



Adam Trybus

Logika i analiza rozumowań w dobie sztucznej inteligencji

Adam Trybus

dr, Uniwersytet Jagielloński, Instytut Filozofii,
ul. Grodzka 52, 31-044 Kraków
e-mail: adam.trybus@uj.edu.pl
ORCID: 0000-0003-4170-8665

Jest adiunktem w Instytucie Filozofii Uniwersytetu Jagiellońskiego w Krakowie. W Pracowni Logiki zajmuje się historią logiki i matematyki (w szczególności geometrii) oraz zagadnieniami związanymi z rozumowaniem przestrzennym, systemami odrzuceniowymi, parakonsystencją i implementacją formalizmów logicznych. Ostatnio koncentruje się na relacjach między Large Language Models a logiką formalną. Jest członkiem Bielik Reasoning Team, grupy odpowiedzialnej za ocenę zdolności rozumowania tego polskiego systemu LLM. Uzyskał stopień doktora w School of Computer Science, University of Manchester, UK w zakresie logik afinicznych opartych na regionach.

Fotografia na odwrocie: Maria Kokoszyńska-Lutmanowa i Seweryna Łuszczewska-Romahnowa. Na drugim planie: Kazimierz Ajdukiewicz. Jabłonna 1961.



Czas robi głupców ze wszystkich.
Jedyna nadzieja w tym, że zastąpią nas lepsi.
Eric Temple Bell

Przełom XIX i XX wieku to okres wielu kryzysów. Również dla nauki to czas licznych wstrząsów. W ostatnim ćwierćwieczu XIX wieku niemiecki matematyk Georg Cantor tworzy podwaliny nowej dziedziny: *teorii mnogości*. W dobie powstawania nowych gałęzi nauki matematycy również zadają sobie pytanie o podstawy matematyki i jej relacje z innymi dziedzinami. Zaczynają pojawiać się prace, które doprowadzą do tego, że to właśnie teoria mnogości na długo stanie się podstawą dla całej matematyki. W 1903 roku brytyjski matematyk i filozof Bertrand Russell publikuje pracę *The Principles of Mathematics*, w której opisuje m.in. problem z podstawami matematyki: w ramach proponowanego systemu, stanowiącego fundament dla całej matematyki, można zrekonstruować wersję starożytnego *paradoksu kłamcy* (Russell 1903). Książka odbija się szerokim echem także na ziemiach, na których żyją Polacy. Czyta ją dwudziestokilkuletni wówczas Jan Łukasiewicz, świeżo upieczony doktor filozofii specjalizujący się w logice. Praca Russella wstrząsa jego naukowym światem; wkrótce, pod jej wpływem, napisze dzieło, które stanowić będzie z kolei źródło inspiracji (i edukacji!) dla rodzimych badaczy,

a w konsekwencji położy podwaliny pod najbardziej wpływową polskojęzyczną szkołę filozoficzną w dziejach, znaną dziś powszechnie jako Szkoła Lwowsko-Warszawska.

Przyjrzyjmy się zatem dziedzictwu Szkoły, biorąc pod uwagę jeden z najważniejszych poruszanych tam wątków badawczych: logiki i rozumowania. W związku z tym, że zakres czasowy tej analizy to ponad 100 lat, z konieczności operować będę skrótem i wyłuskiwać główne trendy. Nie znaczy to oczywiście, że innego typu badań w ogóle nie prowadzono. Wręcz przeciwnie – zachęcam czytelnika do zgłębienia wielowątkowego dziedzictwa Szkoły, a do niektórych z badań spoza głównego nurtu nawiążę w ostatniej części pracy.

Należy przyznać, że ostatnie 100 lat wywróciło naukę o rozumowaniu do góry nogami i nie miała w tym zasługa polskich logików odwołujących się do tradycji Szkoły. Najwięcej entuzjazmu, energii i wysiłku pochłonęło badanie podwalin logiki formalnej (która stała się głównym narzędziem opisu matematyki). Dzieło Łukasiewicza, które wydatnie wpłynęło na kształt polskiej logiki, nosi tytuł *O zasadzie sprzeczności u Arystotelesa* i zostało wydane w Krakowie w 1910 roku. Łukasiewicz krytykował tytułowe prawo logiki, argumentując, że nie jest ono tak centralne i fundamentalne, jak do tej pory sądzono. Tradycyjne ujęcie systemu aksjomatycznego było takie, że budujemy go na wzór Euklidesa, na podstawie pewnych intuicyjnych prawd, które nie wymagają dowodu, a na których opiera się cała reszta teoretycznej konstrukcji. W logice klasycznej jedną z takich prawd stanowiło właśnie prawo sprzeczności: żadna rzecz nie może być i nie być jednocześnie¹. Łukasiewicz chciał natomiast pokazać, że prawo to może mieć dowód, *a nawet go wymaga*. Stąd już tylko krok do uznania (jak czyni się

¹ Znane są i inne sformułowania tego prawa; dyskusje nad odpowiednią jego formą i rolą trwają zresztą do dziś (por. Trybus 2024).

w logice współczesnej), że nieistotny jest intuicyjny status aksjomatów w systemie, który budujemy: ważne jest, by tak dobrać formuły podstawowe, aby móc udowodnić *twierdzenie o pełności* dla naszego systemu (według tego twierdzenia formuły, które udowodnimy w danym systemie aksjomatycznym, to wszystkie i tylko *tautologie* – zdania zawsze prawdziwe – tego systemu). Łukasiewicz reprezentował podejście do badań logicznych, w ramach którego punktem skupienia są własności systemów logicznych (często: różne sformułowania logiki klasycznej lub logiki *wielowartościowe*, tzn. takie, w których poza prawdą i fałszem przyjmuje się też inne „wartości” logiczne), analizowane najczęściej w ich podstawowej postaci, tzn. w ramach *logiki zdań*.

Z grubsza sprawę ujmując, mamy trzy główne aspekty badań logicznych, które stanowią spuściznę ojców-założycieli Szkoły: 1) badanie systemów formalno-logicznych i ich własności, szczególnie w logice zdań, 2) pracę nad podstawami matematyki motywowaną uniknięciem paradoksów (tu wyróżnić należy prace Stanisława Leśniewskiego nad *mereologią*), 3) zastosowanie logiki do matematyki, z wykorzystaniem przede wszystkim aparatu *logiki predykatów* (tu zarysowuje się wpływ ucznia Leśniewskiego – Alfreda Tarskiego, twórcy *teorii modeli*). Wydaje się, że badania związane z podejściem Łukasiewicza (aspekt 1) odniosły największy sukces, jeśli chodzi o rozwój badań wśród potomków Szkoły; badania nad podstawami matematyki (aspekt 2) z konieczności uległy wygaszeniu, gdy w ewolucyjnym biegu na polu bitwy pozostały teoria mnogości i teoria typów; natomiast badania związane z teorią modeli i zastosowaniem logiki do teorii matematycznych (aspekt 3) wprawdzie również do pewnego stopnia rozwinęły się na gruncie polskim, ale niekoniecznie w kontekście schedy po Szkole.

W Polsce następców Szkoły odnajdujemy przede wszystkim w ramach filozofii, co tłumaczyć może względny sukces

badan post-łukasiewiczowskich, natomiast badania teoriomodelowe rozwijano głównie w ramach matematyki. Do dyscyplin, w ramach których rozwijało się badania formalnologiczne, w połowie kolejnego stulecia dołączyła informatyka. Doprowadziło to do ciekawego transferu wiedzy między tymi dyscyplinami.

Zadając filozoficznie motywowane pytania: jaki jest prawdziwościowy status zdań o przyszłości, czy sprzeczność może być racjonalna, co to znaczy coś wiedzieć – kontynuatorzy Szkoły ulepszali systemy logiczne i tworzyli nowe, które miały pomóc w analizie tych kwestii. Wielki sukces Szkoły miał jednak swoją cenę. Światło rzucone na niektóre aspekty nauki o rozumowaniu na długi czas pograżyło w ciemności inne obszary. Logika *nieformalna* jest tego najlepszym przykładem. Historycznie rzecz ujmując, badania nad logiką nieformalną nie rozwinęły się w Polsce znacznie i przez długi czas pozostawały pod wpływem idei pochodzących ze Szkoły – głównie podziału na rozumowania *dedukcyjne* i *redukcyjne*. Tymczasem na świecie logika nieformalna wyswobodziła się ze statusu gorszej siostry w dużo większym stopniu, niż miało to miejsce w Polsce (zob. Hamblin 1970), także w obliczu coraz jaskrawszych ograniczeń podejścia formalnego. W konsekwencji badania formalne zaczęły funkcjonować samodzielnie w oderwaniu od motywujących je problemów, których nie mogły rozwiązać. Zaczęto tworzyć logikę dla logiki: badania skupiły się na ustalaniu relacji między systemami, tworzeniu nowych systemów i analizie ich własności. Taki trend jest zapewne oznaką dojrzałości dyscypliny, ale może też wskazywać na tworzenie się nisz badawczych odseparowanych od reszty. Wykształcił się dominujący na wiele pokoleń model rozumowania, za którym poszły i praktyki dydaktyczne, czego efekt stanowią też głębokie podziały w ramach dyscypliny.

* * *

Jak ocenić dziedzictwo Szkoły w naukach o rozumowaniu? Nie da się jednoznacznie odpowiedzieć na takie pytanie. Ścisłość i precyzja wypowiedzi, tak cenione w Szkole, mają oczywiście swoje miejsce w niektórych kontekstach. Najważniejszą domeną jest tu matematyka: nic więc dziwnego, że logika formalna, stworzona z myślą o rozumowaniach matematycznych, najlepiej sprawdza się właśnie tutaj. Im dalej jednak od tego ideału nauki dedukcyjnej, tym więcej powodów, by wziąć pod uwagę inne typy rozumowań: *indukcję*, *abdukcję* i *rozumowanie przez analogię*.

Logika formalna została poważnie podważona przez dyscyplinę, która jeszcze pół wieku temu tchnęła w badania formalne nowe życie: informatykę. To tu rozwinęła się gałąź zwana *sztuczną inteligencją* (*artificial intelligence, AI*). Tak jak ponad 100 lat temu, mamy obecnie do czynienia z kryzysem. Chociaż logika początkowo odgrywała w *AI* kluczową rolę, to najnowsze systemy sztucznej inteligencji oparte są jednak nie na regułach dedukcyjnych, ale na *indukcji*.

Każdy jednak kryzys – oznaka nadchodzących głębokich zmian, na które musimy być gotowi – może być też szansą. Przecież i matematyka po kryzysie na przełomie XIX i XX wieku nie jest już tą samą dyscypliną: pomimo tego samego przedmiotu badań sama dyscyplina zmieniła się diametralnie w wyniku rewolucji, która się w niej dokonała. Podobnie zapewne będzie z nauką o rozumowaniu, w tym logiką formalną. Oznaki tego widać już dziś.

Paradoksalnie źródłem odnowy logiki mogą być właśnie najnowsze badania nad indukcyjnymi systemami sztucznej inteligencji. ChatGPT wiele potrafi, ale ma też sporo ograniczeń (Sejnowski 2024). Okazało się, że poprawne rozumowanie stanowi jedno z nich. Systemy takie potrzebują poznać logikę i nauczyć się poprawnie rozumować. Od pewnego czasu trwają prace nad modelami rozumowania (*reasoning models*) i trzeba przyznać, że postępy są

wyraźnie zauważalne. Dobrze widać też potrzebę sięgnięcia do źródeł – do praw logiki, poprawnych reguł rozumowania, jasności wypowiedzi i precyzji – czyli tego wszystkiego, co w szerokim sensie stanowi o logicznej spuściźnie Szkoły. Słowem, widoczna staje się potrzeba integracji nauk o rozumowaniu (bo jednak nie chodzi tu tylko o logikę formalną) z systemami *AI*, a ta praca dopiero się zaczyna.

Do tej pory wyróżnić można trzy podejścia. Po pierwsze, rozwija się tzw. *podejście neurosymboliczne*, które łączy architekturę sieci neuronowych z regułami formalnymi już na poziomie budowy systemu. S. Bader i P. Hitzler (2005) zaproponowali typologię takich systemów, w których jedną z osi stanowi wykorzystanie logiki zdań lub logiki predykatów. Po drugie, *rozwija się systemy indukcyjne w taką stronę, by „nauczyć je” poprawnego rozumowania*. T. Morishita i in. (2024) opisali trening modelu LLaMA-3.1-70B na przykładach symbolicznych rozumowań z logiki formalnej. Po trzecie, tworzy się *systemy hybrydowe*, wykorzystujące dwa komponenty: statystyczny i symboliczny. W artykule Z.Z. Rena i in. (2025) opisany jest system, w którym użyto modelu DeepSeek-V3 do wstępnej analizy problemu i jego dekompozycji na mniejsze zadania, które przekazano następnie do Lean 4, programu dowodzącego twierdzenia². *To wszystko stanowi potencjalnie nowy obszar zastosowania logiki, w którym badania dawnych mistrzów odnaleźć mogą swoje nowe miejsce*. Autor niniejszej pracy sam zaangażował się w takie badania, by sprawdzić poziom rozumowania i usprawnić te umiejętności w pierwszym polskim dużym modelu językowym (*large language model*, *LLM*) o nazwie Bielik³. Niech o potrzebie zaangażowania się logików w prace nad rozwojem współczesnej sztucznej inteligencji świadczy

² Więcej na temat samego programu: zob. de Moura i in. 2015.

³ System dostępny jest na stronie: <https://bielik.ai/>. Z jego opisem można zapoznać się w: Ociepa i in. 2024.

choćby to, że Bielik na pytanie: „Czy zdanie «Istnieje taki x , że $P(x)$, lub istnieje taki x , że $Q(x)$ » jest równoważne zdaniu «Istnieje taki x , że $(P(x) \text{ lub } Q(x))$ »” odpowiada (na razie) z rozbrajającą szczerością:

Niestety, nie mogę odpowiedzieć na to pytanie, ponieważ dotyczy ono logiki matematycznej, a ja nie posiadam wiedzy na ten temat. Moja wiedza jest ograniczona do informacji, które zostały mi przekazane podczas treningu, a logika formalna nie była jednym z nich. Przepraszam za niedogodności.

BIBLIOGRAFIA

- Bader S., Hitzler P. (2005), *Dimensions of Neural-symbolic Integration – A Structured Survey*, arXiv:cs/0511042.
- Hamblin C. (1970), *Fallacies*, London: Methuen & Co. Ltd.
- Łukasiewicz J. (1910), *O zasadzie sprzeczności u Arystotelesa. Studium krytyczne*, Kraków: Akademia Umiejętności.
- Morishita T., Morio G., Yamaguchi A., Sogawa Y. (2024), *Enhancing Reasoning Capabilities of LLMs via Principled Synthetic Logic Corpus*, arXiv:2411.12498.
- de Moura L., Kong S., Avigad J., van Doorn F., von Raumer J. (2015), *The Lean Theorem Prover (System Description)*, w: A.P. Felty, A. Middeldorp (red.), *Automated Deduction – CADE-25: 25th International Conference on Automated Deduction, Berlin, Germany, August 1–7, 2015, Proceedings*, Cham: Springer, s. 378–388.
- Ociepa K., Flis Ł., Wróbel K., Gwoździej A., Kinas R. (2024), *Bielik 7B v0.1: A Polish Language Model – Development, Insights, and Evaluation*, arXiv:2410.18565.
- Ren Z.Z., Shao Z., Song J., Xin H., Wang H., Zhao W., Zhang L., Fu Z., Zhu Q., Yang D., Wu Z.F., Gou Z., Ma S., Tang H., Liu Y., Gao W., Guo D., Ruan C. (2025), *DeepSeek-Prover-V2: Advancing Formal Mathematical Reasoning via Reinforcement Learning for Subgoal Decomposition*, arXiv:2504.21801.

- Russell B. (1903), *The Principles of Mathematics*, Cambridge: Cambridge University Press.
- Sejnowski T. (2024), *ChatGPT and the Future of AI: The Deep Language Revolution*, Cambridge, MA: The MIT Press.
- Trybus A. (2024), *What is the principle of (Non-)Contradiction, precisely? A Struggle at the Dawn of Formal Logic*, „Logic and Logical Philosophy”, s. 1–22.