі роботи порівнюють різні форми представлення фактів у графах знань [2, 7], проте жодна з них не використовує декомпозицію до єдиного анонімного примітиву. За результатами проведеного огляду, запропонований підхід є першим, що зводить висловлювання довільної структурної складності до однорідного графа через єдиний тип зв'язку, забезпечуючи уніфіковане представлення, придатне для нейромережевої обробки.

ЖИТЕПАТҮПА / REFERENCE

1. Hogan, A., Blomqvist, E., Cochez, M., d'Amato, C., Melo, G. D., Gutierrez, C., ... & Zimmermann, A. (2021). Knowledge graphs. *ACM Computing Surveys*, 54(4), 1–37. DOI: 10.1145/3447772
2. Hernández, D., Hogan, A., & Krötzsch, M. (2015). Reifying RDF: What works well with Wikidata? *Proceedings of the 11th International Workshop on Scalable Semantic Web Knowledge Base Systems (SSWS)*, 32–47. URL: https://ceur-ws.org/Vol-1457/SSWS2015_paper3.pdf
3. Hartig, O. (2017). Foundations of RDF* and SPARQL* – An alternative approach to statement-level metadata in RDF. *Proceedings of the 11th Alberto Mendelzon International Workshop on Foundations of Data Management (AMW)*, vol. 1912. URL: <https://ceur-ws.org/Vol-1912/paper12.pdf>
4. Galkin, M., Trivedi, P., Maheshwari, G., Usbeck, R., & Lehmann, J. (2020). Message passing for hyper-relational knowledge graphs. *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 7346–7359. DOI: 10.18653/v1/2020.emnlp-main.596
5. Guan, S., Jin, X., Wang, Y., & Cheng, X. (2019). Link prediction on n-ary relational data. *Proceedings of the World Wide Web Conference (WWW)*, 583–593. DOI: 10.1145/3308558.3313414
6. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems (NeurIPS)*, 30. URL: <https://papers.nips.cc/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html>
7. Rosso, P., Yang, D., & Cudré-Mauroux, P. (2020). Beyond triplets: Hyper-relational knowledge graph embedding for link prediction. *Proceedings of The Web Conference (WWW)*, 1885–1896. DOI: 10.1145/3366423.3380257

DECOMPOSITION OF ONTOLOGIES AND KNOWLEDGE GRAPHS VIA ZERO-ORDER RELATIONS

Oleksandr Halushka, Viktor Shynkarenko

Abstract. *Knowledge graphs and ontologies operate with structurally heterogeneous statements – binary triples, n-ary reified relations, and nested facts – which complicates their unified processing. This paper proposes decomposing any ontological statement into a single primitive – an anonymous directed link of zero order (r_0). The predicate is transferred from the edge to a node, producing a homogeneous graph where all ontological elements are nodes connected by a single edge type. The decomposition relies on two principles: role neutrality of concepts, whereby the distinction between «entity» and «relation» is contextual rather than ontological, and recursive statement order, which formalizes structural complexity through a metric. It is shown that r_0 -decomposition performs for knowledge graphs a role analogous to tokenization for natural language texts: it transforms a heterogeneous structure into a homogeneous one suitable for neural network processing.*

Keywords. *Knowledge graph, ontology, reification, logical statement decomposition, knowledge representation, neural networks.*