

Česko-slovenské sympózium o analytickej filozofii
XXIX. ročník

Filozofická fakulta Univerzity Komenského v Bratislave
08.–10. 9. 2025



PROGRAM

08. 09. 2025

12:55 – 13:00	Otvorenie konferencie
13:00 – 13:30	V. Havlík: Emergentní vlastnosti velkých jazykových modelů
13:30 – 14:00	V. Gvoždiak: Inferenční schopnosti LLMs: Nejsem rasista, ale...
14:00 – 14:30	Prestávka
14:30 – 15:00	M. Duží: Syntactic vs Semantic Consistency in a Hyperintensional Lambda Calculus
15:00 – 15:30	F. Gahér: Pojem úplne holého individua a parciálny esencializmus jednotlivín
15:30 – 16:00	P. Dvořák: Morální logika Jana Caramuela z Lobkovic (1606–1682)
16:00 – 16:30	Prestávka
16:30 – 17:00	V. Punčochář: Role neúplné informace v problému svobodného rozhodování
17:00 – 17:30	V. Svoboda: Problém svobodného rozhodování/jednání a determinismus
17:30 – 18:00	J. Šebelová: Logicismus po korekci Fregeho systému
18:00 – 20:00	Raut

09. 09. 2025

09:30 – 10:00	D. Vacek: Systémy umelej inteligencie a ich roly
10:00 – 10:30	F. Tvrdý: Modální dezinformace
10:30 – 11:00	Prestávka
11:00 – 11:30	Z. Rybaříková: Filozofky vědy ve střední Evropě: srovnání Polska a Československa v první polovině dvacátého století
11:30 – 12:00	M. Zach: Od homeostáze k adaptaci: Jak organismy udržují stabilitu?
12:00 – 14:00	Obedná prestávka
14:00 – 14:30	L. Kvasz: Klasická epistemológia a matematické poznanie
14:30 – 15:00	P. Labuda: Aristoteles o jazyku ako konštitutívnom nástroji racionality: contra McCready-Flora
15:00 – 15:30	Prestávka
15:30 – 16:00	L. Zámečník: Mezi silnou a slabou autonomií nekauzálních vysvětlení
16:00 – 16:30	D. Makovník: Rekonštrukcie vedeckých explanácií pomocou

diferenciálnych rovníc.

16:30 – 17:00 **Prestávka**

17:00 – 17:30 T. Kollárik: Groarkeho deduktivizmus a úskalia kruhového uvažovania

17:30 – 18:00 L. Bokroš: Kompatibilita odvodu najlepšieho vysvetlenia a bayesiánskej teórie potvrdenia – kritická analýza

10. 09. 2025

09:30 – 10:00 P. Jošt: Proti spekulatívnej bioetice

10:00 – 10:30 A. Greif: Morálka ako zdieľaná identita

10:30 – 11:00 **Prestávka**

11:00 – 11:30 S. Spodniak: Máme egoistické dôvody konať morálne?

11:30 – 12:00 T. Čana: O objektívnej istote inak

12:00 **Ukončenie konferencie**

Toto sympóziu bolo podporené Agentúrou na podporu výskumu a vývoja na základe Zmluvy č. APVV-21-0405 a projektom VEGA 1/0557/23.

Emergentní vlastnosti velkých jazykových modelů

V příspěvku se zaměřím na otázku emergence nových schopností ve velkých jazykových modelech (LLMs) v závislosti na rostoucí škále parametrů těchto modelů. Současný nositel Nobelovy ceny za fyziku John J. Hopfield již v roce 1982 testoval teoretickou možnost vzniku emergentních (kolektivních) vlastností v rámci integrovaných obvodů, k čemuž byl inspirován svými výzkumy v neurobiologii, kde pracoval na molekulární dynamice molekul. Za emergentní vlastnosti považoval určitou schopnost generalizace, rozpoznávání známých vzorů, kategorizaci, opravy chyb a uchovávání časové posloupnosti (Hopfield 1982). V současných mnohem komplexnějších modelech hlubokých neuronových sítí (DNNs) se nyní zdá, že lze identifikovat kvalitativní skoky ve schopnostech při řešení komplexních úloh, které se objevují při překročení určitých škál parametrů těchto modelů. Pozornost je v této souvislosti věnována pokročilým funkcím, jako je víceúrovňové uvažování, matematické dokazování a kontextuální porozumění, které se objevují spontánně bez explicitního tréninku těchto modelů na tyto úlohy. Cílem těchto výzkumů je odhalení a porozumění nelineární závislosti mezi škálováním velikostí modelů a emergencí těchto schopností. Pozoruhodnou skutečností se zdá být fakt, že jsme schopni využívat těchto technologií umělé inteligence, aniž bychom přesně věděli, jak a proč fungují. Největší modely jsou nyní tak složité, že je třeba je zkoumat, jako zvláštní přírodní jevy, jejichž pochopení vyžaduje provádění experimentů pro vysvětlení jejich výsledků a schopností. Kromě těchto otázek výzkumu emergence (Havlík 2022) v nové, a vzhledem k technologické náročnosti dříve nedostupné doméně (DNNs), se objevuje i mnoho metodologických a konceptuálních otázek o vztahu AI a mysli. Lze oprávněně předpokládat, že emergentní jevy jsou z určité perspektivy zákonitostí vznikání nového ve světě, podobně jako jsou takto chápány některé fyzikální jevy a procesy? Lze touto analýzou přispět k obecnějšímu pochopení emergentních jevů nejen ve fyzice a teorii komplexity, ale i v biologii a fungování mozku? Můžeme předpokládat, že prokazují-li modely neurálních sítí vznik emergentních vlastností, dochází k obdobným jevům i v rámci neuronových sítí biologického orgánu mozku a může tento výzkum přispět i k pochopení mysli a mentálních vlastností? Posledně vyvinuté modely, na rozdíl od stávajících modelů (ChatGPT a dalších), se tyto otázky stávají mnohem aktuálnější a podrobnější rozbor těchto přístupů může poskytnout mnohem přesvědčivější odpovědi. Příkladem může být „Centaur: a foundation model of human cognition“ (Binz et al. 2024), který na rozdíl od modelů jako je ChatGPT a další, vykazuje překvapivé emergentní schopnosti generalizace a „out-of-distribution“ (OOD), že je považován za bránu k naší mysli (Centaur, a Door to Our Minds?).

Literatura

Binz, Marcel, Elif Akata, Matthias Bethge, Franziska Brändle, Fred Callaway, Julian Coda-Forno, Peter Dayan, et al. 2024. 'Centaur: A Foundation Model of Human Cognition'. arXiv. <https://doi.org/10.48550/ARXIV.2410.20268>.

Havlík, Vladimír. 2022. *Hierarchical Emergent Ontology and the Universal Principle of Emergence*. Cham: Springer International Publishing. <https://doi.org/10.1007/978-3-030-98148-8>.

Hopfield, J J. 1982. 'Neural Networks and Physical Systems with Emergent Collective Computational Abilities.' *Proceedings of the National Academy of Sciences* 79 (8): 2554–58. <https://doi.org/10.1073/pnas.79.8.2554>.

Vít Gvoždiak

Inferenční schopnosti LLMs: Nejsem rasista, ale...

Pozornost, která je v posledních letech věnována inferenčním a pragmatickým schopnostem velkých jazykových modelů (LLMs), je charakteristická na jedné straně tím, že je upřena na otázky spojené s konverzačními implikaturami či presupozicemi (viz např. Sravanthi et al. 2024), na straně druhé tím, že se výzkumy orientují zejména na high- či rich-resource jazyky (srov. tzv. English-centered bias; Bolei Ma et al. 2025). Příspěvek prezentuje výsledky výzkumu, který se pokouší obě naznačené charakteristiky obohatit. Základní idea vychází z empirického šetření, které provedli Leen Janssens a Walter Schaeken (2013), kteří se zaměřili na to, zda lidé provádějí inference spouštěné spojkou „ale“ a následované výrazy „takže“ a „nicméně“ v argumentační konstrukcích typu „A, ale B“. Prezentovaný výzkum tento rámec aplikuje v prostředí českých mluvčích a doplňuje ho o konverzační situace, které v argumentační struktuře pracují s vyjadřováním stereotypů (věku, genderu a národnosti/ethnicity), tedy velmi blízké schématu tzv. konvenční implikatury (Grice 1989). Pro účely výzkumu byl vytvořen dataset 48 situací, jejichž cílem je ověřit inferenční tendence v neutrálních a stereotypních ale-konstrukcích. Experiment 1 sleduje způsoby, jakým tyto konstrukce interpretují lidští (čeští) mluvčí (N=30), experiment 2 pak způsoby zpracování různými typy LLMs a komparuje jejich výkony s tendencemi českých mluvčích.

Marie Duží

Syntactic vs Semantic Consistency in a Hyperintensional Lambda Calculus

I address and solve a technical problem arising in the hyperintensional, procedural semantics of Transparent Intensional Logic (TIL). The problem primarily concerns variable binding and correct substitution. In TIL, there is a pair of twin procedures, namely Trivialisation and Double Execution. While Trivialisation raises a context to the hyperintensional level, Double Execution cancels the effect of Trivialisation. The pair of twin procedures work together well in the standard cases. But there are limiting cases where they fail to. Trivialization, when applied to another procedure, makes it technically feasible to operate on this procedure itself rather than on the product it is designed to yield. Any occurrences of variables within the procedure become Trivialization-bound. Trivialization-binding contrasts with lambda-binding. The other

procedure (Double Execution), when applied to another procedure, is typed and structured to, first, obtain the product of this procedure and, second, obtain the product of this product on condition that the latter is itself a procedure. As a result, there are cases where Trivialization-bound occurrences of variables turn into free occurrences. Moreover, and even worse, new free variables can crop up in the process of computing Double Execution. These cases undermine our definition of substitution and yield inconsistencies, thus jeopardizing the validity of the rules of lambda-conversion, not least beta-conversion. To solve the problem, we must distinguish two kinds of computing: syntactic conversions and semantic evaluation of procedures. While syntactic lambda-conversions bear no relation to the valuation of free variables, semantic evaluation is valuation-dependent. The technical result I present is a set of definitions that introduce the restrictions required for syntactic computation, along with a proof of consistency. Additionally, I compare plausible methods of semantic evaluation. The problem I address and solve extends beyond Transparent Intensional Logic. Any logic or programming language (especially if based on lambda calculus) furnished with a high degree of expressive power, in which procedures can occur as operands, is liable to confront a comparable problem with variable binding and substitution.

František Gahér

Pojem úplne holého individua a parciálny esencializmus jednotlivín

Tichý založil reinterpretáciu Wittgensteinovho *Traktátu* na pojme abstraktnej procedúry a pojme holého individua. Táto reinterpretácia Wittgensteinových komplexov ako abstraktných procedúr je veľmi blízka Fregeho explikácii myšlienok, ktoré sú zložené, a to výlučne z abstraktných entít. Pojem holého individua, ako bol v neskorých prácach špecifikovaný a prezentovaný Tichým je pojmom modálneho - *empiricky* holého individua. Tichý odmieta individuový esencializmus a obhajuje intenzionálny esencializmus. Jeho individuový antiesencializmu vychádza z odmietnutia empirickej vševedúcnosti skúmateľa. K otázke holého individua prebehla neskôr v TIL-kárskom kruhu vcelku bohatá diskusia, o čom svedčia najmä práce z pier Koláňa (1999, 2000), Gahéra (1999), Svobodu (2001), Jespersena a Maternu (2002), Duží (2004), Cmoreja (2006, 2010), Raclavského (2007, 2008, 2009). Antiesencializmus, ako ho predstavil (ale neobhajoval) Stalnaker, však pôvodne vychádzal z odmietnutia akejkoľvek vševedúcnosti, nielen empirickej. Vystáva tu však menej pertraktovaná otázka, že ak sme oprávnené odmietli v nejakej podobe empirickú vševedúcnosť či už na základe možno-svetových abstrakcií či na základe rýdzo epistemologických úvah, prečo sme neodmietli aj analytickú (logickú) vševedúcnosť (Stalnaker 1991) a nepracujeme aj s pojmom *úplne* holého individua, t.j. aj analyticky holého individua. Pojem úplne holého individua môžeme charakterizovať ako zložku abstraktných procedúr, ktorá nemá žiadne vlastnosti okrem sebaidentity, teda ani žiadne iné analytické vlastnosti, je čírym individuátorom, prázdny vešiakom pre vlastnosti (a vzťahy) a je stálym prvkom univerza v každom možnom svete. Sémantický omyl logického atomizmu, ktorý považoval konkrétne jednotliviny za zložky komplexov, bol založený na doktríne tzv. tropizmu (*partikularizmu*). V maskovanej podobe tento tropizmus predkladá ako *materiál*, látka, z ktorej je utvorené individuum, a ktorá je zdanlivo nositeľom jeho identity. V skutočnosti je cieľom týchto úvah konkrétna jednotlivina a parciálny esencializmus časo-priestorových

jednotlivín je overené a pragmaticky fungujúce vysvetlenie ich identity počas časovo obmedzenej existencie.

Literatúra

CMOREJ, P., TICHÝ, P. (1998): Komplexy I, II. *Organon F*, 5 (2), 139 – 161; (3) 266 – 289.

CMOREJ, P. (2006): Holé individua a predikácia. In: M. Zouhar (ed.): *Jazyk z pohľadu sémantiky, pragmatiky a filozofie vedy*. Bratislava: Filozofický ústav SAV; Praha: Filozofický ústav Akademie věd České republiky v Prahe, 137-161. Vyšlo aj v práci (Cmorej 2009, 141-171).

DUŽÍ, M. (2004): Intensional Logic and the Irreducible Contrast between de dicto and de re. *ProFil* Vol. 5, No. 1, 1 – 34, http://profil.muni.cz/01_2004/duzi_de_dicto_de_re.pdf (ISSN 1212-9097).

JESPERSEN, B. – MATERNA, P. (2002): Are Wooden Tables Necessarily Wooden? *Acta Analytica* 17, Issue 28, 115 – 150.

GAHÉR, F. (2025): Fregeho *Der Gedanke* ako konkurenčný projekt k Wittgensteinovmu raktátu. *Filozofia*, 80 (3), 285 – 299. DOI: <https://doi.org/10.31577/filozofia.2025.80.3.2>

KOLÁŘ, P. (2000): Clothing Bare Individuals. In: Majer, O. (ed.): *Topics in Conceptual Analysis and Modelling*. Filosofia, Praha, 126 – 140.

MAURIN, A. (2024): *Tropes*. *The Stanford Encyclopedia of Philosophy* (Winter 2024 Edition), Edward N. Zalta & Uri Nodelman (eds.), URL = <https://plato.stanford.edu/archives/win2024/entries/tropes/>.

RACLAVSKÝ, J. (2007): Against Kolář's Criticism of Tichý's Bare Individuals, *Pro-Fil* 8, 2. ISSN 1212-9097. http://profil.muni.cz/02_2007/raclavsky_against_kolars_criticism.pdf

RACLAVSKÝ, J. (2008): Reformulating Tichý's Conception of Bare Individuals. *Organon F* 15, No. 1.

RACLAVSKÝ, J. (2009): Individua v myšlenke-propozici: Tichého přístup. In: *Filozofia*, 64 (7), 669 – 679.

STALNAKER, R. (1984): *Inquiry*. The MIT Press, Cambridge.

STALNAKER, R. (1991). The Problem of Logical Omniscience, I. In: *Synthese*, roč. 89, č. 3, Belief and Rationality, s. 425-440. (1999)

SVODODA, V. (2001): Individua na odpočinku. In: *Filosofický časopis* 49, č. 3, 415 – 424.

TICHÝ, P. (1987): Frege and the Case of the Missing Sense. *Grazer Philosophische Studien*, Vol.27, 27- 48.

TICHÝ, P. (1999): Traktát ve světle intenzionální logiky, *Filosofický časopis* 49, 249 – 262.

Petr Dvořák

Morální logika Jana Caramuela z Lobkovic (1606–1682)

Příspěvek se zabývá dosud všeobecně neznámou a nedostatečně prozkoumanou morální logikou Jana Caramuela z Lobkovic (1606–1682), jak ji načrtl ve svém rozsáhlém pozdním díle, druhém svazku *Moralis seu Politica Logica* (1680). Morální logika se zabývá inferencemi zahrnujícími kvantifikované výroky. Kromě dvou latinských verzí aristotelské univerzální kvantifikace „každý“ a „žádný“ (*omnis* a *nullus*) Caramuel rozlišuje několik dalších partikulárních kvantifikátorů, které jsou jakoby „mezi“ těmito dvěma. Jsou to „téměř všichni“, „většina“, „polovina“, „méně než polovina“, „téměř nikdo“ (v nesylogistické části jeho logiky také další dvě formy kvantifikátorů *plerique omnes* a *fere nulli*). Kromě těchto existují také relativní kvantifikátory „mnozí“ a „nemnozí“. Sylogismy se skládají z kvantifikované partikulární (nebo univerzální) premisy, singulární premisy a singulárního závěru. Epistemický status závěru (pravděpodobnější, pravděpodobný atd.) odpovídá morálnímu kvantifikátoru v premise. Pokud je epistemický status přímo uveden (jako epistemický modus ovlivňující závěr) ve výroku závěru, pak je epistemický status takto rozšířeného závěru „jistý“ a argument je formálně platný. Caramuel věnuje pečlivou pozornost modům v první aristotelské figuře a rozšiřuje platný modus *Darii* na devět módů. Později v knize uvádí dalších devět módů v platónské (čtvrté) figuře a 36 módů ve druhé aristotelské figuře. Platné mody ve třetí aristotelské figuře již ponechává na čtenáři. Nástin morální sylogistiky, který Caramuel předkládá, je míněn jako nástroj argumentace v oblasti praktických věd (morální teologie, etika) a především v právním prostředí (u soudu). Z dnešního pohledu se jedná o fragment epistemické logiky a fragment fuzzy logiky v rámci aristotelské sylogistické logiky termínů.

Vít Punčochář

Role neúplné informací v problému svobodného rozhodování

Ve svém příspěvku se budu věnovat tradiční filosofické debatě o svobodné vůli a možnosti existence svobodně se rozhodujících subjektů v deterministickém světě. Tuto debatu pokládám za relevantní i za předpokladu, že náš svět ve skutečnosti deterministický není, neboť nás nutí ke konceptuálnímu projasňování toho, v čem naše svobodné rozhodování spočívá a v čem nikoli. Ve své přednášce se budu vymezovat zejména proti následujícím tezím, které, domnívám se, reprezentují základní verzi nekompatibilistického postoje:

Kdyby byl svět deterministický, tak by (a) nebyl z hlediska otázky svobodné vůle žádný podstatný rozdíl mezi tím, co běžně pokládáme za svobodné rozhodnutí, a mezi jakýmkoli jiným procesem ve fyzikálním světě; (b) neměli bychom žádnou kontrolu nad svými budoucími činy (jelikož bychom neměli kontrolu nad minulostí, která by tyto činy plně determinovala); (c) nezáleželo by na tom, jak se v dané situaci rozhodneme (neboť by vše již bylo předem determinováno), a nemělo by tedy smysl, abychom se o něco snažili; (d) nikdo by nemohl nést odpovědnost za své činy. Teprve nedeterministický svět otevírá prostor pro skutečnou

svobodnou vůli, neboť plnohodnotné svobodné rozhodnutí předpokládá, že za daných fixovaných podmínek existuje více možných variant budoucnosti, mezi kterými se rozhodujeme.

Mezi výrazné představitele nekompatibilismu patří např. Sam Harris či Robert Sapolsky. Jejich nekompatibilistický postoj se dá také shrnout následujícím konstatováním: pokud je svět deterministický, tak máme jen iluzi toho, že naše rozhodování je svobodné. V kontrastu s tímto postojem budu hájit tezi, že máme sklony k přirozené a systematické kognitivní iluzi, že v deterministickém světě bychom se nemohli svobodně rozhodovat. Pokusím se tento kognitivní klam rozptýlit prostřednictvím analýzy modálních pojmů, které používáme při artikulaci toho, v čem spočívá determinismus, proces svobodného rozhodování, zodpovědnost či kontrola. Na řadě příkladů budu ilustrovat, že běžně používáme modalitu možnosti, které nejsou čistě epistemické, tj. mají kromě subjektivního i výrazný objektivní obsah, i když ten je relativní vůči neúplné informaci, přičemž rozsah této neúplnosti je určen kontextuálně. Pokusím se ukázat, že nenápadné fluktuace takovéto informační neúplnosti, které jsou implicitně přítomné v běžném používání modálních termínů, jsou zdrojem naší přirozené iluze nesvobody v deterministickém světě.

Vladimír Svoboda

Problém svobodného rozhodování/jednání a determinismus

Ve svém příspěvku se budu věnovat tradiční filosofické debatě o svobodné vůli a o (ne)možnosti existence svobodně se rozhodujících subjektů v deterministickém světě. Budu se snažit ukázat, že Punčochářova (a Dennettova) argumentace, podle níž mohou být konatelé svobodní v deterministickém světě (nebo dokonce i ve světě, který je deterministicky zacyklený tak, že se v něm veškeré děje donekonečna opakují), je obecně nepřesvědčivá. Nejprve se pokusím obhájit tezi, že debaty o svobodné vůli trpí tím, že nedoceňují potřebu co nejdůslednějšího vyjasnění klíčových pojmů, zejména pojmů rozhodnutí a svobodného rozhodnutí, resp. svobodného jednání. To podle mne v důsledku vede k podcenění úlohy, kterou ve světě, v němž žijeme, hraje nedourčenost, resp. podjatá/omezená nahodilost. Pokusím se rozlišit různá chápání pojmu rozhodnutí jedince (zvířete, člověka, anděla, boha) a různá pojetí více či méně svobodného rozhodnutí. V návaznosti na to zkusím ukázat, že jenom některá pojetí svobodného jednání jsou slučitelná s představou, že historie našeho světa je adekvátně popsána jako postupná aktualizace světamžiků v jediné předem určené světahistorii, z níž z principu nelze žádným způsobem – a tedy ani na základě nahodilosti či (svobodného) rozhodnutí - vybočit. Budu se snažit obhájit přirozenou představu, podle níž budoucí stavy světa nejsou předem předurčeny a rozhodnutí, resp. spíše jednání, jsou svobodná v různé míře. Jistou míru svobodného jednání lze přitom připsat i zvířatům, případně i robotům. Osudy těchto konatelů nejsou podle této představy předem jednoznačně určeny, i když mohou být určeny do jisté, někdy i značné míry.

Jana Šebelová

Logicismus po korekci Fregeho systému

Na konci 19. století formuloval Gottlob Frege program logicismu s cílem ukázat, že aritmetika je odvoditelná z logiky prostřednictvím logických zákonů. Logický systém, který pro tento účel vytvořil, se však ukázal být nekonzistentní. O několik desetiletí později byla chyba v jeho logickém systému opravena a tím byl tento systém rehabilitován. To ovšem otevřelo otázku, zda korekce Fregeho systému vede k úspěšnému naplnění programu logicismu. Tento problém lze redukovat na otázku povahy jediného mimologického prvku systému, tzv. Humova principu. Pokud je tento princip analytický, pak by bylo skutečně možné matematiku odvodit z logiky a program logicismu by tak byl úspěšně završen. Pokud však analytický není, nelze logicismus považovat za naplněný. Otázka analytičnosti Humova principu je proto zásadní a příspěvek se věnuje analýze toho, jak byla a je tato otázka diskutována v odborné literatuře. V první části příspěvku se zaměříme na vymezení toho, v čem Fregeho program logicismu spočívá a jaké má tento pojem a program historické kořeny. Podrobněji se budeme věnovat Fregeho pojetí analytického výroku a jeho vztahu k tradici. Ve druhé části se soustředíme na novodobou diskuzi o povaze Humova principu, konkrétně na polemiku mezi Crispinem Wrightem, který hájí jeho analytičnost, a Georgem Boolosem, jenž tuto pozici zpochybňuje. Na tomto příkladu představíme hlavní argumenty obou stran a ukážeme, v čem se jejich pozice liší a co je vede k rozdílným odpovědím na otázku analytičnosti Humova principu. Komparace závěrů první a druhé části umožní hlubší analýzu toho, jakou hraje to které pojetí analytičnosti v rozhodování o ne/úspěšnosti Fregeho logicistického programu roli. Tak ukážeme, z jakých perspektiv lze Fregeho program logicismu po rehabilitaci jeho logického systému považovat za naplněný a z jakých nikoliv.

Daniela Vacek

Systémy umelej inteligencie a ich roly

Fráza „zodpovedná umelá inteligencia“ je (nad)používaná, no jej význam ostáva nejasný. Frázu možno vyjasniť nepriamo, prostredníctvom zodpovedného vývoja, nasadzovania, či používania umelej inteligencie. Táto stratégia je v literatúre pomerne častá (Bryson 2018, Constantinescu et. al. 2021, Dignum 2019, Glavaničová, Pascucci 2022, Hakli, Mäkelä 2019, Kiener 2022). V doslovnom význame by sme však mali hovoriť priamo o umelej inteligencii – ako v nejakom zmysle zodpovednej. Táto stratégia je pomerne nepopulárna (k výnimkám možno zaradiť články Sullins 2011, Altehenger, Menges 2024 a Altehenger, Menges, Schulte 2024). V tomto príspevku sa budem uberať nepopulárnou cestou. Narozdiel od vyššie citovaných článkov sa však zameriam na „slabší“ pojem zodpovednosti, tzv. *role responsibility*, zodpovednosť rolí (Cane 2002, Cane 2016, Dare 2016, Dare, Swanton 2019). Budem argumentovať, že aplikácia tohto pojmu v oblasti umelej inteligencie je sľubná a má svoje výhody aj v prípade jeho aplikácie priamo na systémy umelej inteligencie.

Literatúra

Altehenger, H. a Menges, L. (2024) The point of blaming AI systems. *Journal of Ethics and Social Philosophy*, 27(2).

Altehenger, H., Menges, L. a Schulte, P. (2024) How AI Systems Can Be Blameworthy. *Philosophia*, 124.

Bryson, J. J. (2018) Patency is not a virtue: the design of intelligent systems and systems of ethics. *Ethics and Information Technology*, 20(1): 15–26.

Cane, P. (2002) *Responsibility in law and morality*. Bloomsbury Publishing.

Cane, P. (2016) Role responsibility. *Journal of Ethics* 20: 279–298.

Constantinescu, M., Voinea, C., Uszkai, R. a Vică, C. (2021) Understanding responsibility in Responsible AI. Dianoetic virtues and the hard problem of context. *Ethics and Information Technology* 23: 803–814.

Dare, T. (2016) Robust role-obligation: How do roles make a moral difference?. *The Journal of Value Inquiry*, 50: 703–719.

Dare, T. a Swanton, C. (Eds.). (2019) *Perspectives in Role Ethics: Virtues, Reasons, and Obligation*. Routledge.

Dignum, V. (2019) *Responsible artificial intelligence: how to develop and use AI in a responsible way (Vol. 2156)*. Cham: Springer.

Glavaničová, D. a Pascucci, M. (2022) Vicarious liability: a solution to a problem of AI responsibility?. *Ethics and Information Technology* 24(3), 28.

Hakli, R. and Mäkelä, P. (2019) Moral responsibility of robots and hybrid agents. *The Monist* 102(2): 259–275.

Kiener, M. (2022). Can we Bridge AI's responsibility gap at Will?. *Ethical Theory and Moral Practice* 25(4): 575–593.

Sullins, J. P. (2011) When is a robot a moral agent. In Anderson, M., and Anderson, S. L. (eds.). *Machine ethics* (pp. 151–161). New York: Cambridge University Press.

Aplikovaná epistemologie odlišuje dezinformace od misinformací na základě intencionality. Dezinformace jsou považovány za úmyslně vytvořené nepravdivé informace, které jsou šířeny se záměrem uškodit. Misinformace jsou pak neúmyslně zastávané nepravdivé informace, které jsou šířeny bez nekalého záměru. Rozlišení není v běžném epistemickém provozu příliš použitelné, a to hned z několika důvodů. *První potíží* spočívá v kritériu pravdivosti, na jehož povahu existuje ve filozofii několik konkurenčních pohledů – korespondenční, koherenční, pragmatické a deflační teorie. Dezinformace většinou neobsahují nějakou inherentní vadu, ale jsou nedostatečně zdůvodněné, tj. pocházejí od nedůvěryhodných svědků nebo z nespolehlivých zdrojů poznání. Některé dezinformace navíc nejsou lživé, ale „pouze“ zavádějící. Například tvrzení „Senátor podlehl plicní embolii tři týdny po podání vakcíny proti COVID-19.“ nemusí být nutně nepravdivé, ale postrádá kontext potřebný k odmítnutí údajně kauzálního vztahu mezi antecedentem a konsekvencí. *Druhá potíží* se týká metodologického behaviorismu, který přírodní i sociální vědy přijaly jako své paradigma a v němž odvolávání se na záměrnost nemá místo. Lidé nemají bezprostřední přístup k myslím jiných osob, a proto většinou není možné odlišit úmyslně vytvořené dezinformace a neúmyslně zastávané misinformace. Existuje avšak zvláštní druh dezinformací, kterému zatím výzkumníci nevěnovali patřičnou pozornost. Jsou to dezinformace, které využívají aparát kontrafaktuálního myšlení, a jež budu proto označovat jako *modální dezinformace*. Mluvčí v takovém případě explicitně přiznávají nepravdivost prezentovaného sdělení v aktuálním světě, ale zároveň konstatují zdůvodněnost vyššího řádu ve světech možných. Modelovým případem v českém veřejném prostoru jsou výroky „Sice se to nestalo, ale mohlo by.“ nebo „Sice to není pravda, ale mohla by být.“ Ve svém příspěvku uvedu konkrétní příklady modálních dezinformací z politického diskurzu, následně provedu jejich logicko-sémantickou a pragmatickou analýzu. V závěru poukážu na překvapivé selhání analytické filozofie, která se s modálními dezinformacemi nedokáže patřičně vypořádat. Naneštěstí se zdá, že se s obdobnými potížemi potýkají i všechny ostatní směry současné filozofie, což odhaluje závažnou slabinu v samotném jádru filozofické metodologie.

Zuzana Rybaříková**Filozofky vědy ve střední Evropě: srovnání Polska a Československa v první polovině dvacátého století**

Ženy v rané vědecké filozofii zůstávají do značné míry neznámé, přestože některé z nich přispěly významným způsobem k rozvoji oboru. Tuto skutečnost lze přičíst dvěma hlavními faktorům. Za prvé, tyto ženy byly mezi prvními, které navštěvovaly univerzity. Jejich akademická kariéra proto byla vzhledem k převládajícímu společenskému prostředí počátku 20. století náročná. Za druhé, se zabývaly tématy, jako je logika a filozofie vědy, tedy obory, které byly (a někdy stále jsou) považovány za převážně „mužské“. V důsledku toho jsou v diskuzích o filozofkách často přehlíženy. Situace je pak ještě náročnější, pokud taková filozofka

žila v regionu, který je dnes známý jako východní Evropa. Můj příspěvek se bude týkat přínosu filozofek vědy z Polska a Československa v první polovině 20. století. Přednáška stručně představí historické a sociologické pozadí jejich studia a další vědecké práce. Hlavní část je zaměřena na jejich přínos filozofii. Uvedu rozdíly v jejich práci, ale zejména podobná témata, která lze v jejich díle nalézt. Tedy důraz na výuku logiky, otázka metodologie vědy, místo metafyziky ve vědě, problémy zákonů empirických věd, determinizmu a jednoty vědy. Chci tvrdit, že vzhledem k tomu, že v meziválečném období nedošlo mezi těmito filozofkami ke kontaktu, pramení tato podobnost ze skutečnosti, že filozofky v Československu a Polsku vycházely ze stejné tradice. Kromě toho také poukáži na to, že mezi ženami-filozofkami a muži-filozofy, kteří se zabývali vědeckou filozofií, nepanoval co do výběru témat v meziválečném období rozdíl.

Martin Zach

Od homeostáze k adaptaci: Jak organismy udržují stabilitu?

Homeostáze je dlouhodobě vnímána jako klíčový princip biologické regulace, který zajišťuje stabilitu fyziologických systémů. Tradičně se homeostáze spojuje se zdravím a návratem k bazálnímu stavu. Zatímco různé vědní obory používají tento termín, jejich interpretace se liší – kybernetika zdůrazňuje zpětnou vazbu, imunologie rozšiřuje homeostázu na aktivované stavy a onkologie dokonce zkoumá možnost, že homeostáze sama může podporovat patologii. Tato studie se zabývá různými způsoby, jimiž organismy udržují svou stabilitu. Vedle klasické homeostáze hraje zásadní roli fyziologická adaptace, která umožňuje přežití prostřednictvím strukturálních či funkčních změn, namísto pouhého návratu k předchozím stavům. Pomocí případových studií z oblasti imunologie a metabolické regulace představujeme rámec pro kategorizaci mechanismů udržování sebe sama, který rozlišuje mezi přímou homeostatickou korekcí, adaptivními posuny a trénovanými odpověďmi. Je nutné kriticky přehodnotit samotný koncept homeostáze, který může být příliš vágní nebo nedostatečně definovaný. Pokud trénovaná imunita vede k změně bazálního stavu místo návratu k předchozí konfiguraci, zpochybňuje to klasický homeostatický model? Biologickou stabilitu je nutné vnímat jako spektrum, nikoli jako jeden rigidní proces. Tímto přístupem navrhujeme širší model, který lépe postihuje variabilitu napříč biologickými systémy a vědními obory. Tato studie přispívá k hlubšímu porozumění biologické resilience a nabízí nový pohled na to, jak organismy přetrvávají v dynamických podmínkách.

Ladislav Kvasz

Klasická epistemológia a matematické poznanie

V článku „Aristotelova sylogistická logika jako teorie aritmetického typu“ (*Filosofický časopis* 72, 2024/1, s. 3–22) sme navrhli rozlíšiť tri druhy teórií: teórie aritmetického typu, teórie matematického typu a teórie fyzikálneho typu. V spomínanom článku sme sa pokúšali ukázať,

že Aristotelova sylogistická logika je vybudovaná ako teória aritmetického typu. Táto skutočnosť je vysvetlením, prečo sylogistická logika nebola v antike použitá ani v matematike Euklidom alebo Archimedom, a nepoužíval ju ani Aristoteles vo svojom vedeckom diele. V texte, ktorý navrhujem pre ČSSAF ktoré sa uskutoční v Bratislave v septembri 2025 navrhнем použiť podobnú argumentáciu pre prípad klasickej epistemológie. Pokúsim sa ukázať, že tvorcovia klasickej epistemológie ako Descartes, Locke, Hume alebo Kant pri budovaní svojich koncepcií poznania, chápali poznanie v zmysle teórie matematického typu. V prípade Descarta to bolo viac-menej prirodzené, a zdá sa, že Locke, Hume aj Kant prevzali Descartovo chápanie poznania, možno bez toho, aby si to uvedomili. Keď sa z tejto perspektívy pozrieme na charakterizáciu poznania ako zdôvodneného pravdivého presvedčenia, matematické pozadie tejto charakterizácie je očividné. To v matematike považujeme určitú hypotézu za poznatok, keď je dokázané. A dôkaz nie je nič iného než špecifický spôsob *zdôvodnenia*. Tým, že určité tvrdenie dokážeme, uistíme sa v jeho *pravdivosti*. Darom Newton v slávnom *Scholium Generale* upozorňoval, že v prípade empirického poznania si jeho pravdivosťou nikto nemôžeme byť istý, karteziáni, ktorí v tej dobe prevládali, stále poznanie spájali s pravdivosťou. A hypotézu, ktorou proces dokazovania začína, možno vnímať ako určitý druh *presvedčenia*. Teda je to matematické poznanie, ktoré možno považovať za zdôvodnené pravdivé presvedčenie. V príspevku sa pokúsim ukázať, že aj mnohé ďalšie črty, ktorými tvorcovia klasickej epistemológie charakterizovali poznávanie, majú svoje korene vo vlastnostiach špecifického chápania matematického poznania ako jasného a rozlíšeného vnemu, ktoré Descartes použil pri svojom pokuse založiť prírodné vedy (optiku, mechaniku, fyziológiu a ďalšie).

Pavol Labuda

Aristoteles o jazyku ako konštitutívnom nástroji racionality: contra McCready-Flora

Cieľom môjho príspevku bude preskúmať Aristotelov pohľad na vzťah medzi jazykom a racionalitou a obhájiť tézu, že podľa Aristotela je ľudský jazyk nástrojom, ktorým sa konštituuje naša racionalita. Takýto interpretačný prístup má síce medzi bádateľmi pomerne pevné základy (napr. Labarrière 1984, Wedin 1988, Osborne 2007 či Newmyer 2016), nedávno však bola publikovaná extenzívna kritika tohto prístupu (McCready-Flora 2019), ktorá si podľa mňa zasluhuje koncentrovanú pozornosť. McCready-Flora tvrdí, že u Aristotela nie je možné, aby výkon ľudskej reči konštituoval racionalitu, pretože v celej jeho filozofii dôsledne platí princíp kognitívnej priority. Tento princíp hovorí, že existencia určitej kognitívnej schopnosti vysvetľuje existenciu jej realizujúceho nástroja, nie opačne. Podľa Aristotela príroda poskytuje živočíšnym druhom iba také nástroje, ktoré zodpovedajú ich schopnostiam. Skutočnosť, že určitý živočíšny druh (napr. človek) disponuje určitou kognitívnou schopnosťou (racionalitou), teda vysvetľuje to, prečo má tento druh primeraný nástroj na jej prejavenie (reč/jazyk), a nie naopak. Dôležitým momentom tejto interpretácie teda je, že jej autor predstavuje Aristotela ako niekoho, kto chápe ľudský jazyk ako externý prostriedok (voči internej kognitívnej schopnosti). Takúto interpretáciu považujem za neudržateľnú. Vo svojom príspevku poukážem na tie časti Aristotelovho diela, v ktorých ľudský jazyk nie je chápaný ako nástroj prírody (*kata physin*). Na rozdiel od všetkých ostatných foriem zvukovej a hlasovej

komunikácie u neľudských živočíšnych druhov Aristoteles o ľudskom jazyku uvažuje ako o nástroji úplne odlišnej povahy, ako o nástroji normatívnej či symbolickej praxe (*kata synthēkēn*). V inferencialistických pojmoch môžeme hovoriť o Aristotelovom chápaní jazyka ako nástroja, prostredníctvom ktorého my ľudia uskutočňujeme hru na vyžadovanie a uvádzanie dôvodov nášho konania. Takto chápaný jazyk však určite nie je externým nástrojom. Je to skôr spôsob realizácie racionality. Navyše, takáto interpretácia jazyka je v súlade s postrehmi Aristotela, týkajúcimi sa jazykovej štruktúry nášho myslenia: naša racionalita je propozicionálna, má štruktúru a normativitu nášho jazyka. Mylnosť interpretačnej pozície McCreadyho-Floru ďalej preukážem aj na jeho nesprávnom výklade Aristotelom použitej analógie medzi jazykom a flautou. Podľa McCready-Florovho výkladu je v Aristotelovom rozvrhu jazyk analogický s hudobným nástrojom, menovite s flautou, a racionálny činiteľ analogický s flautistom. Ak by však bol takýto výklad správny, tak by zároveň malo byť možné (s ohľadom na vyššie opísanú interpretačnú stratégiu McCreadyho-Floru), aby bol niekto flautistom aj bez toho, aby sa najskôr naučil správne hrať na flaute – čo je samozrejme nezmysel. To, čo podľa môjho názoru McCreadyho-Florov výklad opomína, je práve interná, normatívna dimenzia ľudskej jazykovej praxe. Ľudský jazyk predsa nie je analogický s nehraným či so zle hraným nástrojom. Náš jazyk je analogický so správne používaným nástrojom, s nástrojom používaným podľa určitých, spoločensky akceptovaných noriem. My, ľudia, koordinujeme naše životy prostredníctvom jazyka normatívne, nie prírodne. A v súlade s týmto sú hodné zmienky i ďalšie dva fakty, týkajúce sa Aristotelovej filozofie jazyka: (i) Aristoteles na žiadnom mieste svojho diela nespomína nepropozičné, nejazykové či predjazykové racionálne myslenie, a (ii) podľa Aristotela deti nevykazujú racionalitu, až kým si dostatočne neosvoja jazyk. Téza, že náš jazyk je nástrojom konštituovania našej racionality teda znamená, že normatívna jazyková prax iniciuje a rozvíja našu racionalitu. Jazyk nám pomáha realizovať to, k čomu máme ako druh dispozíciu – byť racionálnymi. Musíme si však zároveň uvedomiť, že podľa Aristotela sú živočíšne druhy večné. Preto skutočnosť, že jazyk konštituuje racionalitu, u Aristotela nemožno vykladať lineárne. Jazyk nie je len externou príčinou a racionalita nie je len jej kognitívnym dôsledkom. Jazyk treba chápať skôr ako prostriedok, ktorým sa racionalita iniciuje a rozvíja u jednotlivcov narodených do spoločenstva racionálnych, jazykovo-kompetentných bytostí. V druhovo-eternalistickom explanačnom rámci sa totiž príčiny a dôsledky uskutočňujú v kruhu.

Lukáš Hadwiger Zámečník

Mezi silnou a slabou autonomií nekauzálních vysvětlení

Příspěvek naváže na probíhající debatu (a na můj loňský příspěvek „O autonomii nekauzálních vysvětlení“) o vztahu některých nekauzálních vědeckých vysvětlení k vysvětlením mechanistickým. Většina argumentů ve prospěch neredukovatelnosti nekauzálních vysvětlení se opírá o nějaké pragmatické kritérium (Kostić 2023, Kostić 2018) a zaštiťuje se vědeckou praxí (např. v neurovědách Kostić a Khalifa 2022, v lingvistice Zámečník 2023). Cílem příspěvku bude argumentace ve prospěch autonomie některých typů nekauzálních vysvětlení, především topologických vysvětlení (Kostić a Khalifa 2022). Bude zvažována slabá a silná varianta autonomie: (1) autonomie ve slabém smyslu znamená neredukovatelnost

nekauzálních konstitutivních prvků vědeckých explanací (Craver Povich 2017, Morrison 2015, Zámečník 2018), (2) autonomie v silném smyslu znamená neredukovatelnost (některých) nekauzálních vědeckých vysvětlení (Kostić a Khalifa 2022, Kostić 2023) Příspěvek ukáže, proč je mnohem snazší hájit autonomii ve slabém smyslu a jaké překážky brání upevnění pozice autonomie v silném smyslu.

Literatura

Craver, C., F., Povich, M. (2017) The Directionality of Distinctively Mathematical Explanations. *Studies in History and Philosophy of Science, Part A*, 63: 31–38.

Kostić, D. (2023) On the Role of Erotetic Constraints in Noncausal Explanations. *Philosophy of Science*, Volume 91, Issue 5, pp. 1078 – 1088.

Kostić, D. (2022) Topological Explanations, an Opinionated Appraisal. Lawler, I., Shech, E., Khalifa, K. (eds.): *Scientific Understanding and Representation: Mathematical Modeling in the Life and Physical Sciences*. Routledge: 96–115.

Kostić, D. (2018) Mechanistic and Topological Explanations: An Introduction. *Synthese* 195, 1: 1–10.

Kostić, D., Khalifa, K. (2022). Decoupling Topological Explanations from Mechanisms. *Philosophy of Science*, 22, 1–24.

Lange, M. (2017) *Because Without Cause: Non-Causal Explanations in Science and Mathematics*. Oxford: Oxford University Press, 2017.

Morrison, M. (2015) *Reconstructing Reality: Models, Mathematics, and Simulations*. Oxford: Oxford University Press.

Zámečník, L. (2023) *Investigations of Explanatory Strategies in Linguistics*. Berlin: De Gruyter Mouton.

Zámečník, L. (2018) Mathematical Models as Abstractions. *Organon F*, 25, 2: 244–264.

Dalibor Makovník

Rekonštrukcie vedeckých explanácií pomocou diferenciálnych rovníc

Účelom tejto prezentácie je ďalej prehĺbiť pochopenie (a rozdiely) medzi nekauzálnymi a kauzálnymi vysvetleniami. Pokiaľ ide o jeho povahu, nekauzálné vysvetlenie by sa mohlo definovať negatívne ako postrádajúce znaky, ktoré tvoria kauzálné vysvetlenie. To znamená, že ak faktory predstavujúce možné vysvetlenie nemajú bežné znaky kauzálného vzťahu, treba ich považovať za súčasť nekauzálného vysvetlenia. To znamená, že ak nejaké f pomáha vysvetliť g , potom g umožňuje aj vysvetlenie f . Pojmy symetria a asymetria, ktoré opisujú vzťah medzi zložkami, často s pomocou časového faktora, sa stávajú hlavnými pojmami pri určovaní, či sa vysvetlenie považuje za kauzálné alebo nekauzálné. Zastávam názor, že rozlišovanie

medzi kauzálnymi vysvetleniami a ich nekauzálnymi náprotivkami možno efektívnejšie zachytiť zavedením diferenciálnych rovníc (DR) do explanačnej praxe. Tieto matematické entity odrážajú nielen vzťah medzi funkciami a premennými, ale aj medzi deriváciami funkcií, čím poskytujú vysvetľovateľovi ďalšiu vrstvu reality, ktorú je potrebné zvážiť pri rekonštrukcii vedeckých vysvetlení. V tejto prezentácii ukážem, ako možno kauzálna vysvetlenia adekvátne zrekonštruovať poukázaním na rôzne vzťahy, do ktorých funkcie a ich derivácie vstupujú. Ako taký je zdôraznený rozdiel medzi kvantitami, ktoré zostávajú konštantné, a tými, ktoré sa môžu meniť. Navrhнем explanačný rámec, ktorý využíva diferenciálne rovnice na určenie správneho smeru vysvetľovania: s inferenčnou koncepciou je možné efektívne určiť adekvátnosť kauzálna vysvetlenia skúmaním, či funkcia a jej derivácia vyplývajú z premís argumentu alebo len samotná funkcia. Ak je prvý prípad pravdivý, je to znakom toho, že v explananse sú prítomné faktory prispievajúce k zmene fenoménu, ktorý je opísaný v explanande. Tento mechanizmus bude demonštrovaný na príklade stožiaru a tieňa. Tiež tvrdím, že prostredníctvom môjho explanačného rámca sa na nekauzálna vysvetlenie nemožno pozeráť iba ako na negatívnu entitu, ktorá nemá atribúty kauzálna vysvetlenia, ale funguje tak, že identifikuje tie premenné, ktoré predstavujú takzvané kritické body, ktoré zase odhaľujú body rovnováhy. Ukážem, ako sa táto optimálna hodnota dosiahne bez ohľadu na počiatočnú hodnotu nezávislých premenných – všetky sa približujú k tomuto stavu, čím ignorujú základnú náhodnosť počiatočných podmienok. Kauzálna-dynamické pozadie sa teda ignoruje, čím sa stanovuje nekauzálna charakter vysvetlenia. Pomocou tohto prístupu ukážem rekonštrukciu nekauzálna vysvetlenia stálej populácie lienok a vošiek, pričom sa budem spoliehať na model Lotka-Volterra, ktorý môže údajne opísať interakciu korisť-predátor v akomkoľvek biologickom prostredí.

Tomáš Kollárik

Groarkeho deduktivizmus a úskalia kruhového uvažovania

Táto práca kriticky hodnotí deduktivistický prístup Lea Groarkeho k rekonštrukcii a hodnoteniu argumentov. Hoci Groarkeho deduktivizmus ponúka analytickú jasnosť a metodologickú jednoduchosť tým, že robí argumenty deduktívne platnými prostredníctvom doplnenia implicitných premís, naráža na problémy pri významnom množstve nededuktívnych typov argumentov. Tvrdím, že rekonštrukcia induktívnych zovšeobecnení, argumentov autoritou a svedkom v tomto rámci vedie ku kruhovitosti alebo dokonca aj k redundancii niektorých kľúčových premís daných typov argumentov. Tieto problémy odhaľujú nielen technickú chybu, ale aj hlbší problém s logickou rozmanitosťou uvažovania v reálnom svete. Práca na záver vymedzuje úzku oblasť, v ktorej môže byť Groarkeho deduktivizmus životaschopný – ako napríklad pri niektorých kauzálnych alebo entymematických argumentoch – a naliehavo vyzýva k venovaniu vážnej pozornosti logickej heterogenite pri hodnotení argumentov.

Literatúra

Groarke, Leo. 1992. *In Defense of Deductivism: Replying to Govier*. Amsterdam: International Society for the Study of Argumentation.

Groarke, Leo. 1999. Deductivism Within Pragma-Dialectics. *Argumentation* 13(1):1–16.

Groarke, Leo. 2002. Johnson on the Metaphysics of Argument. *Argumentation* 16(3):277–86.

Groarke, Leo. 2026. The Deductivist Approach to Argument Evaluation. In *The Routledge Handbook of Argumentation Theory*, ed. Scott Aikin, John Casey & Katharina Stevens. New York and London: Routledge.

Liliana Bokroš

Kompatibilita odvodu najlepšieho vysvetlenia a bayesiánskej teórie potvrdenia – kritická analýza

Odkedy van Fraassen v roku 1980 v úplne odmietol odvodenie najlepšieho vysvetlenia (z angl. Inference to the Best Explanation, ďalej IBE) v prospech bayesianizmu, rozmohla sa vo filozofii vedy diskusia ohľadom ich kompatibility. Kritika van Fraassena primárne spočíva vo formuláciách rôznych kompatibilistických koncepcií IBE a bayesianizmu. Napr. Psillos (2004) a Weisberg (2009) uviedli, že takéto spojenie by vyžadovalo radikálnu transformáciu bayesiánskeho rámca – jeho objektivizáciu. Iní, napr. Okasha (2000) či Lipton (2004), nachádzajú možnosti ich kompatibility hneď na niekoľkých úrovniach. Okasha navrhol zahrnúť explanačné úvahy IBE do nastavenia východiskových a prediktívnych pravdepodobností, ktoré vstupujú do Bayesovej teóremy. Podobne podľa Liptona by explanačné úvahy mohli v rámci bayesianizmu pomôcť pri určení: i) východiskovej (priors) a ii) podmienenej pravdepodobnosti (likelihood); iii) relevantnej evidencie. V práci síce kategoricky nepopieram možnosť ich kompatibilného vzťahu, predpokladám však, že na to, aby sme mohli uvažovať o modeli založenom na ich spolupráci, je potrebné vyjasniť niekoľko problematických kategórií oboch modelov. Zastávam totiž názor, že kompatibilistická pozícia, ktorá považuje explanačné úvahy za určujúci faktor východiskových pravdepodobností v rámci Bayesovej teóremy, je problematická, pokiaľ nie je jasne charakterizovaný epistemický status evidencie. Hlavným dôvodom môjho stanoviska je fakt, že to, čo sa považuje za najlepšie vysvetlenie, je určené konkrétnymi prípadmi evidencie, ktoré už boli získané. V dôsledku toho nie je možné priradiť hypotéze pravdepodobnosť $Pr(H)$ bez toho, aby sme zohľadnili, už dostupnú, evidenciu (nie sme odkázaní len na doterajšie presvedčenia). Navyše by pravdepodobnosť $Pr(E)$ nadobudla hodnotu 1, čo by tiež pre bayesianizmus nepredstavovalo veľmi priaznivé okolnosti. Na základe týchto zistení dochádzam k záveru, že IBE a bayesianizmus operujú s odlišným pojmom „evidencie“, čo by malo byť zohľadnené pred akceptovaním kompatibilistickej pozície.

Literatúra

DELLSÉN, F. (2018): The heuristic conception of inference to the best explanation. *Philosophical Studies*, 175 (7), 1745 – 1766.

LIPTON, P. (2004): *Inference to the Best Explanation*. 2nd edition. London: Routledge.

OKASHA, S. (2000): Van Fraassen's Critique of Inference to the Best Explanation. *Studies in History and Philosophy of Science*, 31, (4), 691 – 710.

PSILLOS, S. (2004): Inference to the Best Explanation and Bayesianism: Comments on Ilkka Niiniluoto's „Truth-seeking by Abduction“. In: STADLER, E. (ed.) (2004): *Induction and Deduction in Sciences*. Netherlands: Kluwer Academic Publishers, 83 – 91.

WEISBERG, J. (2009): Locating IBE in the Bayesian Framework. *Synthese*, 167 (1), 125 – 143.

VAN FRAASSEN, B. C. (1980): *The scientific image*. Clarendon press.

Petr Jošt

Proti spekulativní bioetice

Spekulativní bioetika bývá chápána jako snaha o proaktivní řešení v praxi dosud neexistujících etických dilemat, která nicméně mohou vzniknout v důsledku biomedicínského pokroku. V současné bioetice se tak dostávají do popředí úvahy o možnostech ektogeneze, mind-uploadingu či biomedicínského vylepšování, a to navzdory tomu, že jen málo technologií podněcujících tyto debaty dospělo do stadia, v němž by skutečně vyvolávaly tolik diskutovaná etická dilemata. Bioetikové přesto zdůrazňují potřebu anticipovat budoucí technologický vývoj a jeho možné dopady na společnost s cílem zajistit včasnou regulaci a osvětu veřejnosti. Byť považuji ambice spekulativní bioetiky za legitimní, ve svém příspěvku se pokusím ukázat, že její úsilí představuje v mnoha ohledech sisyfovské snažení, jež navzdory dobrým úmyslům selhává v teoretické i praktické rovině. V první části se zaměřím na pochybnou metodologii, která vychází z nepodloženého přesvědčení o údajné možnosti spolehlivé predikce a následné kontroly a regulace biotechnologií. Má kritika bude vycházet ze dvou směrů. *Za první* poukážu na řadu příkladů z nedávné minulosti a současnosti, které ukazují, že predikce v bioetice jen zřídka odpovídaly následnému vývoji. Zvláštní pozornost bude věnována programu „ELSI“ (*Ethical, Legal and Social Implications*). Ten představoval od počátku 90. let minulého století bezprecedentní zdroj financování bioetického výzkumu zaměřeného na očekávané důsledky vyplývající z Projektu lidského genomu (*Human Genome Project*). A *za druhé* představím principiální námitku v podobě tzv. Collingridgeova dilematu, podle něhož je případná proaktivní regulace možná jen v raných fázích vývoje, tedy v době, kdy její potřebu nelze předvídat. Spekulativní bioetika nicméně není jen neefektivní, ale také potenciálně škodlivá pro status aplikované etiky. Ve druhé části se proto zaměřím na některé její neblahé důsledky, jakými jsou především upozadění skutečných aktuálních problémů či stigmatizace technologií, které bývají filozofy často značně dezinterpretovány a zveličovány ve vztahu k jejich skutečným možnostem. V závěru příspěvku předložím alternativní možnosti, jak v rámci bioetiky pracovat s nastupujícími technologiemi, a vymezím specifické situace, v nichž může být predikce opodstatněná.

Adam Greif

Morálka ako zdieľaná identita

Navrhujem metaetickú teóriu morálky – takú, ktorá sa snaží vysvetliť podstatu morálky a základ morálnych dôvodov. Teória je druhom antirealistického kognitivismu a jej cieľom je vykresliť morálku ako kategorickú, jednotnú a naturalistickú. Opieram sa o niekoľko bežných metaetických predpokladov, vrátane fundamentalizmu dôvodov, morálneho racionalizmu a internalizmu dôvodov, a ukazujem, ako táto teória rieši centrálny problém internalizmu dôvodov: ako ustanoviť vnútorné, no zároveň kategorické morálne dôvody. Zavádzam pojem psychologické identity a predkladám kľúčovú domnienku – že správne fungujúce psychologické ja je bezhraničné a zahŕňa univerzálnu identifikáciu. V dôsledku toho je v rámci tejto teórie zmysluplné tvrdiť: „Ja som ty“ a „Ty si ja.“ Proti-intuitívnu povahu tejto domnienky a týchto tvrdení riešim rozlíšením medzi ja a egom. Zatiaľ čo ego má hranice, ja – ako predpokladám – žiadne nemá. Predkladám tiež niekoľko dôvodov v prospech tejto domnienky, najmä na základe bežných morálnych praktík a diskurzu, ako napríklad Zlaté pravidlo. Na tomto základe rozvíjam kľúčový pojem zdieľanej identity a ukazujem, že môže slúžiť ako základ pre vnútorné, no kategorické morálne dôvody, a zároveň pre celý morálny systém. V dôsledku toho táto teória umožňuje pravdivo a zmysluplne tvrdiť: „Tým, že ti pomáham, pomáham sebe“ a „Tým, že ti ubližujem, ubližujem sebe.“ Hoci je teória založená na hypotéze, hlavným dôvodom, prečo ju brať vážne, je jej explanačná sila. Ak ju prijmem, nielenže sa vyrieši centrálny problém internalizmu dôvodov, ale objaví sa aj vierohodný morálny systém. Vierohodným morálnym systémom mám na mysli taký, ktorý umožňuje hovoriť o morálnej pravde a morálnom poznaní, ktorý ustanovuje jednotnú, naturalistickú a kategorickú morálku, a pritom sa vyhýba tradičným problémom konkurenčných metaetických teórií.

Stanislav Spodniak

Máme egoistické dôvody konať morálne? K podmienkam racionálneho sebaobetovania

Teória racionálneho egoizmu predpokladá, že nutnou a postačujúcou podmienkou racionálneho konania je, aby prispelo k blahu samotného aktéra. Niektorí autori (napr. Brink 1990 a Hocutt 2008) sa zaoberajú otázkou, či racionálny egoista môže vykonávať prosociálne, altruistické, ba dokonca sebaobetavé činy bez toho, aby sa spreneveril princípom racionálneho egoizmu. Neusilujú sa pritom o obhajobu samotného racionálneho egoizmu, ale o obhajobu morálky prostredníctvom argumentu, že aj egoisticky orientovaný aktér môže mať subjektívne (agent-relative) dôvody konať morálne. Tieto dôvody vychádzajú z predpokladu, že naše vlastné blaho do určitej miery závisí od blaha iných osôb, ktoré sú nám blízke. V našom záujme preto môže byť konať v ich prospech. Tézou tohto príspevku je, že skutočnosť, že aktér koná v prospech osoby, od ktorej do určitej miery závisí jeho vlastné blaho, je síce nutnou, ale nie postačujúcou podmienkou racionálneho konania v zmysle racionálneho egoizmu, pretože v niektorých prípadoch by racionálny egoista takéto konanie nemohol považovať za racionálne. Ide najmä o niektoré prípady tzv. patologického altruizmu, teda o situácie, v ktorých aktér koná v prospech blízkej osoby, avšak takým spôsobom, ktorý prekračuje zdravú mieru, a to aj napriek dobrým úmyslom. Kde presne leží hranica medzi rozumným a nadmerným

altruizmom, zostáva predmetom diskusie. Bez jasného vymedzenia nielen nutných, ale aj postačujúcich podmienok racionálneho konania v prospech druhých však nemožno jednoznačne určiť, či konkrétne konanie možno považovať za racionálne z hľadiska racionálneho egoizmu, a teda ani to, či má egoista dôvod konať morálne.

Literatúra

Brink, D. O. (1990): Rational egoism, self, and others. In: Flanagan, O. – Rorty, A. O. (eds.): *Identity, character, and morality: Essays in moral psychology*. Cambridge – London, MIT Press. 339 – 378.

Hocutt, M. (2008): In defense of Herbert Spencer. *The Independent Review*, 12 (3), 433 – 445.

Tomáš Čana

O objektívnej istote inak

Príspevok sa venuje kategórii istoty. To znamená, že patrí medzi dlhý rad príspevkov, ktoré sa pokúšali – a pokúšajú – objasniť unikátnu skutočnosť, na ktorú narážame ako v bežnom živote, tak i na poli špeciálnych vied. O čo ide? Ide o to, že ako na ulici, na súde, vo vedeckom laboratóriu, tak i všade inde sa stretávame so situáciami, v ktorých rámci nám o niečom nedáva zmysel pochybovať. Presnejšie, narážame na presvedčenia, o ktorých pochybovať by na nás pôsobilo komicky. Obvykle o takých presvedčeniach hovoríme, že sú *isté*. (Niekedy sa o nich tiež hovorí, že sú *evidentné, samozrejmé, nespochybniteľné, zreteľné, pevné* atď.) Ako príklad možno uviesť to, že som si istý, že *mám dve ruky*, že *som mal dve ruky aj včera*, že *svet existoval aj včera*, že *existoval aj pred sto rokmi*, že *každý má svoje meno*, že *ak je pravda*, že *„sme v podniku, kde ponúkajú pivo alebo kofolu“ a „pivo nechce“, tak je pravda*, že *„chce kofolu“, že mačky nerastú na stromoch* atď. Sú to presvedčenia, pre ktoré je charakteristické, že ich vlastne ani nepotrebujeme vyslovovať: za normálnych okolností nám o nich totiž ani nenapadne uvažovať, nie ešte pochybovať. (Samozrejme, s výnimkou malých detí v období, keď sa oboznamujú s výrazmi materinského jazyka: tie sa dokážu aj na presvedčenia tohto druhu so všetkou vážnosťou opýtať.) Hovoríme tu teda o presvedčeniach, pre ktoré platí, že ich vnímame ako pravdivé a zároveň ich pokladáme za akoby „imúnne voči kritike“. Z toho vyplýva, že okrem pravdivých presvedčení, o ktorých nemáme problém pochybovať, narážame v živote aj na také, pri ktorých sa niečo také zdá absurdné (nie nemožné, no akoby neprirodzené). Sú to *pravdivé presvedčenia, o ktorých nepochybujeme*. Jednoducho to (aspoň ako dospelí) nerobíme. Práve tieto presvedčenia konštituujú unikátnu skutočnosť, na ktorú sa pokúsim systematicky v tomto príspevku zamerať našu pozornosť. Napriek tomu, že sa tento príspevok venuje kategórii istoty, jeho primárnym cieľom nie je predložiť nový prístup k presvedčeniam, o ktorých pochybovať by na nás pôsobilo komicky. Vlastný prínos by nemal spočívať v presadzovaní originálnej pozície v tejto súvislosti. Cieľom je niečo iné: navrhnúť novú terminológiu. Predstaviť *novú optiku* a pokúsiť sa jej prostredníctvom urobiť pozície, ktoré doposiaľ boli v tomto ohľade sformulované, navzájom zrozumiteľnejšie. Prínos tohto príspevku by teda mal spočívať v tom, že sa pokúsi polemiky o kategórii istoty v modernom

filozofickom myslení prezentovať novým spôsobom: aby viac nepôsobili ako vzájomne obchádzanie. Práve tak táto výmena názorov totiž pôsobí. Nová terminológia má vo svojom centre dvojicu protikladných termínov: 1) *intelektualizmus o objektívnej istote*; 2) *antiintelektualizmus o objektívnej istote*. Ich prostredníctvom sa pokúsim odlíšiť dva typické pohľady na uvažovanú kategóriu. Predpokladom odlíšenia medzi 1) a 2) je však iná konceptuálna dištinkcia. Je ňou rozdiel medzi dvoma pojmi istoty: a) *subjektívnym*; b) *objektívnym*. Domnievam sa, že bez explicitného vyznačenia tejto dištinkcie sa na tomto poli nedá nikam dopracovať. Z toho vyplýva, že cieľom tohto príspevku je v prvom rade odlíšiť dve roly, v ktorých sa obrat „som si istý, že...” môže v komunikácii vyskytovať a v druhom rade čo najexaktnejšie vymedziť dve typické tendencie, ktoré sa *skrývajú* za dlhým radom konkrétnych filozofických pozícií zaoberajúcich sa objektívnou istotou.