



ČASOPIS
PRO
SOUČASNOU
FILOSOFII

ČÍSLO 2
ROČNÍK 11
2019

FILOSOFIE DNES

Časopis pro současnou filosofii
vydává KFSV FF UHK

Šéfredaktor: Ladislav Koreň

Výkonná redaktorka: Iva Svačinová

Redakční rada: Ondřej Beran, Jaroslav Daneš, Petr Glombíček, Tomáš Hirt, Juraj Hvorecký, Vojtěch Kolman, Ladislav Kvasz, Róbert Maco, Tomáš Marvan, Martin Paleček, Jaroslav Peregrin, Marie Hrdá, Ondřej Švec, Eugen Zeleňák

Editoři: Markéta Panoušková, Hynek Kaplan

Sazba a grafická úprava: Hynek Kaplan

Adresa: Filosofie dnes, KFSV FF UHK Rokitanského 62
500 03 Hradec Králové

E-mail: iva.svacinova@uhk.cz

vychází elektronicky,
dvakrát ročně

www.filosofiednes.ff.uhk.cz

ISSN 1804-0969

Ročník 11, Číslo 2 (2019)

Obsah

Studie

Jan Romportl	Naturalness and Artificiality Revisited Through Natural Infinity	4
Tomáš Šmirák	Reflexivita a propoziční postoje	34
David Kocourek	Collective Belief and the Intentional Strategy	57

Různé

Martin Paleček Matej Drobňák	10 let LMS centra – historie a současnost	76
---------------------------------	---	----

Naturalness and Artificiality Revisited Through Natural Infinity¹

Jan Romportl

*Katedra filozofie
Filozofická fakulta
Západočeská univerzita
Univerzitní 2732/8, 301 00 Plzeň
romportl@me.com*

Discussions about naturalness, artificiality and unnaturalness in this article are motivated by the field of Human Cognitive Enhancement (HCE) because of its potential for altering human personality and identity. This article at first proposes a concept of human naturalness as interaction between physis and logos. Then it presents an intuitive understanding of naturalness in terms of the inherent inability of language to fully describe all attributes of an object that is natural. The analytical core of the article proposes a formal model of naturalness utilising Vopenka's phenomenologically grounded Alternative Set Theory (AST), comprising and formalising the concept of natural infinity. A brief introduction to AST is presented as well. Naturalness and artificiality are modelled as two structurally different naturally infinite semisets within AST. This key structural difference is then analysed and applied back to the domain of HCE.

Key words: naturalness, artificiality, infinity, logos, physis, philosophy of language, human cognitive enhancement, alternative set theory, semiset, bioethics

1 Introduction

The contribution of this article to the theoretical analysis of the concepts of *natural*, *artificial* and *unnatural* is motivated by research in the field of Human Cognitive Enhancement (HCE). This article has been written

1 The research leading to these results has received funding from the Norwegian Financial Mechanism 2009–2014 and the Ministry of Education, Youth and Sports of the Czech Republic under Project Contract no. MSMT-28477/2014, Project no. 7F14236.

in the broader context of a transdisciplinary research project that aims at a complex assessment of the whole range of HCE-related issues, questions, and aspects, including an experimental analysis with neuroimaging and psycho-cognitive testing, health risks and benefits, public acceptance, and bioethical grounding. Especially the latter are very much influenced by discussions about *naturalness* – whether extending and enhancing human cognition by various technological interventions is natural or unnatural, and whether this is related to how good or bad and how acceptable or unacceptable such interventions are. However, in many of these discussions, it is very often left completely aside what the concept of *natural* actually refers to and what its underlying logical structure is.

HCE is a subarea of Human Enhancement (HE) addressing interventions in human cognitive capabilities and processes, such as those conducted as part of medical research (restorative/therapeutic) as well as those aiming at enhancing or replacing given biological human faculties², including also the domain of moral enhancement³. Cognitive enhancement can alter the human personality and identity much more profoundly than does corporeal enhancement, and therefore it requires more urgently a proper philosophical and ethical analysis.

HE technologies and systems, when coupled with a human and her body, strongly elicit the feeling of *intrus* ('the intruder'), as discussed in Nancy's seminal paper 'L'Intrus'⁴. The question of their acceptance is thus indeed of great philosophical importance, far overreaching ideological and biopolitical concerns. One of the key aspects of the acceptance of HE (individual, social, cultural) is *naturalness*, as can be seen e.g. from the work of Francis Fukuyama⁵, Ryuichi Ida⁶, or generally the whole HE overview book edited by Nick Bostrom and Julian Savulescu⁷. The advocates of enhancement either assert that HE systems are natural, or they that they might be artificial but that it is itself natural (or at least is not unnatural) to extend/enhance human natural body/cognition by these

2 Bostrom & Sandberg (2009), BMA (2007).

3 Savulescu & Maslen (2014).

4 Nancy (2002).

5 Fukuyama (2003).

6 Ida (2009).

7 Bostrom & Savulescu (2009).

systems. The bio-conservatives claim basically the opposite⁸. Therefore, the pro- and anti-enhancement feud can be at least partially transformed into a discussion about naturalness (in its more general sense), or even about the ‘nature of naturalness’. In other words: HE acceptance can be seen through the looking glass of naturalness.

Even though the concept of naturalness does not find its place very often in the current discussions within mainstream bioethics⁹, its persisting importance can be illustrated by the recent focused project, study, and report of the Nuffield Council on Bioethics¹⁰ that reviewed the usage of the terms *natural*, *unnatural* and *nature* in various discursive contexts including bioethics, commercial sector and public or parliamentary debates. However, even that study focuses only on one side of the whole problem: a textual analysis of how *natural* and *unnatural* are used discursively as placeholders for many different value-laden or value-neutral concepts and connotations. See, for example:

“Often when people describe something as *natural* or *unnatural* in the context of a debate about science, technology, or medicine they are using those terms to make claims about that thing being good or bad, or right or wrong.”¹¹

The Nuffield Council’s study does not provide an analysis of whether something like *naturalness* really exists *objectively*, perhaps as a predicate that assigns a certain immutable quality to all phenomena (objects, processes) that are natural, much in the same way as for example *red* assigns a quality of *redness* to all the objects that are red. Moreover, rather than explaining what the *logical* structure of naturalness is, that study explores various *ideological* structures of naturalness in various biopolitical discourses, which is also very important, but for other purposes than those that I want to pursue here.

8 Moreno (2011), Kass (2004), McKibben (2004).

9 Smajdor (2015).

10 NCB (2015).

11 NCB (2015, p. 14).

It is therefore the goal of the present article to supplement and extend studies such as that of the Nuffield Council report, with an analytical account of the concept of naturalness, which can perhaps help in understanding what is so specific about *naturalness* that makes it always so difficult to truly use it in a different way than just as a discursive placeholder. I will also try to make explicit distinction between *artificial* and *unnatural*, even though these two terms are apparently often used interchangeably¹² because of their similar negative values and connotations in many discourses.

As the author, I am aware that I am trying to transdisciplinarily blend such different fields as phenomenologically motivated non-standard set theory, bioethics, analytic philosophy and even somewhat speculative metaphysics as a source of intuition. And I am also aware that this mixture can indeed be hazardous and can upset experts from literally all these fields due to various technical, methodological, and even conceptual issues. However, I also honestly feel that many traditional correct approaches are somewhat stuck in an inability to provide guidance in acute and pressing questions and problems of rapid technological development that can profoundly change the human (one of which is, e.g. the here discussed HCE). So this article has been prepared with motivation driven from the mission of Epstein's *Transformative Humanities*¹³, which can perhaps help constructively soften some of its rough edges.

2 Building an Intuition

Before I attempt to build a formal set-theoretical model of naturalness, I will explain the intuition behind this proposal of how 'natural' relates to 'artificial', how 'artificial' is different from 'unnatural' and how 'naturalness' differs from *human naturalness*.

2.1 Human Naturalness

I would like to bring to attention two concepts from ancient Greek philosophy: *physis* and *logos*. It is true that in their classical philosophical

¹² NCB (2015).

¹³ Epstein (2012).

usage they usually do not form mutual opposites, but I believe that contrasting them can be useful for our purposes.

Physis is a concept referring to all things that grow on their own, intrinsically; to things that are in and created by Nature. It was used by the Greek god Hermes when he pulled out a plant to show Odysseus its intrinsic way of growth. It is something that all the discursive usages of the words *nature* and *naturalness* (as analysed by the Nuffield Council Report) have in common: it is their conceptual core. However, we can intuitively feel that it is very difficult to truly grasp what *physis* really is. All the uses of *nature* and *naturalness* in all the contexts and pragmatic domains refer to it, but it is as if the substance of *physis* always slipped away. We will see in the next sections why this is so, but for now let us intuitively understand *physis* as the core of *naturalness*, as the intrinsic way of growth shown by Hermes.

Logos, on the contrary, means the whole complex and metaphorical concept of speech, meaning, human reason, rationality. Ever since the philosophy of Heraclitus, it has been used to denote the principle of order and knowledge. Just as *physis* is the core of all the natural things that grow by themselves without any human intervention, *logos* is the core of the logic behind all rational arguments that inherently pertain to the set of aspects delimiting the human from the rest of Nature.

The Nuffield Council Report shows how often in public, media and policymaking discourse the naturalness-related words are value-laden, and especially with positive values. From there, the public tends to resort to the *naturalistic fallacy*¹⁴. The concepts of *physis* and *logos*, on the other hand, should be value neutral, or better said, they connote such a diverse set of opposing values that I would call them rather value ambivalent. *Physis* connotes with raw untamed things, growth, procreation, reactivity, wilderness, warmth, dynamics, spontaneity, vortex, chaos, vagueness, mud, dust, rot, worms, turmoil, and also excessive growth (as in cancer). *Logos* connotes clarity, purity, intentionality, geometry, logic, cold, static, copying, algorithm, and also excessive loss (as in Alzheimer's disease).

We can intuitively construct a continuous spectrum representing the scale of the degree to which we judge a given object or process as pertai-

¹⁴ Smajdor (2015).

ning to *physis* or to *logos*, having abstract pure *physis* on one extreme side of the spectrum and abstract pure *logos* on the opposite extreme side. The roots of an herb happily and undisturbedly growing through the dark of the soil would be very close to the *physis* end of the spectrum, while the roots of a polynomial function are just at the opposite end, very close to *logos*.

I then assert that we should put human at the middle of this spectrum. When we define ‘human’, we automatically delimit this concept from ‘lesser’ animals by an inherent interconnection with rational and abstract thinking, hence as a *logos*-creature, but at the same moment we stress the human’s strong embedding in a ‘natural’ biological origin in order to delimit the human from all the artificial entities and interventions triggered by technological development (as often appearing in public discoursed analysed e.g. by Smajdor or Nuffield Council Report¹⁵, hence as a *physis*-creature. Metaphorically speaking, human is a being of the tension between *physis* and *logos*, a being producing and produced by this tension, a being that possesses about the same from both realms. Therefore, ‘being as natural as human’ means ‘having the same ability to balance in the equilibrium between *physis* and *logos*’. So we would expect also HCE systems to respect this kind of *human naturalness*: neither dragging human beings too much towards *logos* nor towards *physis*.

Every thing, every object, has its share of the artificial and the natural. There is no purely natural object because the objectness itself is the first trace of artificialisation. Understanding a fragment of reality as an object is a matter of the *logos* and thus it gives the first blow to its pure naturalness. The tools of *logos* pull out pieces of inherently non-structured *physis* and construct shapes and objects from them. When Hermes showed the herb, drawing it from the ground to demonstrate its nature (its *physis*, the *physis* was already retreating. It was still somehow very strongly there, much more strongly than if Hermes had shown a plastic bag or a microchip, but no longer in its pure form because artificiality had already crept in: the herb ‘being shown’ is not the herb ‘being natural’ inseparably in its *physis*.

¹⁵ Smajdor (2015), NCB (2015).

What Hermes did was something very typical for a human, or even delimiting for a human mind, thought, and intelligence. If an animal were to interact with an herb instead of a human, it would not show/objectify/name the herb – it would simply transparently share with the herb their ‘unspoiled’ and non-conceptualised *physis* together. On the other hand, if it were a machine instead of a human, it would operate only with purely symbolical representations completely detached from the intrinsic substance of the herb, and the herb itself would be replaced by a single symbol, or a symbolic representation of its geometrical model, or a symbolic representation of its molecular structure, or something similar. In other words, *logos* dissolves *physis*, and the human is, in its nature, a steersman constantly oscillating around this unstable equilibrium where objects appear from the mud of *physis* before they disappear into the void of *logos*. We can also see this metaphor as a keen variation of Wiener’s and Ashby’s¹⁶ cybernetics.¹⁷

Human naturalness is thus something significantly different from *naturalness* seen merely as *physis*. Human naturalness is an indivisible and intrinsic combination of the natural and the artificial, continuously re-enacted by the process of life itself. Therefore, should there be any bioethical account of naturalness in HCE, it must be grounded in the aforementioned *human naturalness*, because otherwise if it again resorts only to the purely *physis*-based understanding of *naturalness*, we will not be able to prevent it from drowning in the naturalistic fallacy.

2.2 Natural and Artificial Objects

We can say that a major tool for such a steering between *physis* and *logos* is language. Language in general is a long bridge between *physis* and *logos*, with deixis and protolanguages close to the bank of *physis*, formal languages, mathematics, geometry, etc., close to the bank of *logos*, and natural language somewhere in between, where human minds operate.

When we use language to further analyse a freshly objectified (shown) object in more detail, we go step by step over this bridge and we start

¹⁶ Ashby (1957).

¹⁷ As Prof. Cyril Höschl, director of Czech National Institute of Mental Health, informally noted during one of his lectures: “Life is a process of sailing between Scylla of cancer and Charybdis of dementia.”

losing more and more of the object's naturalness. The object in itself stays the same but we receive more and more detached abstract concepts. Although such concepts are new objects transferable by means of language, they lose their connections to the inherent givenness of the original object.

For example, no human being can describe in words how the root of a herb (or a cloud, a bird's nest, a coral) *exactly* is. The moment closest to the root's naturalness is when we *show* it (deixis), and after that the more we say about it in an attempt to fully describe it, the more artificial a construct we get. At some point, the length of the description reaches beyond the limit of any human being and becomes manageable only by symbol processing machines, having no meaning for a human while being in this *logos* domain. For example, a 'sentence' with 10 million 'words' might be quite a good description of how the root is, but only as long as we interpret it as a 10-megapixel photograph of the root, forgetting everything about the language and humbly returning back to *showing* the root or at least its image, i.e. back to a much more *physis*-related deixis. The *logos*-based interpretation of those 10 million symbols (i.e. reading and understanding them one by one) has absolutely no meaning for us.

So what is the *natural*? The natural is that which defies being captured by language. *Naturalness* – or *physis* – is everywhere where we feel tension between what we wanted to capture by our words and what we really captured. The more the tension, the more naturalness we just encountered. The natural is something that we have to abstract away from in order to capture it by language.

In contrast, the *artificial* is imposed by language: the artificial is that whose essence is fully determined by language. The artificial is a language-implemented abstraction drawn from the soil of *physis* and attracted by the clarity of *logos*.

Let's imagine an old rustic wooden table. What is artificial about it? That which we can grasp with words: the shape and size of its geometrical idealisation, its weight, colour tone, purpose, or perhaps a description of the way it was made by a carpenter with an axe, a saw, and a jack plane. However, we cannot describe how *exactly* it looks, how it feels

when being touched, the exact look of its texture and wood structure, its smell.

Now let's imagine a three-legged white round plastic garden table. How to grasp it with words? Just take its designer's drawings and the description of the technological aspects of its manufacture, and we have it right in front of us. We do not need to see and touch and feel this table to fully know *how* and *what* it really is. Hence, it is almost completely artificial. Yet even such an artificial thing has something natural about it: various scratches, defects, imperfections, shabbiness, but most importantly its inherent *qualia* potential that we exploit when we meet the table right here and now. All these aspects defy being captured by words, and therefore are natural.

It is important to mention in this place that terms such as 'natural variation' and 'natural sugars', or sentences such as 'hedgehogs are a part of the UK's natural wildlife' use *natural* in a different meaning than what we have discussed so far: they use *natural* in the most common and simple meaning, referring to those things that have not been subject to human intervention (which is a very exclusive meaning because it excludes all human activities from being natural) or originate in a source that was not created by humans. However, we can still get the question whether fully grown forests manually planted by humans hundreds of years ago are a part of 'natural wildlife'. Probably yes, but are they also natural? And this is the moment when we must get back to our definition of *natural* formulated in the paragraphs above.

2.3 The Artificial and the Unnatural

The term *artificial* should be perceived as relatively value neutral, whereas *unnatural* clearly carries a heavy load of negative values. Both are, however, used as the opposite of *natural*:

“Our work identified an asymmetry between use of the terms *natural* and *unnatural*. It found that, proportionately, there is a contrast between the regularity with which value is invoked by use of the term *natural* and the term *unnatu-*

ral [...] Within the sources reviewed, the term *natural* was used much more commonly than the term *unnatural* and was typically used in a value-neutral way. In contrast, when the term *unnatural* was used, it was often used to suggest something is wrong or bad.”¹⁸

From the textual examples that the Nuffield Council’s report presents (e.g. “The instinctive desire within many of us not to consume something that is unnatural – the fear of so-called Frankenfoods.”, or “I was also unhappy with the idea of having something as unnatural as a silicone implant in my body.”) it is apparent that almost always when the term *unnatural* is used, it means something not only wrong or bad, but also repulsive or abhorrent.

It seems that in many discourses, especially public and in the media, the term *artificial* is used in the same sense, i.e. losing its value-neutrality, e.g. in “Most of the 600 people working at Thanet Earth appear to believe passionately in what they do, and angrily reject the charge that they are somehow perverting nature or creating something artificial.” However, I will put this usage of *artificial* aside here because these cases simply look like common mistakes in word selection resulting from ignorance or semantic carelessness, because ‘creating something artificial’ is actually very natural for humans: it is an inherent part of *human naturalness*.

What I suggest here is to understand *unnatural* strictly as the opposite of *human naturalness*, not as the opposite of *natural* in the sense of *physis*. This will fit perfectly most of the discursive usages of *unnatural*. *Artificial* then can be the opposite of *natural* in the sense of *physis*.

‘Frankenfood’ is thus unnatural because it repulses *human naturalness*, no matter whether it is artificial or coming from *physis*. Analogically, bread or sushi or traditional English Christmas pudding are quite artificial in the way they are made, yet they are perfectly natural in the sense of *human naturalness*. And a human being having two heads (one of them perhaps necrotic) and three arms is very frightening, repulsive, and unnatural, even though it could quite possibly happen in an entirely natural way through the blind powers of *physis*.

¹⁸ NCB (2015, p. 19).

3 A Formal Model of Naturalness

The presented reflections show that in discussions about HCE, it is not necessary to give up on the concept of *naturalness* as being too inclusive (everything that appears in the natural world, including all the human artifacts, is natural), or too exclusive (only those things that have not been subject to human intervention are natural), or just a conversational placeholder for various value-laden judgements (as analysed by Nuffield Council Report¹⁹). *Naturalness* is not an empty concept – it is just a concept with a different internal structure than we are generally used to, and it is much more difficult to represent this structure and operate with it. So in this section, I will try to outline a way to formally grasp and anchor *naturalness* – both in the sense of *physis* and in the sense of *human naturalness*.

The question now is *how* the internal structure of *naturalness* is different from what we would have naively expected. I will analytically tackle this problem by using a set-theoretical approach for modelling the underlying structure of *naturalness*. However, the standard axiomatic set theory (Zermelo-Fraenkel set theory in the classical Cantorian universe of sets) does not offer suitable tools able to properly represent the phenomenal aspects of *naturalness*, into which we have intuitively gotten insight in the previous section. The necessary (yet standardly missing) formal tools can be summarised in a simplified way as *infinitesimal* (infinitely small, beyond any measurable size) elements or *vague* elements – and the standard set theory has none of them.

A well-known nonstandard set theory is Nelson's internal set theory²⁰ which builds on the non-standard analysis introduced by Robinson in the 1960's²¹. The problem with Nelson's and Robinson's approaches is that they, while being very well elaborated and axiomatised and while introducing infinitesimals, are built without any phenomenal intuition that would help us link them easily with the phenomenally perceived internal structure of *naturalness* (or any other concept used in natural language). However, there is another nonstandard set theory, introduced

19 NCB (2015).

20 Nelson (1977).

21 Robinson (1966).

by Vopenka, usually called Alternative Set Theory (AST) or the Theory of Semisets. It is very similar to Nelson's system in terms of its formal mathematical aspects (including similar mathematical advantages but also various technical difficulties), but unlike Nelson, Vopenka has built the whole system of AST on very profound phenomenal intuitions, making it much more accessible outside the scope of pure mathematics.

Unfortunately, Vopenka's AST faces another problem: for various historical and political reasons, the bulk of Vopenka's work has been published mostly in Czech, without a proper English translation (Vopenka was persecuted by the communist regime in Czechoslovakia, banning him from publishing internationally). Only an early version of AST was published, in 1979, in an English monograph²², i.e. in the period of high interest in nonstandard analysis among mathematicians, but all the subsequent developments were published mostly in Czech, Slovak or Russian²³, after the mathematical mainstream had taken a different way. This makes it much harder now to build and publish anything on top of Vopenka's AST framework because he has been internationally well known only in a very specialised group of theoretical mathematicians (e.g. Holmes²⁴, or Kanamori²⁵ elaborating Vopenka's earlier works on large cardinals and Vopenka's principle). Only recently, shortly before he passed away in the spring of 2015, did Vopenka finish the final draft of an English monograph comprising a fully rewritten and updated version of AST. This had been in its first version available as a preprint²⁶, and now the final edited version is pending publication²⁷. So I will try to very briefly introduce, for the non-mathematician reader of the present article, some of the most important principles of Vopenka's AST (even though this makes the article significantly longer). Then I will use these principles to model the concept of *naturalness*.

22 Vopěnka (1979).

23 Vopěnka (1989, 1996, 2001, 2014, 2015).

24 Holmes (2012).

25 Kanamori (2009).

26 Vopěnka (2012).

27 Vopěnka (2019).

3.1 Horizon, Natural Infinity, and Semisets

The key concepts in AST are built around what Vopenka calls *natural infinity*. Unlike *actual* or *potential* infinity, natural infinity represents such an infinite quantity that can be found in the real natural world and can be phenomenally processed by a human-like observer. Classical Greek geometry operated with *potential* infinity where a geometric line is always finite but can be potentially prolonged anytime by any finite length, hence being infinite in potentiality. This infinite prolongation was given to the powers of Greek gods such as Zeus, so it got out of reach by any human; but not even Zeus was able to oversee the whole line in its actual infinity after all the (infinite number of) potential prolongations had been expanded. Such an omnipotent sight, encompassing all the actualised infinities at once, has been given to the Christian God, and in the form of the *actual* infinity has been firmly installed in the classical Cantorian set theory, where all the possible crystal-clear finite and infinite sets and all their powersets²⁸ exist at once. It is clear that most of the abstract entities existing in Cantor's set universe do not have any counterparts in the whole physical universe, let alone in its phenomenally conceivable part.

Natural infinity, on the contrary, was defined by Vopenka with a strong link to the concept of *horizon*:

“Every look²⁹ we cast, no matter in what direction, is limited. Either there is a firm boundary which disrupts (or deflects) it sharply, or it is limited by a *horizon* in whose direction clarity decreases and sharpness blunts.”³⁰

Thus, a naturally infinite abstract structure would be the class of natural numbers that number the steps needed for an observer (who casts the look) to take her directly to the horizon. In the ordering of those natural numbers, there is no such number about which we could say that it is the

²⁸ Powerset is the set of all subsets of the given set.

²⁹ A *look* here does not refer merely to a look by sight (with physical eyes); we understand it in the broad sense of regarding something that has been encountered. [The footnote is a part of the original quotation].

³⁰ Vopěnka (2019, p. 62).

last number reaching the horizon, i.e. if n belongs to the class, then $n + 1$ belongs to it too. In this sense, the class is infinite. However, we know that there is a finite (perhaps very big) number N about which we know that if we make that many steps, we will definitely get *past* the horizon. This means that naturally infinite classes are subclasses of large classical finite sets. These subclasses are called *semisets* and they are a crucial structure of AST that, apart from being naturally infinite, also model the phenomenal concept of *vagueness*.

We can see here that the horizon is a rather general epistemological concept, but it is also very often helpful, for gaining an initial intuition, to illustrate it on plain visual examples, such as a long straight railway track with railway sleepers marking the steps to the horizon and beyond.

The fact that we are constructing a set theory where infinite classes are subclasses of finite sets might seem very counterintuitive at first. This situation is, however, quite similar to the hard first step that had to be done on the way from Euclidean geometry to non-Euclidean geometries: to concede that two parallel lines can have a point of intersection.

There are many intuitive examples of real-life naturally infinite semisets (just as there are intuitive examples/models of non-Euclidean geometry, e.g. meridians on the globe that are all parallel and still have two intersections at the poles). The whole class of *sorites* paradoxes is one of them: a naturally infinite semiset represents the class of sand grains that have to be removed from a sand heap before it is not a heap anymore (or it can be the ‘bald man paradox’ if we want to speak about hair instead of sand). The same actually can hold for the extensions of most natural language predicates, such as ‘tall’, ‘old’, ‘pretty’, ‘smart’, ‘green’, etc. In the case of colour gradients (e.g. continuous red-to-green transitions), semisets are the classes of all the adjacent colour tints that we perceive as a single colour, e.g. ‘green’. We know that if a particular given tint is green, its immediate neighbour will also be green, but if we skip over enough tints we will reach something that we quite clearly identify as red.

Semisets also represent the classes of all those colour tints that are mutually *indistinguishable*. However, as I will show later, the internal structure of these *indistinguishability* semisets is inherently different

from the structure of semisets that represent the aforementioned predicates. And it is exactly this structural difference that pertains to *naturalness* as a discursive placeholder analysed by the Nuffield Council Report (its structure is similar to the structure of predicates like ‘green’ or ‘tall’) versus *naturalness* as *physis* (its structure is similar to the structure of *indistinguishability*).

Years ago, Vopenka gave another good example³¹ (originally meant as rather humorous and relaxed) of a natural infinity, which now has a strong link to the topic of HCE: let’s imagine a set of individuals starting with some prehistorical primate (let’s call her ‘the monkey’ for the sake of simplicity, even though this will quite likely upset many biologists, and not just them), then going through all the direct descendants of the monkey up to some human actually living these days (for example, me). This set is clearly a finite set. A naturally infinite semiset is a subclass of this set that corresponds to all the individuals from that direct hereditary line who are monkeys, because we know that if an individual is a monkey, her direct offspring is also a monkey; but we also know that there are some individuals further in the same hereditary line who definitely are not monkeys anymore.

This very well illustrates the close connection between *vagueness* (as inherently present in the natural language predicates) and *infinitesimals*. The aforementioned monkey-human sequence can actually be also understood in such a way that a direct offspring of each monkey individual is infinitesimally less monkey than this individual, and vice versa, a parent of each human individual is infinitesimally less human than this individual. The infinitesimality here means that the change is so small that it is beyond any measurable means within a given system.

When we speak about the set-theoretical universe itself, the infinitesimal entities are smaller than any real number (where a real number is the measure), which is exactly the case in the systems of Robinson, Nelson, and also Vopenka. But outside of the strictly formal set-theoretical apparatus, when applied to the real world, we can see that infinitesimal features can be found even in the objects normally perceivable by

31 This example is from (Vopěnka 2001) but even there Vopenka stated that he had originally presented it ‘long time ago’.

the bodily senses – just like the sand grain in the case of the heap of sand is infinitesimal with respect to the heap. Infinitesimality is an epistemological aspect of the *look* being cast at a fragment of reality. And this is where the phenomenological intuition behind Vopenka’s formal system is very useful. Moreover, it immediately draws our attention to the fact that even certain HCE interventions may exhibit exactly this kind of infinitesimality with respect to human naturalness.

3.2 Two Internal Structures of Semisets

It is unavoidable to use at least some very basic mathematical notation in this section, but I will limit it to the minimum (comprehensible to anyone with some basic high-school knowledge of mathematics) and also the explanation will be illustrative rather than axiomatic. Note on terminology: the term *set* in AST is used strictly for a finite structure (the same as the finite sets in standard set theory), the term *semiset* is used for a naturally infinite structure (this is specific to AST; it is not present in standard set theory), and the term *class* can be used both for sets and semisets, i.e. all sets are also classes, but not all classes are sets because some of them can be semisets. This also means that a semiset is a naturally infinite subclass of a finite set.

We start with a less formal and more intuitive definition of *function*. A function F is a structure that takes an input value x and on the basis of a fixed mapping returns exactly one y for this given x , formally written as $F(x) = y$. Both x and y can generally be any abstract entity modelled within the set universe, such as numbers (natural or real), sets, sets of sets, sets of numbers, etc. However, in this article, it will be enough for us to suppose that our functions take only numbers and return only sets of numbers. For example, consider the function that, for the input number 4, returns the set of three numbers 2, 6, 9, i.e. in the formal notation $F(4) = \{2, 6, 9\}$, and for the input number 5 it returns the set $\{2, 6, 8, 10\}$, i.e. $F(5) = \{2, 6, 8, 10\}$. And by coincidence, when our function takes the number 6, it returns the same output as for the number 4, i.e. $F(6) = F(4) = \{2, 6, 9\}$. Our function accepts only these three input numbers, it is not defined for any other input number. An important point is that

a function itself is a set within the set universe: it is a set containing ordered couples $\langle y, x \rangle$.

The set of all those x for which the function F is defined is called its domain. The notation for the domain of F is $dom(F)$. In our case, $dom(F) = \{4, 5, 6\}$. The set of all possible y that the given function can return is called its range, and its notation is $rng(F)$. In our case, $rng(F) = \{\{2, 6, 9\}, \{2, 6, 8, 10\}\}$. Note that the range of this F is a set of two sets, one containing three numbers, the other containing four numbers.

Every reader is probably familiar with the operations of set union (denoted \cup) and set intersection (denoted \cap). So if we have a set A that contains any number of sets as its elements, then the set $\cup A$ is constructed as the union of all the sets that A contains. Analogously, $\cap A$ is constructed as the intersection of all the sets that A contains. In case of our example F , we have $\cup rng(F) = \{2, 6, 8, 9, 10\}$ and $\cap rng(F) = \{2, 6\}$. We can see that the sets constructed by \cup can be ‘growing’ if we add more sets to A , whereas the sets constructed by \cap can be only ‘shrinking’.

The *segment* of the natural numbers determined by n is the set of natural numbers $\{0, 1, 2, \dots, n-1\}$, i.e. the set of n consecutive natural numbers starting with 0. On this basis, AST introduces a new structure called a *horizon segment*, which is such a segment H where two conditions must be met: 1) in the natural number ordering, H has no last element m , i.e. it is infinite; 2) any function F defined on H (i.e. $dom(F) = H$) can be prolonged to a finite set. These two conditions, their justification and their consequences are discussed by Vopenka in detail³² but these discussions are far beyond the scope of this article. For us it is important to know that a horizon segment H is a basic semiset and a basic set-theoretical model of natural infinity.

Now we are getting close to grasping two important internal structures of semisets. Let X be a class, either a set or a semiset – it does not have to be a segment, it can simply be any finite set or any naturally infinite semiset. And let H be a horizon segment.

If it is possible to find a function F defined on H (i.e. $dom(F) = H$) such that $X = \cup rng(F)$, we call X a σ -class. And if it is possible to find a func-

³² Vopěnka (2001, 2015, 2019).

tion F defined on H such that $X = \cap \text{rng}(F)$, we call X a π -class. From the meaning of the operators used in them, we can say that σ -classes have the ‘growing’ quality in them while π -classes have the ‘shrinking’ quality, just as, a couple of paragraphs above, we discussed in the example with the simple finite function, only now the constructing function F (i.e. the function that ‘constructs’ X) is defined on a naturally infinite semiset, not on a finite set -- which makes it far more interesting.

Vopenka has constructed a lot of formal propositions and proofs about σ - and π -classes and has built a rich theoretical apparatus around them³³. At this moment, we will focus only on two of those propositions, which assert the following (rather simplified here):

1. A class X is both a σ - and π -class at the same time if and only if X is a set. This means that if X is a set (i.e. a finite structure, just as in standard set theory), we can find both ways of constructing it from different functions defined on a horizon segment H : the ‘growing’ way \cup as well as the ‘shrinking’ way \cap . However, if X is a semiset (i.e. a naturally infinite structure), it can only be either a σ -class, or a π -class, not both at once. It means that there are two different kinds of semisets: those that we can construct only in the ‘growing’ way \cup , and those that we can construct only in the ‘shrinking’ way \cap – based on their internal structure.
2. If X is a σ -class and w is a set, then $w - X$, i.e. the complement of X with respect to w , is a π -class. And the other way around, if X is a π -class, then $w - X$ is a σ -class.

As we have already mentioned, predicates within natural language are – due to their inherent vagueness and lack of absolutely sharp boundaries – very well modelled by semisets. Many of these predicates have the structure of a σ -class, more specifically, those predicates standing for phenomena that Vopenka calls *primarily evident phenomena*:

³³ Vopěnka (1989, 2001, 2015, 2019).

“A primarily evident phenomenon is a phenomenon which we are able to evidence, that is, we are able to see it and know that we are seeing it as soon as it can be evidenced, that is, as soon as it has appeared to us.

For example, if we look at this page of this book, we will probably agree that it is not red. In the sensorily perceptible world [...], interpreted in the usual way as a community of objects, we interpret non-redness as a unary attendant phenomenon. [...] The phenomenon we are discussing is before our eyes on this page. But to be able to evidence it, we must first know redness from somewhere: in other words, we must have evidenced redness at an earlier date. Non-redness is thus not a primarily evident phenomenon. By contrast, redness is a primarily evident phenomenon because even if we had not come across this phenomenon before, we would notice the red colour as soon as we first saw it.”³⁴

It is exactly this distinction between primarily evident and non-primarily evident phenomena that can be captured as a distinction between σ -classes and π -classes, where σ -classes are formal models of primarily evident phenomena and π -classes are formal models of non-primarily evident phenomena. The predicate ‘red’ is a σ -class which can be formally constructed for example by a function F_{red} that returns a set of red objects (or rather their abstract placeholders, e.g. numbers) for each time point when somebody (e.g. me, or the whole of mankind, it does not matter in this example) sees some objects, and the class of these time points is a horizon segment H_{time} on which F is defined.

In case of the complementary predicate ‘non-red’, one might think that we can formally construct it the same way as ‘red’, i.e. analogously by applying the union operator \cup over sets of non-red objects given by the function F_{nonred} defined on the horizon segment H_{time} (and the standard set theory that does not support semisets would actually lead us to this conclusion). However, the aforementioned proposition (2) tells

³⁴ Vopěnka (2019, p. 89).

us that this is not possible. The extension of ‘red’ clearly is not a set but a semiset, so following from (1), ‘red’ is only a σ -class, not a π -class at the same time, hence following from (2), ‘non-red’ as the complement of ‘red’ is only a π -class, not a σ -class at the same time, and therefore there is no way to find a function that would construct ‘non-red’ by the union operator \cup in the ‘growing’ way. The only way how to construct ‘non-red’ is by the ‘shrinking’ intersection operator \cap . And this is indeed supported by our intuition: from observing only and only ‘non-red’, there is absolutely no way how to understand (and phenomenally process) what ‘red’ is – we even would not know that we observe ‘non-red’.

This way of constructing the π -class ‘non-red’ places very different demands (in terms of their semantics) on what entities the constructing function F_{nonred} can return. While the constructing function F_{red} for the σ -class ‘red’ is doing fine by returning sets of red objects (their abstract placeholders) because the final extension of the predicate ‘red’ is then the union \cup of all these sets of objects, the constructing function F_{nonred} for the π -class ‘non-red’ cannot analogously return plain observed non-red objects because the operator of intersection applied to them would soon produce the empty set \emptyset , as no non-red *object* would appear in absolutely all (naturally infinite number of) sets of non-red object observations given by F_{nonred} on H_{time} . Instead, the function F_{nonred} must return, for example, very big finite sets of *attributes* (i.e. actually other predicates) of the observed non-red objects, not the objects themselves, because after applying \cap to all these naturally infinite number of sets of attributes, we would get the class of attributes that all the non-red objects share: the π -class of the ‘essence’ of ‘non-red’.

This is a very important difference between the internal structures of the primarily evident phenomena such as ‘red’ and non-primarily evident phenomena such as ‘non-red’. From the difference in demands placed on the constructing functions of primarily and non-primarily evident phenomena, we can somewhat eloquently conclude that the underlying structure of primarily evident phenomena (and their respective predicates) is the aggregate of *what* they are, whereas for non-primarily phenomena, it is the essence of *how* they are.

It is not easy at first to intuitively fully comprehend the somewhat strange structure of a π -class. So for further illustration, let's imagine the situation that I have a piece of paper from which I start to tear off little pieces one by one so that the remaining piece of paper is smaller and smaller after each step. I can do this until the remaining piece of paper falls behind the horizon of distinguishability of my bodily senses (which determine the horizon of my look at these objects), leaving an infinitesimally small (now invisible for me) 'trace on the horizon' – an infinitesimally small geometric point. However, this geometric point actually still consists of a huge number of molecules, as if 'hidden' behind this infinitesimal geometric point, while the geometric point is their projection on the horizon (hence the 'trace' – as if the molecules disappeared beyond the horizon of my bodily senses, leaving only this trace in the form of an infinitesimally small geometric point being able to form the continuum with nearby similar traces/points). And the class that these remaining molecules form is a π -class. The class of the steps of tearing off the pieces of paper until this situation happens is a σ -class.

This brings us to another important part of Vopenka's AST, which I will mention here only very briefly: a set-theoretical framework for topology and continuum analysis based on naturally infinite semisets. To put the matter in a very simplified way, this framework gives us tools, for example, for the description and analysis of how *continuous shapes* of objects are created from the underlying structures of *discrete* entities, e.g. how the continuous shape of a piece of paper appears on top of the huge class of its discrete molecules. It is based on a formal definition of the *relation of indistinguishability*, which is a π -class defined as a class of underlying discrete entities that are mutually indistinguishable, e.g. the class of the paper's molecules that behind the horizon merge into a single geometrical point of the paper's continuous shape, or a class of mutually indistinguishable pixels on a high-resolution screen. This class is a π -class because it is defined as the complement of a *relation of distinguishability*, which is a σ -class of a primarily evident phenomenon of distinguishability of two objects. The relation of distinguishability contains, e.g. all pairs of the paper's molecules that I can mutually distinguish by cutting the paper into two pieces so that one molecule stays in

one piece of the paper and the other molecule stays in the other piece of the paper.

This is the reason why also all the colour tints that I cannot mutually distinguish constitute a π -class, whereas all the mutually distinguishable colour tints create a σ -class. And all the colour tints that I perceive as green create another σ -class.

3.3 Naturalness as Semiset

Having the intuition about naturalness together with the formal models of primarily and non-primarily evident phenomena, we can now link these two. We will do this here with the dichotomy natural–artificial, but the very same steps can also be taken for the dichotomy natural–unnatural.

From what we have discussed in Section 2.2, we can now see that *artificial* is the primarily evident phenomenon represented by a σ -class constructed from all objects (both physical as well as abstract, such as geometrical shape, quantity, etc.) that we can grasp by language and thus account for them as products of human *logos*-driven activity, hence as artificial. Therefore, the predicate ‘artificial’ says *what* ‘artificial’ is. Apparently, this σ -class is naturally infinite, i.e. it is a semiset, not a set, and therefore it is not a π -class.

In contrast, *natural* as the complement of *artificial* must be a naturally infinite π -class. It is constructed from all the features that we fail to grasp by language while constructing an understanding of the objects of the sensorily perceptible world. The constructing function of this π -class ‘collapses’ a naturally infinite number of these features into a ‘core’ that is common to all the instances of *naturalness* – saying *how* ‘natural’ is. This structure is also naturally infinite but in a very different way than the naturally infinite class of the extension of the predicate ‘artificial’. Just like in the example of ‘red’ and ‘non-red’ in the previous section, by observing only *natural* we would not be able see what *artificial* is, and we even would not know that it is *natural* what we are observing.

The quality *artificial* is, for example, that of a class of all the objects that somebody can see in her life. The quality *natural*, in contrast, is that

of a class of all the molecules of my table that I cannot distinguish one from another because they are all ‘behind’ one infinitesimal point that contributes to formation of the perceived topological continuum of my table.

Naturalness has a structure similar to that of a geometric point in Vopenka’s sense. It is also similar to the structure of *now* in Vopenka’s phenomenologically inspired model of time and the temporal continuum:³⁵ on the temporal axis underlaid by a huge number of very dense yet still discrete atomic time events, the *past* is a σ -class of all the time events that have already happened, the *future* is a σ -class of all the time events that will happen, and between these two σ -classes is a π -class of time events belonging to *now*. This *now* also has a structure similar to a geometric point, and it can be ‘wrapped’ or ‘embedded’ into a larger σ -class of *presence*. However, due to its σ -class nature, the *presence* class of the time events that I perceive as presently happening cannot be disjoint from the *past* and *future* classes. We can see that *artificial* (or *unnatural*) is metaphorically (and mathematically) similar to *past* (or *future*), while *natural* is similar to *now*.

We can also see that the ‘nature’ of *naturalness* is far less intuitively imaginable and visualisable than the ‘nature’ of *artificiality*. This is probably the reason why most of the discussions about what *natural* is treat it as a discursive placeholder with a σ -class structure typical for the easier-to-comprehend predicates of primarily evident phenomena (similar to the aforementioned structural difference between *now* and *presence* – with *natural* as a discursive placeholder being analogous to *presence*). This can lead to a textual analysis of the discursive usage of various σ -class predicates commonly labeled by the word ‘natural’³⁶, but it usually does not lead to a deeper understanding of what *naturalness* is -- or rather of *how* naturalness is, because we already know we should not ask *what* naturalness is in terms of extension of this predicate. So when making some decisions (perhaps ethically motivated) about *naturalness*, we must consistently take into account that we cannot get the answer to *what* natural is; instead we must pose the question as *how*

³⁵ Vopěnka (2001), Havel (2016).

³⁶ NCB (2015).

natural is, and possibly whether we interfere by our actions with this *how* or not.

We can also recall again the example with the wooden and plastic tables from Section 2.2 and model this situation with semisets. This means that now we are not modelling what *naturalness* is in general. Instead, we are trying to model a particular manifestation of *artificial* and *natural* in a given object, i.e. the structure of these predicates as they appear in the particular phenomenon that we cast our look upon.

Let H_{inst} be a horizon segment of instances when we are able to assert something about the observed object (e.g. the table). How H_{inst} exactly looks like is very much inherent to the particular look that is being cast – but in all cases it must be something compatible with the spatio-temporal givenness of a human being because it is exactly this givenness that we relate *natural* and *artificial* to. We know that if we can assert a proposition, we will be able to assert one more, but we also know that we will never assert billion propositions.³⁷ Then we can conceive a function F_{desc} defined on H_{inst} by which we are describing the observed object. For each instance from H_{inst} , F_{desc} returns a set of propositions being described about the object. The class $Urng(F_{desc})$ is apparently the σ -class of all that ‘can be grasped by language’ (namely, the language pertaining to the look being cast), hence the class modelling the object’s *artificiality*.

Let w be a theoretically conceivable set of all possible propositions about the object made in all conceivable languages within all conceivable looks being cast at the object (not only the looks cast from a human’s givenness). This is a bit problematic if we do not want to resort to the concept of an omnipotent God in whose mind these propositions are. However, we can for example imagine that there is a list of the actual states, positions, processes and configurations of all the molecules and other structural entities which the observed object is made of (or we can go to the atomic level, it is not that relevant for the argument here), including their causal connections to some higher level phenomena. Such a list might potentially be achievable one day as a result of scientific dis-

³⁷ Perhaps the aforementioned Zeus can assert a billion propositions, but this only means that his view of *naturalness* will be somewhat different from a human’s, because he casts different looks. The only speciality would be with the Christian God: He can assert the actually infinite number of propositions, thus for Him nothing is natural and all is artificial. But after all, this might not be eventually so surprising.

covery but it will never be accessible in its wholeness to a human mind. And we can imagine this list to be one of many subsets of w .

The class $w - \text{Urn}g(F_{desc})$ is apparently a π -class and it is exactly this class that models the *naturalness* of the observed object, i.e. all that is natural of this object. This class is extremely huge, yet still has the quality of a geometric point, as we have discussed above. All that is in this class is as if ‘sunk’ beyond the horizon on which it projects its trace, just as the pixels on a high-resolution screen are ‘sunken’ beyond the horizon of indistinguishability, projecting a continuous shape on the screen as their trace. It is not those propositions in w that are natural – we do not see them anyway. What is natural is their trace $w - \text{Urn}g(F_{desc})$. All that is natural in the given object is created by those entities from the internal structure of the object that are beyond the horizon of our look cast at the object – for example the π -class image of the mutual relations of billions and billions of molecules forming the shape of the wooden table. We know they are there but it is simply beyond the horizon of any language to describe how exactly they all are.

4 Conclusion

Getting back to our example with Hermes and his herb’s roots: if we took a 10-megapixel photograph of it and watched it on a screen, a lot of things would be artificial – i.e. those things that we could say about it, having the structure of a σ -class. However, something would still remain natural: the exact root’s shape that we see on the screen as the trace of the underlying pixels. The pixels themselves are indeed fully artificial (just as all the propositions in the set w) but we do not see them, we see only their trace in the form of a π -class. We may know (perhaps from the process of scientific discovery, or from the knowledge of how a computer screen works) that they are somewhere behind the horizon in the underlying structure of the shape that we see, but the only way to get to them, see them and distinguish them is to change the look that we cast at the object – looking at the screen from such a close distance (or even perhaps with a magnifying glass) so that we see the individual pixels – but losing the identity of the root itself, making it obsolete to ask about its naturalness.

What we have just described in the previous sections is actually the same principle we can observe in *emergent* phenomena³⁸. For example, within the emergentist paradigm of Artificial Intelligence, cognitive processes (or ‘mental’ phenomena) emerge on large complexes of locally quite simple units, such as in neural networks. Cognitive processes are π -class projections of the processes in neural networks running behind the horizon of our look cast at the systems to which we ascribe cognition. Neural networks are, metaphorically, like the pixels that we know are down there, yet we do not see them, only their images or projections on the ‘screen’ of the horizon pertaining to the look that we cast. And in the same way, the shape, colour and texture of the table are emergent phenomena over the molecular structures and processes within the table. If we look closer, the emergence disappears, but so does the shape; we can see the molecules, but we do not see neither the table, nor its shape and colour anymore.

What emergence needs (as I also argue elsewhere³⁹) is a naturally infinite semiset σ -class, such as the one constructed by F_{desc} , and its complementary π -class; and the emergent phenomenon that we then observe is the π -class constructed by what is behind the horizon. However, this leads us to the conclusion that what we see as the emergent phenomena has the same structure as what we see as the natural – i.e. *naturalness* and *emergence* have the same structure. In other words, we can say that what is emergent that is also natural: the emergence is the way how the naturalness (in the *physis* sense, not *human naturalness*) is manifested in the objects that we observe. Emergence is also *naturalness* that ‘creeps back’ into artificially created systems. We may create a perfectly artificial neural network, but it is its emergent cognitive level where the *naturalness* gets back. I am not saying that naturalness equals emergence because the presented arguments are too weak to plausibly support the assertion that whatever is natural that is also emergent. However, I personally dare to propose so and open it for further discussion.

For HCE systems, technologies and interventions, we can propose the same conclusion in terms of their naturalness too. An HCE system is as natural as it is emergent. This is indeed a completely value neutral

38 Havel (2000), Silberstein (2001).

39 Romportl (2016).

understanding of the *natural*. It will not help any ideological dispute between proponents and opponents of the particular system, nor will it say whether the system is good or bad.

For instance, if a dose of selective serotonin reuptake inhibitors (SSRI, a type of antidepressants) results in a moral enhancement of a subject so that she is more fair-minded and willing to cooperate⁴⁰, then this very high cognitive and social phenomenon is clearly the emergent phenomenon over the vast network of brain causal interactions triggered by the SSRI, and as such, it is *natural* (just like remission of severe depression symptoms is natural in a depressive patient after the SSRI treatment). Similarly, the general sociocultural acceptance of coffee drinking as a stimulant is also an emergent phenomenon over the vast network of social and cultural interactions; and so if the same happens one day to SSRI usage for increasing fair-mindedness, then the overall society-wide increased fair-mindedness will also be natural.

It seems that the only naturalness-relevant criterion for bioethical considerations is the dichotomy *unnaturalness* versus *human naturalness* – and not *artificiality* versus *naturalness*. We can either postulate that some HCE intervention or system is unnatural (hence unacceptable), or we can design a logical argument why it is unnatural (based, e.g., on religious or ethical argumentation). In both these cases, the *unnaturalness* will be created artificially, by language, in front of the horizon. However, the *unnaturalness* (especially as a collectively perceived repulsion) can also emerge over an incredibly large structure of many individual emotional attitudes interacting mutually within a complex sociocultural discourse, in which case we get the natural *unnaturalness*.

Therefore, further bioethical debate within HCE should elaborate this complex non-linear functioning of how the inherent *artificiality* and *naturalness* of HCE systems (as determined by emergence within those systems) interact and influence the *artificiality* and *naturalness* of the meta-level *unnaturalness* and *human naturalness* (as determined by emergence within the meta-level sociocultural systems absorbing, judging and interacting with those HCE systems). In all other contexts, the concepts of *artificial* and *natural* should be left completely value neutral, which is, I believe, good for them.

40 Tse & Bond (2002).

References

- Ashby, W. R. (1957): *An Introduction to Cybernetics*. Chapman & Hall, London.
- BMA (2007): *Boosting your brainpower: ethical aspects of cognitive enhancements*. Discussion paper, British Medical Association, London.
- Bostrom, N. & Sandberg A. (2009): "Cognitive enhancement: Methods, ethics, regulatory challenges." *Science and Engineering Ethics* 15 (3): 311–341.
- Bostrom, N. & Savulescu, J., eds. (2009): *Human Enhancement*. Oxford University Press, New York.
- Epstein, M. (2012): *Transformative Humanities: A Manifesto*. Bloomsbury Academic, London.
- Fukuyama, F. (2003): *Our Posthuman Future: Consequences of the Biotechnology Revolution*. Picador, New York.
- Havel, I. M. (2001): "Causal domains and emergent rationality." In *Rationality and Irrationality: Proceedings of the 23rd International Wittgenstein Symposium, 2000, Kirchberg, Austria*, eds. B. Broggard & B. Smith, öbv & hpt Verlagsgesellschaft, Vienna, 2001, pp. 129–150.
- Havel, I. M. (2016): "Model prožívaného času podle Petra Vopěnky." (English resume: "The Model of Experienced Time according to Petr Vopěnka") *Filosofický časopis / Philosophical Journal* 64 (4): 539–560.
- Holmes, R. M. (2006): "Alternative axiomatic set theories." In *Stanford Encyclopedia of Philosophy* [online]. 30. 5. 2006, substantive revision 12. 8. 2017 [cit. 3. 5 2019]. Available at: <https://plato.stanford.edu/entries/settheory-alternative/>.
- Ida, R. (2009): "Should We Improve Human Nature? An Interrogation from an Asian Perspective." In *Human Enhancement*, eds. N. Bostrom & J. Savulescu, Oxford University Press, New York, 2009, pp. 59–70.
- Kanamori, A. (2009): *The Higher Infinite: Large Cardinals in Set Theory from Their Beginnings*. Springer.

- Kass, L. (2004): *Life, Liberty and the Defense of Dignity: The Challenge for Bioethics*. Encounter Books, New York.
- McKibben, B. (2004): *Enough: Staying Human in an Engineered Age*. St. Martin's Griffin, New York.
- Moreno, J. D. (2011): *The Body Politic: The Battle Over Science in America*. Bellevue Literary Press, New York.
- Nancy, J.-L. (2002): "L'intrus." Translated by S. Hanson. *CR: The New Centennial Review* 2 (3): 1–14.
- NCB (2015): *Ideas about naturalness in public and political debates about science, technology and medicine*: Analysis paper. Nuffield Council on Bioethics.
- Nelson E. (1977): "Internal set theory: A new approach to nonstandard analysis." *A Bulletin of American Mathematical Society* 83 (6): 1165–1198.
- Robinson, A. (1966): *Non-Standard Analysis*. North-Holland Publishing, Amsterdam.
- Romportl, J. (2016): "Polomnožiny a emergence v umělé inteligenci." (English resume: "Semisets and Emergence in Artificial Intelligence") *Filosofický časopis / Philosophical Journal* 64 (4): 587–599. Available at: <http://filcasop.flu.cas.cz/index.php?id=casopis&cislo=4-2016&obsah=735>.
- Savulescu, J. & Maslen, H. (2014): "Moral enhancement and artificial intelligence: Moral AI?" In *Beyond Artificial Intelligence: The Disappearing Human-Machine Divide*, eds. J. Romportl, E. Zackova & J. Kelemen, volume 9 of Topics in Intelligent Engineering and Informatics, Springer, 2014, pp. 79–93.
- Silberstein, M. (2001): "Converging on emergence: Consciousness, causation and explanation." *Journal of Consciousness Studies* 8 (9–10): 61–98.
- Smajdor, A. (2015): "Naturalness and unnaturalness in contemporary bioethics." Technical report, Nuffield Council on Bioethics.
- Tse, W. S. & Bond, A. J. (2002): "Serotonergic intervention affects both social dominance and affiliative behaviour." *Psychopharmacology* 161 (3): 324–330.

- Vopěnka, P. (1979): *Mathematics in the Alternative Set Theory*. Teubner-Verlag, Leipzig.
- Vopěnka, P. (1989): *Úvod do matematiky v alternativnej teórii množín*. Alfa, Bratislava.
- Vopěnka, P. (1996): *Calculus Infinitesimalis, Pars prima*. Práh, Praha.
- Vopěnka, P. (2001): *Meditace o základech vědy*. Práh, Praha.
- Vopěnka, P. (2012): *The Great Illusion of 20th Century Mathematics and Its New Foundations*. University of West Bohemia, Plzeň.
- Vopěnka, P. (2014): *Alternativnaja teorija mnozhestv*. Izdatel'stvo Instituta matematiki, Novosibirsk.
- Vopěnka, P. (2015): *Nová infinitní matematika I.-IV*. Karolinum, Praha.
- Vopěnka, P. (2019): *New Infinitary Mathematics*. Pending publication.

Abstrakt

Přirozenost a umělost uchopená pomocí přirozeného nekonečna

Diskuze o přirozenosti, umělosti a nepřirozenosti v tomto článku jsou motivovány oblastí tzv. Human Cognitive Enhancement (HCE), čili technologickým vylepšováním kognitivních schopností člověka. Důvodem je potenciální schopnost HCE výrazně měnit a zasahovat do lidské osobnosti a identity. Článek nejprve navrhuje koncept lidské přirozenosti jako interakci mezi fysis a logos. Poté předkládá rámec pro intuitivní porozumění přirozenosti jako inherentní neschopnost jazyka plně popsat všechny vlastnosti objektu, který je přirozený. Analytické jádro článku pak navrhuje formální model přirozenosti založený na Vopěnkově fenomenologicky zakotvené alternativní teorii množin, jež je zde zároveň i stručně představena. Přirozenost a umělost jsou modelovány jako dvě strukturálně odlišné přirozeně nekonečné polomnožiny v rámci alternativní teorie množin. Tato klíčová strukturální odlišnost je následně analyzována a aplikována zpět na problém HCE.

Klíčová slova: přirozenost, umělost, nekonečno, logos, fysis, filozofie jazyka, vylepšování kognitivních schopností, alternativní teorie množin, polomnožina, biotika

Romportl, J. (2019): „Naturalness and Artificiality Revisited Through Natural Infinity.“ *Filosofie dnes* 11 (2): 4–33. Dostupné z www.filosofiednes.ff.uhk.cz

Reflexivita a propoziční postoje

Tomáš Šmiřák

*Filozofická fakulta
Ostravská univerzita
Reální 5, 701 03 Ostrava
smirakt@email.cz*

Teorie přímé reference vede k několika problematickým důsledkům týkajícím se propozičních postojů. Jeden z paradoxních důsledků je ukázán na příkladu Lois Lanové, z jejíhož přesvědčení, že Superman je silnější než Clark Kent, bychom údajně byli nuceni vyvodit, že je též přesvědčena o existenci někoho, kdo má tu vlastnost, že je silnější, než skutečně je. Snad žádný racionálně uvažující člověk by však přesvědčení o existenci někoho, o kom by tato nemožná vlastnost měla platit, zřejmě nepřijal. Klíčovým pojmem, na němž odmítnutí platnosti tohoto vyvození stojí, je *reflexivita*. Příspěvek se pokouší odhalit, jak je informace o reflexivitě v rámci sémantického obsahu získávána a jak je možné daný paradox vyřešit. Analyzována jsou nejprve stávající řešení: reflexivita je buď sémantickou vlastností vztahů, nebo anaforických zájmen. Obě možnosti budou vyvráceny prostřednictvím argumentace, že ve větách popisujících vztah mezi obsahy dvou koreferujících singulárních termínů nemůže být reflexivita součástí sémantické informace, jelikož znalost identity reference není podmínkou sémantické kompetence. Následně je představeno řešení, podle kterého je reflexivita součástí pouze pragmatického obsahu. Reflexivitu by v takových případech bylo možné získat jedině koordinací znalostí mezi mluvčími.

Klíčová slova: propoziční postoje, reflexivita, identita, přímá reference

Úvod

Text se zabývá paradoxy Churcha a Richarda ohledně analýzy propozičních postojů. Tyto paradoxy jsou důsledkem neurussellianismu. Základními východisky této teorie je, že propozice jsou strukturované a někte-

ré výrazy přispívají do obsahu propozice výhradně svou referencí (též millianismus či teorie přímé reference). Cílem je nalézt řešení, které tyto předpoklady zachovává a nevede přitom k paradoxním důsledkům.

Millianismus vede k případům, kdy objekty vztahu mohou být sémanticky identické, ale jejich identita nemusí být aktérům zřejmá. Reflexivitou se v kontextu této debaty myslí informace o identitě objektů vztahu. Problematická by byla například propozice, že Superman je silnější než Clark Kent. Jelikož má tato propozice formu aRa, mohlo by se zdát, že reflexivitu obsahuje. Přestože je Lois Lanová přesvědčena o pravdivosti této propozice, není přesvědčena o identitě Supermana a Clarka Kenta. Reflexivita tudíž není součástí této propozice.

Pokud má být zachována teorie přímé reference, tak je nutné najít původ reflexivity jinde než ve zmíněné sémantické formě. Matematický pohled na reflexivitu může sugerovat, že se jedná o vlastnost vztahů. Dalším běžným pohledem je, že reflexivita je výsledkem užití anaforického zájmena. Obě možnosti jsou analyzovány a vyvráceny. Podle následně předloženého řešení je reflexivita u diskutovaných typů vět součástí pragmatické části informačního obsahu. Tato koncepce umožňuje také řešení Churchova a Richardova paradoxu.

Nejprve je krátce něco řečeno o tom, co jsou to propozice, propoziční postoje a také, co stojí za přijetím teze o přímé referenci a strukturovaných propozicích. Následně je popsán Churchův a Richardův paradox. Oba tyto paradoxy spočívají v nerozlišení mezi tím, jestli propozice obsahují či neobsahují reflexivitu. Ve studii je rozebíráno několik možností, jak je možné reflexivitu v rámci propozice získat, a je ukázáno, proč a v kterém bodě je Richardův argument neplatný. Stejně tak je ukázána chyba v Churchově příkladu, která spočívá v přijetí mylného předpokladu.

Propozice a přímá reference

Většinou se zdá, že předmětem řady našich myšlenek a postojů jsou věci vnějšího světa. Pokud by například nyní někdo zaklepal na moje dveře, můžu nabýt přesvědčení, že je to moje sestra, jelikož říkala, že se v tuto dobu zastaví. Ona je pak součástí mého přesvědčení o tom, kdo nyní klepe. Informaci o tom, kdo klepe, můžu sdělit někomu dalšímu. Ve snaze

identifikovat posluchači danou osobu můžu v závislosti na tom, s kým se bavím, použít různé výrazy jako „sestra“, „Hanka“, „teta“ apod. Ve všech případech by bylo předmětem sdělení, že daná osoba klepe (posluchač nemusí nutně vědět, že se jedná o moji sestru).

Přesvědčení, které jsem prvně nabyl (že daná osoba klepe), je tím, co jsem následně sdělil prohlášením zvolené věty, a pokud jsem byl přesvědčivý, tak je i tím, o čem nabyl přesvědčení i můj posluchač. Ve filozofickém žargonu se informacím tohoto typu říká propozice. Propozice mají plnit úlohu předmětu postojů (přesvědčení, obavy aj.) a sémantického obsahu vět. Podle uvedeného příkladu to také vypadá, že by se mělo jednat o komplexy, jejichž konstituenty by byly mj. individua jako Hanka a vlastnosti jako zvoní.

Mezi zastánci propozic je problematický zejména tento poslední bod. Nejznámější je zřejmě Fregův paradox, který Frega vedl k odmítnutí přímé reference a postulování teorie smyslů.¹ Od vzniku Fregova článku vznikla řada teorií ohledně struktury a konstituentů propozic či obecněji ohledně toho, co je předmětem postojů a sémantickým obsahem vět.

Jiným přístupem je snaha zodpovědět možné námitky a zachovat přitom předpoklad o přímé referenci. Pokusy o „obohacení“ obsahu propozic o způsoby prezentace věcí, kauzální řetězce, slova aj. často přináší více problémů, než kolik jich řeší.² Podle zastánců přímé reference jsou to objekty samotné, z čeho se propozice skládají. Problematické předpoklady, z kterých vychází paradoxy Churcha a Richarda, lze shrnout do tří bodů:

- (i) Propozice jsou předmětem postojů a sémantickým obsahem vět. Mezi konstituenty propozic patří objekty, vlastnosti a vztahy.
- (ii) Jedinou sémantickou hodnotou singulárních termínů je reference.
- (iii) Postojová věta „x p, že S“ popisuje vztah mezi referencí x a sémantickým obsahem S, vyjádřený postojovým slovesem p. Postojová věta je pravdivá, nachází-li se referent x

1 Frege (1892, s. 26).

2 Shrnutí např. v Richard (1997).

vůči sémantickému obsahu vedlejší věty S v daném postojo-
vém vztahu vyjádřeném p.

Přesvědčení o existenčních propozicích a reflexivita

Argumenty Churcha i Richarda mají společné to, že je agentovi při-
psáno přesvědčení týkající se dvou objektů, které jsou sice identické,
jejichž identitu si však daný agent neuvědomuje. Oba z toho vyvozují sice
jiné, ale vždy paradoxní důsledky.

Churchův³ argument se týká přesvědčení o existenčních propozicích.
Vychází z toho, že zřejmě žádný racionálně uvažující člověk není pře-
svědčen, že by nějaký objekt mohl být odlišný od sebe sama. Nikdo tu-
díž nemá přesvědčení, které by mělo formu, že existuje nějaké x takové,
že $x \neq x$. Přírozeným důsledkem je rovněž to, že nikdo není přesvědčen,
že existují x a y takové, že $x \neq y$, mají-li být objekty daného přesvědčení
identické. Churchův příklad se odvolává na krále Jiřího IV., jenž byl ales-
poň v jednom okamžiku svého života přesvědčen, že Walter Scott není autor
románu Waverley. S použitím proměnných bychom postoj krále Jiřího
popsali jakožto přesvědčení o existenci x a y takových, že $x \neq y$. S ohledem
na to, že nikdo, a tedy ani král Jiří, nemůže být přesvědčen, že nějaký
objekt je odlišný od sebe sama, tak platí, že pokud je přesvědčen, že $x \neq y$,
pak objekty ohodnocující x a y skutečně musejí být odlišné. Jenomže,
jak známo, Walter Scott skutečně je autor románu Waverley, přičemž
samotný fakt, že si král Jiří myslí, že tomu tak není, je ještě odlišnými
osobami nečiní. To je ovšem v rozporu s důsledky jeho zdánlivě chybějícího
přesvědčení, že existuje x takové, že $x \neq x$, respektive že existují x a y
takové, že $x \neq y$ (kdy x a y jsou identické). Buď jsme tedy nuceni přijmout,
že lidé mají schopnost učinit objekty odlišnými už na základě svého pře-
svědčení ohledně jejich odlišnosti, což se totiž zdá být přírozeným dů-
sledkem existenčních přesvědčení, nebo jsme nuceni odmítnout, že by
obsahem podobných přesvědčení měly být konkrétní objekty vnějšího
světa, což by v případě krále Jiřího znamenalo, že se jeho přesvědčení
netýká Scotta samotného.

3 Church (1982, s. 61–62).

Churchův argument je problematický zejména pro teorie identifikující obsah přesvědčení s pravdivostními podmínkami, jelikož jakékoli přesvědčení by s ohledem na pravdivostní podmínky mohlo být identifikovatelné s konjunkcí této propozice a její existenční generalizace.⁴ Teorie uznávající strukturované propozice se tomuto problému v jistém ohledu vyhýbají. Jakákoli propozice vyjádřená větou se singulárními termíny je v rámci těchto teorií obsahově odlišná od existenční generalizace této propozice. To ale ještě neznamená, že se tyto teorie Churchově námitce vyhýbají úplně. Nejspíše bychom připustili, že je-li někdo přesvědčen, že *a je P*, pak je zřejmě přesvědčen, že existuje nějaký objekt, pro nějž daná vlastnost *P* platí.⁵ Dejme tomu, že by někdo svědomitě prohlásil větu „Honza je šťastný“ – pak bychom usoudili, že je přesvědčen o singulární propozici (1).⁶

(1) Honza je šťastný

Současně je rozumné se domnívat, že přestože to mluvčí přímo neuvedl, je přesvědčen o existenci někoho šťastného. Taková propozice by mohla být znázorněna jako (2).

(2) $[\exists x](x \text{ je šťastný})$

Z hlediska obsahu jsou propozice (1) a (2) odlišné. Předpokládanému mluvčímu na základě jeho prohlášení „Honza je šťastný“ nemůžeme přímo přisoudit přesvědčení o (2), jelikož (2) není obsahem prohlášení „Honza je šťastný“. Přesto bychom zřejmě komukoli, kdo by tuto větu svědomitě prohlásil, a o němž bychom tak mohli říci, že je přesvědčen o propozici (1), mohli přisoudit rovněž přesvědčení o (2). Je tudíž možno odvodit, že existuje určitý psychologický zákon (P) pro existenční generalizaci obsahů přesvědčení.

4 Propozice, že o je P, má stejné pravdivostní podmínky jako konjunkce propozic, že o je P a že existuje x takové, že x je P.

5 Salmon (1986, s. 39, 47).

6 Předpokládá se, že platí, že když někdo svědomitě prohlašuje větu S, tak je přesvědčen o obsahu věty S v daném kontextu. Podrobněji Kripke (1979, s. 248–249).

(P) Kdokoli je přesvědčen o singulární propozici φa , je přesvědčen o existenční generalizaci této propozice $[\exists x](\varphi x)$

Pokud je (P) platným zákonem, aplikovatelným na Jiřího přesvědčení, že Scott není autor Waverley, pak je Churchova námitka stejně kritická i pro teorie identifikující sémantický obsah se strukturovanými propozicemi. Uznáme-li tedy, že se Jiřího přesvědčení vztahuje k objektům vnějšího světa (přesvědčení de re), tak je z jeho prohlášení „Scott není autor Waverley“ vzhledem k (P), Kripkovu principu⁷ a Leibnizovu zákonu⁸ možné odvodit, že je přesvědčen, že $[\exists x](x \neq x)$.⁹

Argument představený Richardem¹⁰ poukazuje na další problematické důsledky pro teorie uznávající singulární propozice. Richard tímto argumentem demonstruje, že i kdyby singulární propozice mohla být sémantickým obsahem některých vět, postojová věta přesto vztah mezi agentem a singulární propozicí vyjadřovat nemůže. Nějaký hrdina z komiksového světa, snažící se vysvětlit, co si Lois Lanová myslí, by například mohl prohlásit „Lois je přesvědčena, že Superman je silnější než Clark Kent“. Lois je skutečně přesvědčena o vztahu mezi Supermanem a Clarkem Kentem, kdy první z těchto dvou je silnější než ten druhý. Propozici, která je obsahem jejího přesvědčení, je možné znázornit jako (3a).

(3a) Superman je silnější než Clark Kent

Pokud jsou uznány předpoklady (i)–(iii), pak je na základě Leibnizova zákona možné odvodit, že Lois je přesvědčena o propozici (3b), která je s (3a) identická.

7 Kdokoli svědomitě prohlašuje větu S, je přesvědčen o sémantickém obsahu věty S. Kripke (1979, s. 248–249).

8 Záměna jednoho členu celku za jiný člen se stejným významem nevede ke změně významu celku; srov. Leibniz (1968, s. 308).

9 a. Jiří: „Scott není autor Waverley“

b. Scott je autor Waverley

c. Jiří je přesvědčen, že Scott není autor Waverley

d. Jiří je přesvědčen, že existuje x a y takové, že $x \neq y$

e. Jiří je přesvědčen, že existuje x takové, že $x \neq x$

premisa

premisa

a, Kripkův princip

c, (P)

de re, b, d, Leibnizův zákon

10 Richard (1983, s. 39–47).

(3b) Superman je silnější než Superman

Každý, kdo by byl přesvědčen o (3b), by zároveň měl být schopen odvodit (3c).

(3c) Superman je silnější než on sám

Situaci je možné dále zkomplikovat tím, že je možné uznat, že pro dvoumístné reflexivní vztahy existují jednomístné vlastnosti, které jsou s těmito vztahy synonymní. Na tomto základě bychom z (3c) mohli rovněž odvodit (3d), kde je Supermanovi připisována reflexivní vlastnost, že je silnější, než jak silný skutečně je.

(3d) Superman má vlastnost být silnější, než skutečně je.

Aplikací (P) na (3d) bychom získali důsledek (3e), který reprezentuje Loisino přesvědčení o existenci někoho, o němž platí nemožná vlastnost, že je silnější, než jak silný skutečně je.

(3e) $[\exists x](x \text{ je silnější, než jak silný skutečně je})$

(3b)–(3e) jsou problematickými důsledky údajného Loisina přesvědčení. Richarda absurdita tohoto vyvození vede k formulování vlastní sémantické teorie, která jde proti předpokladu (iii).¹¹ Problém tohoto vyvození spočívá v tom, že se z přesvědčení, které nezahrnuje reflexivitu, vyvozuje přesvědčení, které reflexivitu zahrnuje. Loisino přesvědčení sice zahrnuje dvakrát totéž individuum, na základě neznalosti o jejich identitě je však vztah mezi nimi podle Lois nereflexivní. Na konci tohoto vyvození¹² je však Lois přisuzováno přesvědčení, které zahrnuje znalost identity referentů, a tudíž si má být vědoma, že daný vztah je reflexivní.

I když je zřejmé, proč je dané vyvození problematické, není jasné, jak jej vyřešit v teoretickém rámci přímé reference. Salmon nezaujímá konkrétní stanovisko ohledně toho, který přechod je neplatný. Argumentuje

¹¹ Richard (1990, s. 133–141).

¹² V kterém bodě přesně je předmětem sporu, který je rozebírán dále.

pouze, že vyvození je neplatné, přičemž je sporné, jestli je (3c) platným vyvozením z (3b).¹³ Soames se sice věnuje stejnému problému,¹⁴ ale explicitně se nevyjadřuje k tomu, kde Richardův argument selhává. Z jeho článku o anaforách u propozičních postojů¹⁵ je ovšem možné odvodit, že se kloní k mylnosti přechodu z (3b) na (3c).

Následující kapitoly se věnují několika možnostem, jak je možné reflexivitu v rámci propozice získat. Argumentuje se, že reflexivita není součástí obsahu vztahů ani anaforických zájmen. Následně je zaujato stanovisko, podle kterého by byl neplatný teprve přechod z (3c) na (3d).

Reflexivita vztahů

Jedna z možností, jak si zachovat představu, že Lois skutečně je přesvědčena o singulární propozici (3a), a zároveň odmítnout, že je přesvědčena o jejím zdánlivém důsledku (3e), spočívá v přiřazení reflexivity k vlastnostem vztahů. Můžeme tak rozlišit vztahy, které jsou sémanticky reflexivní či nereflexivní. Zdánlivě je možné považovat vztahy za reflexivní, pokud jsou objekty tohoto vztahu identické. Lidé ale mohou být přesvědčeni, že nějaký vztah je reflexivní, přestože objekty tohoto vztahu jsou odlišné, nebo nereflexivní, i když tyto objekty jsou identické. Jedná se o důsledek toho, že naše myšlení podléhá omylům, v důsledku čehož býváme přesvědčeni o nemožných propozicích.¹⁶ Logicky reflexivní vztah tedy může být sémanticky nereflexivní, má-li totiž sloužit k popisu vztahu mezi dvěma zdánlivě neidentickými objekty. Takový vztah by byl taktéž obsahem Loisina přesvědčení. Pokud je reflexivita sémantickou vlastností vztahů, pak jakákoli reprezentace propozice (3a)–(3c) může být správná či chybná v závislosti na tom, jaký vztah je zde reprezentován. Lois je přesvědčena o propozici, která obsahuje nereflexivní vztah mezi Clarkem Kentem a Supermanem, přičemž (3d) by důsledkem takové propozice být nemohlo, jelikož vlastnost, jež je v rámci ní Supermanovi připisována, koresponduje výhradně s reflexivním vztahem. Mezi (3a)–

13 Salmon (1992, s. 59).

14 Soames (1985, s. 60).

15 Soames (1994, s. 125–134).

16 Klasickým příkladem by byla přesvědčení o matematických omylech či logicky sporných propozicích *p a ne-p*.

(3c) je tak možné rozlišit dva typy propozic (4) a (5), z nichž (4) obsahuje nereflexivní vztah R mezi jedním a tímž individuem.

(4) xRx ¹⁷

Otevřená propozice (4) obsahuje dva výskyty x , ale neobsahuje informaci o jejich identitě. Má-li naproti tomu propozice obsahovat sémanticky reflexivní vztah, pak si ten, kdo tuto propozici zvažuje (tvrdí ji, je o ní přesvědčen apod.), musí být vědom také toho, že dané výskyty x , jež jsou objektem reflexivního vztahu, jsou výskyty identického objektu. Tuto informaci o identitě je kvůli rozlišení od (4) možné v rámci propozice reprezentovat pomocí λ -operátoru vázajícího objekty vztahu R, jak je znázorněno v (5).

(5) $(\lambda x')(x'R'x')x$

Lois je přesvědčena o propozici typu (4), týkající se totožného individua a vyjadřující nereflexivní vztah k tomuto individuu. Propozice tohoto typu neobsahují informaci o identitě x , a tak se (4) zdá být vhodnou formou k reprezentaci přesvědčení Lois. Naproti tomu zvažování propozice typu (5) vyžaduje znalost identity výskytů x , a jednalo by se tak o něco, o čem Lois přesvědčena není.

Na základě tohoto přístupu bychom Loisino přesvědčení o (3d) mohli odvodit pouze, pokud by její přesvědčení, že Superman je silnější než Clark Kent, mělo formu (5), nikoli (4). Pouze kdyby Loisino přesvědčení zahrnovalo kromě dvou stejných individuí také informaci o jejich identitě, byla by schopna ze vztahu *silnější než*, jenž by podle ní v takovém případě platil pro totožné individuum, odvodit, že pro toto individuum platí synonymní reflexivní vlastnost *silnější než skutečně je*. Lois si ovšem identity Clarka Kenta a Supermana vědoma není, a tak nemůže být přesvědčena o propozici (3) formou (5), jež informaci o identitě zahrnuje. Jakýkoli vztah, mající platit mezi Supermanem a Clarkem Kentem, je

17 Sémantickým důsledkem přímé reference je, že v případě Lois by se propozice reprezentovala jako $\$R\$$, kde $\$$ symbolizuje Supermana čili Clarka Kenta. Přestože tato propozice obsahuje totéž individuum, Lois si jejich identitu neuvědomuje, a tak součástí této propozice nemůže být informace o jejich identitě. Stejně tak by informace o jejich identitě nebyla součástí obsahu prohlášení „Superman je silnější než Clark Kent“.

z jejího pohledu nereflexivní. Žádná propozice ve formě $xR^{neref}x$, složená z nereflexivního vztahu R^{neref} , neobsahuje informaci o identitě daných výskytů x , a tak z přesvědčení o této propozici není možné vyvodit přesvědčení o xP^{ref} zahrnující reflexivní vlastnost P^{ref} , kde by xP^{ref} bylo synonymní s $xR^{ref}x$, čili $(\lambda x')(x'R'x')x$.

V Loisině případě by tak (3d) nebylo platným vyvozením ohledně jejího přesvědčení.¹⁸ Nyní je nutné zjistit, jestli je možné, aby věty ve formě (6), jež popisují vztah mezi obsahy dvou běžných koreferujících singulárních termínů, mohly vyjadřovat dvoje propozice typu (4) a (5).

(6) a. $\alpha R \alpha$

b. $\alpha R \beta$

Zodpoví se tím, jestli by bylo možné najít řešení daného paradoxu již v bodě (3a) či (3b). Pokud by se například ukázalo, že věty ve formě (6a) i (6b) mohou být obě ohledně reflexivity víceznačné, pak by celé vyvození (3) mohlo být platné i neplatné v závislosti na tom, jestli věta (3a) obsahuje reflexivitu či nikoli. Pokud by se ovšem ukázalo, že pouze věta ve formě (6b) může být víceznačná, bylo by možné odmítnout platnost vyvození z (3a) na (3b). Pokud by ani jedna z vět nemohla být víceznačná, bylo by nutné hledat řešení reflexivity jinde.

Věta ve formě (6a), která obsahuje dvakrát totéž jméno, by snad mohla vyjadřovat propozici s reflexivním vztahem v rámci následujícího příkladu:

Dva lidé se baví o tom, zda existuje lepší zpěvák než Karel Gott. Při odchodu z jeho vystoupení se první z nich ptá „Myšlíš, že někdo bude mít někdy lepší vystoupení než dnes Karel Gott?“ a druhý na to odvětlí „Jedině Karel Gott může mít (někdy) lepší vystoupení než (dnes) Karel Gott“.

Zde je zřejmě Karlu Gottovi připisován reflexivní vztah, že jedině on může někdy podat lepší výkon, než jaký dnes podal on sám. Reflexiv-

¹⁸ Zůstává otázka, kam například zařadit (3c) – zda tato propozice nutně obsahuje reflexivitu, či nikoli. Této otázce je věnována vlastní kapitola.

ní interpretace je možné docílit také užitím dvou různých koreferujících termínů. Řekněme, že by se Daniel Landa alias Žito těšil podobné slávy jako Karel Gott. Někdo by mohl v podobné situaci prohlásit „Jedině Žito může mít lepší vystoupení než Daniel Landa“. Řečník tak namísto stejného jména může použít dvě různá jména, aby například zdůraznil, že koncert mohl být lepší pod šaškovskou záštitou značky Žito (ve smyslu, že kdyby Landa projevil více šaškovské manýry, jež jsou s ním spjaty od doby, kdy se prohlásil za Žita, vystoupení by mohlo být ještě lepší).

Zmíněné příklady vět ve formě (6a) a (6b) by vyjadřovaly propozici typu (5). V obou případech je Gottovi, respektive Landovi, připisován reflexivní vztah, že jen on sám může předvést lepší vystoupení. Příklad s Lois ukazuje, že věta ve formě (6b) může vyjadřovat také propozici s nereflexivním vztahem. Poslední možností, jež chybí, je případ, kdy věta ve formě (6a), zahrnující dva stejné koreferující termíny, vyjadřuje nereflexivní propozici typu (4). Taková situace může nastat, když se někdo mylně domnívá, že vlastní jméno referuje ke dvěma různým lidem. Příkladem může být Kripkeho Pierre,¹⁹ jenž zná muzikanta Paderewského a slyšel o politikovi Paderewském. Ve skutečnosti se jedná o jednu osobu, Pierre si však myslí, že jsou to dva různí lidé. Pierre by tak mohl popisovat určitý vztah týkající se Paderewského v domnění, že popisuje vztah mezi dvěma lidmi.

Důsledek tohoto přístupu spočívá v tom, že všechny obecné výrazy, schopné vyjadřovat jak reflexivní, tak nereflexivní vztah, jsou víceznačné. Význam výrazu *mít lepší vystoupení než* se tak mění v závislosti na tom, jaký má mluvčí postoj k identitě objektů tohoto vztahu. Kompetentní mluvčí pro každý výskyt takového obecného výrazu musí umět rozpoznat, o jaký typ vztahu se konkrétně jedná. Jednoduchou odpovědí na otázku, na základě čeho ohodnocuje mluvčí daný vztah jako reflexivní, by bylo, že daný vztah je reflexivní, pokud jsou objekty tohoto vztahu identické. To však není možné, jelikož propozice (4) i (5) vyjadřují vztah k identickým individuům. Podobně není reflexivní vztah ani součástí Loisina přesvědčení o propozici (3), přestože se jedná o vztah týkající se totožné osoby. Tím, co je pro ohodnocení daného vztahu jakožto reflexivního nutné, je tudíž přesvědčení o identitě objektů tohoto vztahu. Pokud se mluvčí domnívá, že objekty, mezi něž je určitý vztah vkládán, jsou identické, pak tento vztah ohodnotí jakožto reflexivní.

19 Kripke (1979, s. 265).

Tím je ovšem kladena nemožná podmínka pro porozumění sémantickému obsahu vět (6). Mluvčí sice rozumí větě „Ema mele maso“, sémantický obsah takové věty však tento mluvčí zná jen tehdy, pokud ví, o jakou Emu se konkrétně jedná – tj. zná referenci daného vlastního jména. To je ovšem nedostatečná podmínka pro věty ve formě (6). Třebaže mluvčí zná referenci pro α a β , nemusí proto ještě vědět, zda konkrétní výskyt této věty vyjadřuje propozici typu (4), nebo (5). Mohlo by se případně uznat, že je se sémantickým obsahem těchto vět obeznámen, protože má názor na koreferenci termínů α - α či α - β . V takovém případě by si však jiný mluvčí, rovněž znající referenci těchto termínů a mající odlišný názor na jejich koreferenčnost, nebyl schopen s prvním mluvčím porozumět. Pro tyto dva mluvčí by daná věta vyjadřovala odlišnou propozici. Měli bychom tak dva kompetentní mluvčí, kteří použitým výrazům rozumí, pro něž však daná věta vyjadřuje odlišnou sémantickou propozici.

Další problém tohoto přístupu se týká případů, kdy by měl mluvčí popisovat někomu přesvědčení o obsahu věty ve formě (6). Vraťme se k příkladu s Lois: daný mluvčí si může být vědom identity Supermana a Clarka Kenta, na základě čehož by měl větě „Superman je silnější než Clark Kent“ rozumět reflexivně. Pokud by však užitím této věty popisoval Loisino přesvědčení, tak by tím popisoval něco, o čem Lois přesvědčena není. Lois sice je přesvědčena, že Superman je silnější než Clark Kent, vztah mezi nimi je však podle ní nereflexivní. Oproti tomu mluvčí, popisující Loisino přesvědčení užitím věty „Superman je silnější než Clark Kent“, si jejich identity vědom je, a tak je vztah v jeho prohlášení reflexivní. Vzniká tudíž paradox, kdy mluvčí není schopen popsat Loisino přesvědčení jen proto, že má jiný názor na identitu Supermana a Clarka Kenta. Jediná možnost, jak v rámci dané strategie tento paradox vyřešit, spočívá v opuštění některé z výchozích tezí (i)–(iii).²⁰ Pokud se mají tyto teze zachovat, pak je nutné odmítnout, že reflexivita je sémantickou vlastností vztahů.

²⁰ Zřejmě by se muselo uznat, že se do vět jako „Lois je přesvědčena, že Superman je silnější než Clark Kent“ promítá znalost mluvčího ohledně znalostí druhého (Lois). Pokud by si mluvčí myslel, že Lois ví, že Superman a Clark Kent jsou jednou osobou, tak by byl vztah v domněnkové větě reflexivní, v opačném případě by byl nereflexivní. V takovém případě by byl obsah této věty závislý nejen na znalostech jednotlivých mluvčích ohledně identity Supermana a Clarka Kenta, relativizoval by se také vůči přesvědčením těchto mluvčích ohledně Loisina povědomí o této identitě.

Anafora²¹

Existuje jiná strategie, jak přijmout Loisino přesvědčení o (3a) a zároveň se vyhnout důsledku (3d) a (3e). Reflexivita totiž nemusí být sémantickou vlastností vztahu samotného, mohla by být produktem anaforického zájmena. Na základě tohoto přístupu by propozice (3a) a (3b) byly obsahově identické a tudíž by obě tvořily obsah Loisina přesvědčení. Obsahově by se od nich lišila teprve propozice (3c) zahrnující anaforu, z níž by rovněž vyplývaly propozice (3d) a (3e).²² Účelem tohoto přístupu je docílit, aby věty jako (7) vyjadřovaly na rozdíl od vět (6) propozici zahrnující reflexivitu.

(7) Honza miluje svoji matku

Teorie tohoto typu lze souhrnně označit jako teorie reflexivních anafor (TRA) a jako takové kontrastují s teorií prosté anafory (TPA), podle níž je sémantický obsah anafory identický s obsahem jejího antecedentu, a jež tudíž reflexivitu neprodukuje.²³ Řekněme, že Jitka je přesvědčena o obsahu věty (7). Na základě TPA by Jitka byla přesvědčena o propozici (8), v rámci níž je Honza hodnotou proměnné x .

(8) Honza miluje x (ovu) matku

Jitka by o této propozici byla přesvědčena, kdykoli by svědomitě prohlásila větu, v které by byl výraz (jméno či zájmeno) pro x , který referuje k Honzovi. Není nutné, aby Jitka věděla, že daný výraz pro x referuje k téže osobě, jejíž lásku k matce popisuje. Mohla by například ukázat na nějakého muže, kterého nepoznává jako Honzu, o jehož matce chce říct, že k ní Honza chová jinou než mateřskou lásku, a prohlásit (9).

(9) Honza miluje jeho matku

²¹ Anafora je výraz, jehož obsah je závislý na určité sémantické informaci nacházející se ve větě. Předmětem zde jsou anaforická zájmena se singulárním termínem v antecedentu. Jedná se o zájmena jako „on“, „jeho“ apod., jež se vztahují k dříve zmíněným singulárním termínům.

²² Uvedené platí pouze v případech, že bychom přijali, že pro reflexivní vztahy existují (alespoň v některých případech) vlastnosti, jež přispívají k obsahu věty stejnou hodnotou. Tento předpoklad je možné odmítnout a s ním i důsledek (3d). Pro daný problém je to nepodstatné, jelikož by byl blokován už přechod z (3b) do (3c).

²³ Salmon (1992, s. 59).

Zájmeno „jeho“ v tomto prohlášení funguje tak, že vybírá referenci v závislosti na tom, na koho se ukazuje. Prohlášení (9) by tak rovněž vyjadřovalo propozici (8). Obecně se má za to, že věty (7) a (9) (v daném kontextu) nemohou tvrdit totéž. TRA mezi propozicemi těchto vět rozlišuje a umožňuje tvrdit, že kdo je přesvědčen o obsahu věty (9), nemusí být ještě přesvědčen obsahu věty (7). Podle TRA je zkratka opakování obsahu antecedentu nedostačující. Propozice by měla reflektovat, že se jedná o anaforu a propozice tak musí zahrnovat informaci o vztahu anafory ke svému antecedentu. Tuto informaci je možné opět reprezentovat λ -operátorem, jenž by obsah anafory svázal s obsahem jejího antecedentu. Taková propozice by byla reprezentovaná v rámci (10).

(10) $(\lambda x)(x \text{ miluje } x(\text{ovu}) \text{ matku}) \text{ Honza}$

Kdokoli zvažující propozici (10) si je vědom identity výskytů x a je si tak vědom, že ten, jehož láska k něčí matce se popisuje, je tím, o jehož matku se jedná. Nemá tudíž problém odvodit, že o Honzovi platí vlastnost, že miluje vlastní matku. Ke stejnému vyvození by Jitka, tvrdící propozici (8) prohlášením (9), nedospěla, jelikož nerozpoznává identitu toho, jehož lásku k něčí matce popisuje, s tím, na nějž ukazuje. Provedená analýza je tudíž schopna elegantním způsobem rozlišit mezi obsahem demonstrativně užitého zájmena, součástí jehož obsahu je prostá reference, a zájmena anaforického, jehož obsah je vázán k obsahu antecedentu. Z téhož důvodu se analýza pomocí TPA zdá být neadekvátní, protože by v rámci věty (7) identifikovala pouze reflexivně indiferentní vztah k matce někoho, jímž je ve skutečnosti on sám, což ovšem v rámci propozice není dostatečně znázorněno.

Pokud se uzná, že obsah anafory funguje na zmíněném reflexivním principu, tak je obtížné vysvětlit, jak anafora funguje v případech, kdy se antecedent nenachází ve stejné větě jako anafora. Následující diskurz (D) je příkladem, kdy se antecedent nachází v jiné větě než anafora:

(D) Jirka se vloupal do Honzova bytu (...) Martin si myslí, že (on) se tam dostal oknem.

Otázkou je, jaká propozice je sémantickým obsahem věty (11) diskurzu (D).

(11) Martin si myslí, že on se tam dostal oknem.

Ve větě (11) chybí prvek, k němuž by anafora svůj obsah vázala. Základním předpokladem strukturovaných propozic je, že struktura propozice odráží strukturu věty, jež tuto propozici vyjadřuje. V rámci této věty se však nachází pouze jediný výraz, referující k Honzovi. Pokud by byl obsah propozice obohacen o referující antecedent, jenž by nějakým způsobem přesahoval do první věty (D), v které se antecedent *Jirka* nachází, tak by obsahově neodpovídal větě (11). Přesto se zdá, že anaforické zájmeno „on“ i v tomto případě musí obsahovat více než jen pouhou referenci k Jirkovi. Účelem je tedy docílit toho, aby bylo z obsahu diskurzu patrné, že si je Martin vědom identity jedince, jenž se vloupal do Honzova bytu, s tím, kdo se tam dostal oknem. Ze stejného důvodu jako v rámci příkladu (9) by TPA takové interpretace nedosáhla. Je-li účelem docílit reflexivní interpretace (D), pak je nutné kromě vázané proměnné formulovat ještě jiné sémantické mechanismy.²⁴

Nesémantická reflexivita

Jiná možnost, jak docílit reflexivní interpretace, spočívá v analyzování těchto zájmen jakožto běžných singulárních termínů, opakující referenci svého antecedentu, a reflexivitu vysvětlit pomocí jiných než sémantických procesů. Důsledkem takového přístupu by bylo, že neexistuje sémantický rozdíl v obsahu tvrzení „Ona šla do kina“ v závislosti na tom, zda je zájmeno „ona“ použito demonstrativně nebo anaforicky. Rozdíl spočívá výhradně v tom, jak je obsah tohoto termínu určován. Buď se na tuto ženu ukazuje, nebo se k ní referuje prostřednictvím antecedentu, obsah prohlášení však zůstává stejný.

Výše bylo ukázáno, proč může být problematické hodnotit (3a)–(3c) jako propozice zahrnující reflexivitu, ať už by se jednalo o produkt vztahu nebo anaforického zájmena. Současně se projevilo, že v případě vět ve formě (6) je nutné znát více než jen sémantický obsah daných výrazů, abychom dokázali rozhodnout, zda je součástí vyjádřené informace takéž propozice s reflexivitou (5). Reflexivní interpretaci je možné získat

²⁴ Soames (1994, s. 122–134) popisuje, že ani odmítnutí předpokladů (ii) a (iii) nevede k lepším výsledkům.

jiným způsobem než přes sémantický obsah. Obecnou představou milliánů je, že v rámci domněnkových vět se tvrdí, že někdo je přesvědčen o sémantickém obsahu věty užití k popisu aktérova přesvědčení. Millián taktéž uzná, že existují různé způsoby, jak někdo může být o dané propozici přesvědčen, a alespoň část informace o agentově způsobu přesvědčení může být zahrnuta coby pragmatický obsah promluvy, závislý na koordinaci znalostí mezi mluvčími. Reflexivita by mohla být právě takovým typem informace. Tento přístup tak nevede k důsledku, že existuje dvojitý význam pro každý obecný vztahový nebo anaforický výraz. Vystačí si s významem jediným, jenž je vůči reflexivitě indiferentní. Zda je daná věta použita se záměrem vyjádřit reflexivní informaci, závisí teprve na epistemickém pozadí.

Kompetentní mluvčí sémantickému obsahu věty (6) rozumí, pokud ví, o jaký vztah se jedná, a zná referenci těchto výrazů. Nemusí mít znalost ohledně identity referentů, tudíž nemusí ani vědět, jestli prohlášení této věty zahrnuje informaci o reflexivitě. Totéž platí pro postojovou větu ve formě (12) s přidanou znalostí sémantického obsahu výrazu *c* a postojového slovesa *p*.

(12) *c p*, že α R β

Na základě dalších znalostí mezi mluvčími mohou tito mluvčí interpretovat věty ve formě (6) tak, že zahrnují reflexivitu. Tato další znalost by mohla vypadat následovně: mluvčí *A* si je vědom koreference α a β a je si vědom, že posluchač *B* ví, že *A* si je vědom stejnosti reference α a β atd. Ohledně použití postojové věty ve formě (12) bylo řečeno, že může zahrnovat informace týkající se toho, jak aktér postoje uvažuje o propozici, jež je mu připisována. Osoby *A* a *B* hovoří o přesvědčení *C*. Přidaná informace, již by *A* a *B* pro vyvození reflexivní informace potřebovali, se tak vztahuje k přesvědčení *C*. Tudíž *A* a *B* mohou prohlášení (12) vzájemně interpretovat tak, že *C* je přesvědčen o reflexivním vztahu mezi referenty α a β , pokud *A* ví, že *C* si je vědom stejnosti reference α a β , a zároveň ví, že *B* ví, že *A* ví, že *C* si je vědom stejnosti reference α a β atd.

Mezi mluvčím *A* a posluchačem *B* může selhat vědomostní koordinace ohledně koreference α a β . I v takovém případě *A* i *B* stále rozumí obsahu věty (6), třebaže nerozumí plnému rozsahu informací (tj. jejich prag-

matickým implikacím), ke kterému použití této věty sloužilo. Stejně tak může selhat koordinace mezi A a B ohledně znalostí C při prohlášení věty (12). B se může mylně domnívat, že A ví, že B ví, že C si je vědom stejnosti reference α a β . Jedná se ovšem o pouhé nedorozumění mezi mluvčími. I v takovém případě B rozumí, co A tvrdí, třebaže plně nerozumí, co tím A myslí.

Jedinou (mylnou) alternativou by bylo prohlásit, že tito lidé nejen že nerozumí, co tím či oním ten či onen myslí, nýbrž že prostě nerozumí rovnou všem těmto větám. Následně by bylo nutné specifikovat, čemu nerozumí a proč. V případě Pierra se zřejmě jedná o jméno „Paderewski“. Pierrův problém spočívá v tom, že si myslí, že existují dvě specifická vlastní jména *Paderewski* referující ke dvěma různým osobám. Jednou z možností by bylo prohlásit, že nerozumí žádnému ze jmen, jež v rámci svého idiolektu odlišuje. Pierre však může být Paderewského dobrým kamarádem, běžně o jeho vystoupeních diskutovat s jeho rodinou a nevědět pouze to, že je rovněž politikem. Bylo by odvážné tvrdit, že k Paderewskému v těchto situacích není schopen referovat užitím jeho jména. Další možností by tak bylo tvrdit, že jméno „Paderewski“ rozumí výhradně v případě těch výskytů, kdy si myslí, že referuje k osobě, s níž je obeznámen jako muzikantem a svým kamarádem. Neexistuje ovšem důvod, proč tyto výskyty upřednostňovat před druhým typem výskytů, při nichž si myslí, že se referuje k Paderewskému-politikovi. Názorněji je to patrné v případě idiolektu Lois.

V případě Lois je opět problematické tvrdit, že je nekompetentním²⁵ uživatelem jména „Superman“ nebo „Clark Kent“. Nerozumí snad jméno „Clark Kent“, jehož užívá při referování k osobě, s níž pracuje v Daily Planet, nebo „Superman“, jež odkazuje k hrdinovi, o němž píše? Ani jedna z variant se nezdá příliš pravděpodobná. Obecně vzato není možné klást požadavek, aby byl mluvčí kompetentním uživatelem nějakého jména,

25 Sémanticky nekompetentní mluvčí se mýlí v sémantickém obsahu výrazu. V tomto případě by Lois musela referovat k někomu, kdo není Supermanem (čili Clarkem Kentem), což by také vedlo k mylnému popisu daného individua. Sémanticky kompetentní mluvčí naopak referuje ke správné osobě, i když se může mýlit v jednotlivých popisech. Inferenční dovednosti jsou pro určení sémantické kompetence nedostačující. Nikdo k výrazu nedokáže doplnit všechny inference a není možné určit, jaký soubor inferencí, ať už zahrnuje znalost identit či nikoli, by byl dostačující k tomu, abychom někoho mohli označit za sémanticky kompetentního uživatele daného výrazu.

výhradně pokud by při všech setkáních s danou osobou rozpoznal tuto osobu jakožto nositele daného jména. Jsem si jistý, že jsem mnohé ze svých přátel alespoň jednou v životě nepoznal, když šli naproti mně přes ulici, což snad ještě neznamená, že nejsem schopný referovat k nim užitím jejich jména či rozpoznat, ke komu tento výraz referuje, když ho někdo použije ve svém prohlášení. Dejme tomu, že by Jirka viděl mě a mého kamaráda Honzu stát v takové vzdálenosti od sebe, že je rozumné předpokládat, že jsme se viděli a poznali. Honza má nový účes a Jirku zajímá, co o jeho účesu soudím. Jirka se může zeptat „Co říkáš na Honzu?“, předpokládá, že jsem jej včera poznal. Jelikož jsem jej ale nepoznal, dotaz si vyložím jinak. Mohu například předpokládat, že Jirku zajímá jeho rozchod s přítelkyní. Tím, co v rámci naší komunikace selhalo, není rozpoznání výrazu či reference, nýbrž naše epistemická koordinace.²⁶ Záměrem mého kamaráda bylo odkázat mě k Honzovi v situaci, kdy jsem jej nepoznal, omyl tedy nastal na úrovni porozumění obsahu použité věty, nýbrž na úrovni toho, co má tento obsah evokovat. Přestože jsem Honzu v dané situaci nepoznal, zůstávám i nadále kompetentní referovat k němu pomocí jeho jména.

Rozpoznání identity reference tedy není podmínkou sémantické kompetence mluvčích. Lois je schopna mluvit o Clarku Kentovi užitím jména „Clark Kent“ a stejně tak je schopna mluvit o Supermanovi užitím jména „Superman“. Je si vědoma, ke komu tato jména referují a nikdy jejich užitím nerefereje k někomu, komu nepřínáleží. Přesto si není vědoma toho, že referují ke stejné osobě, takže při každém setkání s touto osobou nerozpoznává danou osobu jako nositele jednoho z jmen. Pokud tedy rozpoznání identity reference není podmínkou sémantické kompetence, pak informace o této identitě nemůže být součástí sémantického obsahu vět. Nezahrnuje-li však propozice informaci o identitě dvou referujících výrazů, nemůže zahrnovat ani informaci o reflexivitě vztahu, jenž má mezi danými výskyty reference platit. Konečně nezahrnuje-li propozice informaci o reflexivitě vztahu, není z ní možno vyvodit, že o daném objektu platí reflexivní vlastnost. Z Loisina přesvědčení o (3a)–(3c) tak není možné vyvodit její přesvědčení o (3d).

26 Salmon (2012, s. 432).

Existenční generalizace

Lois není možné připsat přesvědčení o (3d), protože by si to vyžádalo reflexivní přístup. (3d) by mohl odvodit pouze někdo, kdo má vedle informace ohledně Loisina přesvědčení o (3a) ještě informace o Loisině vědomí identity Supermana a Clarka Kenta. Pouze pokud by někdo věděl, že Lois ví, že Superman a Clark Kent jsou totožnou osobou, mohl by odvodit její přesvědčení o (3d). Bez této znalosti je však odvození (3d) z (3a) neoprávněné a s ním i důsledek (3e). Přesto však Loisinu přesvědčení, že Superman je silnější než Clark Kent, zahrnuje pouze jediné individuum, a pokud na toto její přesvědčení vztáhneme zákon (P), pak musíme uznat, že je přesvědčena o (13).

(13) $[\exists x](x \text{ je silnější než } x)$

Jedná se o totožný problém jako ten, o němž píše Church. Král Jiří i Lois jsou přesvědčeni o existenci někoho, o němž platí vztah, jenž nemůže platit pro jedno a totéž individuum. Ani Jiří, ani Lois nemají za to, že se jejich přesvědčení vztahuje k totožné osobě, přitom jej však oba vztahují ke konkrétním osobám. Lois je přesvědčena, že někdo, kdo má nadpřemyselské schopnosti, je silnější než někdo, koho zná z Daily Planet, přestože ve skutečnosti jakožto dvě různé osoby, k nimž se její přesvědčení vztahuje, neexistují. (13) je existenční generalizací singulární propozice zahrnující pouze jedno individuum, a reprezentuje tak obsah Loisinu přesvědčení.

Předpoklad, že nikdo nemůže být přesvědčen, že x je silnější než x nebo že x je odlišné od x , není apriorně platným zákonem. Nutnou podmínkou platnosti této formulace by bylo, že agent rozpoznává objekt x v rámci těchto vztahů jako identické individuum. Zahrnutím informace o identitě do obsahu propozice by skutečně bylo možné vyvodit, že nikdo nemůže být přesvědčen o propozici ve formě (14).

(14) $[\exists x](\lambda x')(x' \text{ je silnější než } x')x$

Kdyby král Jiří nebo Lois získali nové informace ohledně identity, jejich přesvědčení by se jistě změnila.

Stejně jako je Lois přesvědčena, že Superman je silnější než Clark Kent, je zároveň přesvědčena, že Superman není Clark Kent. Tím je vzhledem k (P) přesvědčena o existenční propozici (15).

(15) $[\exists x](x \neq x)$

Dejme tomu, že by někdo pozoroval Supermana, snažícího se přesvědčit Lois, že není Supermanem. Kdyby byl ve své snaze úspěšný, pozorovatel by mohl prohlásit „Superman přesvědčil Lois, že (on) není Superman“. Tato věta o Supermanovi by byla pravdivá, čímž by zároveň existoval někdo, o kom by platila, což lze reprezentovat v (16).

(16) $[\exists x](x \text{ přesvědčil Lois, že } x \neq x)$

Pozorovatel by tak zjistil, že existuje konkrétní jedinec, o němž je Lois přesvědčena, že není tím, kým je.²⁷ V určitém smyslu by tak Lois skutečně byla přesvědčena, že nějaké x není x .

Singulární propozice (3a) ani její existenční generalizace (13) nepodávají celkový obrázek ohledně přesvědčení Lois. Určují pouze vnější konstituenty, jichž se její přesvědčení týká, jež stejně tak mohou být předmětem postojů druhých lidí a mohou být druhými uchopeny za účelem porozumění, k čemu se přesvědčení Lois vztahuje. Pokud uznáme, že stavy přesvědčení disponují stejnou strukturou jako obsahy přesvědčení,²⁸ pak není možno vyjádřit stav přesvědčení korespondující s Loisiným přesvědčením o (13) pomocí stejné reprezentace. Lois disponuje dvěma způsoby, jimiž je pro ni Superman reprezentován a spojuje si s nimi dva různé rozsahy informací. Pro reprezentaci stavu přesvědčení korespondujícího s přesvědčením o (13) by bylo zapotřebí dvou odlišných proměnných, jež by vázaly její odlišné způsoby uvažování o Supermanovi.

²⁷ Soames (2015, s. 158–159).

²⁸ Obsahem přesvědčení je propozice, o které je někdo přesvědčen. Stav přesvědčení je stav, v kterém se při přesvědčení někdo nachází. Stav přesvědčení zahrnuje dodatečné informace, které nejsou součástí singulární propozice (popisy, index). Perry (1977, s. 494).

Závěr

Článek se věnoval tomu, jak vyřešit Churchovu a Richardovu námitku ohledně analýzy propozičních postojů v teoretickém rámci přímé reference. Churchův-Richardův argument vychází z přesvědčení, které nezahrnuje reflexivitu, a vyvozuje z něj přesvědčení, které reflexivitu zahrnuje. Úkolem tedy bylo zjistit, jak se v propozici získává informace o reflexivitě a v kterém bodě tak paradoxní vyvození selhává.

Reflexivita se běžně považuje za sémantický typ informace. První část práce se věnovala vyvrácení této představy. Informace o reflexivitě podle předložené analýzy není sémantickou vlastností vztahů ani anaforických zájmen. Nejprve se argumentovalo, že v prohlášeních obsahujících vztahy a anaforická zájmena reflexivita není součástí sémantické informace. Následně bylo představeno řešení, podle kterého je reflexivita u těchto vět součástí pouze pragmaticky podmíněné informace, která vychází ze znalostí mluvčích o sobě navzájem.

Když se toto řešení vztáhlo na Richardův paradox, tak se ukázalo, že Richardovo vyvození je sice neplatné, ale v jiném bodě, než v kterém si mysleli jiní zastánci přímé reference. Salmon a Soames argumentovali, že propozice, že Superman je silnější než Superman, a propozice, že Superman je silnější než on sám, se liší v tom, že první z nich reflexivitu neobsahuje, zatímco druhá ano. Podle předložené analýzy ale reflexivitu neobsahuje ani jedna z těchto propozic, a tak je vyvození druhé propozice z první stále platné. Neplatné je teprve následující vyvození, kdy se Supermanovi přisuzuje vlastnost, že je silnější, než jak silný skutečně je.

Argument proti přímé referenci, který vychází z Churchova paradoxu, je rovněž neplatný. Východiskem argumentu je fakt, že nikdo není přesvědčen o existenci objektu, který by byl odlišný od sebe sama. Ukázalo se, že propozice, že existuje x takové, že $x \neq x$, není vhodnou reprezentací tohoto východiska. Existují totiž případy, kdy aktéři mohou být o této propozici přesvědčeni, což ještě neznamená, že by uznávali existenci objektu, který by měl vlastnost, že je odlišný od sebe sama.

Literatura

- Church, A. (1982): „A remark concerning Quine’s paradox about modality.“ *Analisis filosófico* 2 (2): 25–32; přetištěno v *Propositions and attitudes*, eds. N. Salmon & S. Soames, Oxford university press, Oxford, 1988, s. 58–65.
- Frege, G. (1892): „Über Sinn und Bedeutung.“ *Zeitschrift für Philosophie und philosophische kritik* 100 (1): 25–50.
- Leibniz, G. W. (1686): „Discourse on Metaphysics.“ In *G. W. Leibniz: Philosophical Papers and Letters*, ed. L. Loemker, Reidel, Dordrecht, 1969, s. 303–330.
- Kripke, S. (1979): „A puzzle about belief.“ In *Meaning and use*, ed. A. Margilit, Reidel, Dordrecht, 1999, s. 239–283.
- Perry, J. (1977): „Frege on demonstratives.“ *The philosophical review* 86 (4): 474–497.
- Richard, M. (1983): „Direct reference and ascriptions of belief.“ *Journal of philosophical logic* 12 (4): 425–447; přetištěno v *Context and the attitudes*, ed. M. Richard, Oxford university press, Oxford, 2013, s. 26–47.
- Richard, M. (1990): *Propositional attitudes: An essay on thoughts and how we ascribe them*. Cambridge university press, Cambridge.
- Richard, M. (1997): „Propositional attitudes.“ In *A companion to the philosophy of language*, ed. B. Hale, C. Wright & A. Miller, Blackwell, Oxford, 2017, s. 324–356.
- Salmon, N. (1986): „Reflexivity.“ *Notre Dame journal of formal logic* 27 (3): 401–429; přetištěno v *Content, cognition, and communication*, ed. N. Salmon, Oxford university press, Oxford, 2007, s. 32–57.
- Salmon, N. (1992): „Reflections on reflexivity.“ *Linguistics and philosophy* 15 (1): 53–63; přetištěno v *Content, cognition, and communication*, ed. N. Salmon, Oxford university press, Oxford, 2007, s. 58–66.
- Soames, S. (1985): „Lost innocence.“ *Linguistics and philosophy* 8 (1): 59–71.

Soames, S. (1994): „Attitudes and anaphora.“ *Philosophical perspectives* 8: 251–272; přetištěno v *Philosophical essays, 2: The philosophical significance of language*, ed. S. Soames, Princeton university press, Princeton, 2009, s. 111–135.

Soames, S. (2015): *Rethinking language, mind, and meaning*. Princeton university press, Princeton.

Abstract

Reflexivity and propositional attitudes

The theory of direct reference leads to some problematic consequences about propositional attitudes. One paradoxical consequence is shown on the example of Lois Lane. Given her belief that Superman is stronger than Clark Kent, we are apparently forced to deduce that she also believes that there is someone with a property that he is stronger than he actually is. No rational being would commit to the belief that there is someone with this impossible property. The deduction is based on the notion of *reflexivity*. This paper explores solutions of how reflexivity is obtained within semantic content and how the paradox can be resolved. First two contemporary options are analyzed: Reflexivity is either a semantic property of relations or that of anaphorical pronouns. Both options are refuted on the basis that reflexivity cannot be part of semantic information in sentences describing relations between the contents of two coreferential singular terms because knowledge about identity of reference is not a condition of semantic competence. Afterwards a solution is presented according to which reflexivity is only part of the pragmatic content. In the given examples, reflexivity can only be obtained with speakers coordinating their shared knowledge.

Key words: propositional attitudes, reflexivity, identity, direct reference

Šmiřák, T. (2019): „Reflexivita a propoziční postoje.“ *Filosofie dnes* 11 (2): 34–56. Dostupné z www.filosofiednes.ff.uhk.cz

Collective Belief and the Intentional Strategy¹

David Kocourek

*Filozofická fakulta
Univerzita Hradec Králové
Rokitanského 62, 500 03 Hradec Králové
david.kocourek@uhk.cz*

What do we mean when we say that some group believes something? Do we simply mean that all the members of the group believe it, or are we acknowledging the existence of some kind of group agent? According to Margaret Gilbert, talk about group mental states refers to the specific kind of agreements she calls joint commitments – that is, to collectively believe something means to be committed with others to believe it. In my article, I will first present Gilbert’s approach in more detail but will ultimately show that this approach is problematic and will refute it. I will briefly consider the most common solution to the problems Gilbert’s account faces, which lies in replacing collective beliefs with acceptances, but I will show that this solution will not do either. The solution I will then present will be based on Daniel Dennett’s intentional strategy, which is a method of interpreting the things around us and predicting their behaviour by treating them as rational agents with relevant intentions. I will try to show that all the problematic cases of collective belief can be explained by applying the intentional strategy to the groups in question.

Key words: Collective belief, Daniel Dennett, Intentional strategy, Joint Commitment, Margaret Gilbert

What do we mean when we utter sentences such as: “Christians believe in the afterlife”, “The government wants to raise the minimum wage” or

¹ Work on this paper was supported by the joint Lead-Agency research grant between the Austrian Science Foundation (FWF) and the Czech Science Foundation (GAČR), Inferentialism and Collective Intentionality, GF17-33808L.

“Students like to party on Friday nights”? On one hand we could surely interpret these sentences as saying that every member of the group in question (or at least an overwhelming majority of its members) has the property in question – that every Christian believes in the afterlife, every member of the government wants to raise the minimum wage, and every student likes to party on Friday nights. This interpretation is often called a *simple summative account* – but we would probably feel that there might be something wrong with this interpretation, because while it seems quite plausible to claim that every Christian believes in the afterlife, it seems less plausible to say that every student likes to party on Friday nights – and it seems very implausible to think that all members of any government would ever agree on anything. Yet, we do often use sentences like these, and we generally do not consider them to be plainly false. But how then should we interpret them? British philosopher Margaret Gilbert has proposed the so-called *plural subject account* of collective belief, according to which it is possible for social groups to have beliefs, emotions and mental states in general that are not necessarily shared by all their members. In the following article I will first present in detail Gilbert’s approach to collective belief, then I will attempt to point out some of its problematic implications and offer a different solution to the phenomenon in question.

The key concept in Gilbert’s plural subject account is *joint commitment*. What is joint commitment? Gilbert usually explains it by contrasting joint commitment with a simple *personal commitment*: we create personal commitments every time we make a decision to do something. For example, when I decide to, say, write a grant proposal tonight, this decision will in itself not only give me a reason to perform this action (“I am going to write a grant proposal tonight because I have decided to do so”), but it will also create an obligation to act on that decision (“I must write a grant proposal tonight because I have decided to do so”). Therefore, should I forget about my decision and go to the pub instead, I could, when I have remembered that I have made a personal commitment to write a grant proposal tonight, say something like “Oh no, I should have been writing” – that is, I could recognize my current action as wrong, since this action contradicts my previous personal commitment. Of course,

the obligations created by personal commitments are always only to myself, and I can easily change them just by changing my mind – but this will not be the case once we get to joint commitments.²

Joint commitments arise when two or more people decide to do something together. The most important thing about joint commitments is that they are not the sum of two (or more) separate commitments to do the respective parts of a collective task – rather, a joint commitment is one commitment binding a number of people at once. If, for example, you and I have a joint commitment to go see a movie together, it is not just a combination of my commitment to go see a movie and your commitment to go see a movie. If I manage to deliver on my part and get to the cinema on time, but you do not, I would not say that I have succeeded in our task and you have not – no, if you do not show up, then we both fail. The formula that Gilbert herself uses to define joint commitment is this: *persons a, b, c, etc. are jointly committed to do x as a body*³ – meaning all the individual members of the group are committed to successfully carrying out action *x* as if one metaphorical person was doing *x*. Sure, they will most likely divide the work somehow – one person will do this, another that – but they can only all succeed or all fail, there is only one commitment that they all share together.

This might, at least at first sight, seem like an interesting approach to the problem of collective action, but Gilbert also uses joint commitment when she explains collective beliefs and other collective mental states.⁴ The formula here is very similar to the one we saw above: *a group (population, collective, etc.) A believes that p if all the members of A are jointly committed to believe that p as a body*.⁵ And this is where the problems with Gilbert's account start, because while it makes sense to say that members of a group agree to commit themselves to carry out an action, it is not easy to see how they could commit themselves to feel or believe something. We generally do not think that an individual could do such a thing, so how come group members can? To get a better idea of how Gil-

2 Gilbert (2002a, p. 31), Gilbert (2004, p. 173).

3 Gilbert (2002b, p. 138).

4 For example, Gilbert (2002c) uses the plural subject account to analyse collective guilt feeling.

5 Gilbert (2004, p. 173).

bert thinks this works, we should look at one specific example she often presents: imagine a poetry club that regularly reads and discusses poetry. At the end of every meeting the members of the club try to formulate their opinions on the poem they were reading that day and attempt to formulate some collective assessment of the poem in question. Now, let us say that after reading a poem one day, one of the members opens the discussion by claiming that she loves the poem. Another member joins her, and the next one says she kind of likes it too. The rest of the members maybe nod in agreement or at least do not explicitly disagree. Now we could – according to Gilbert – say that the poetry club likes the poem in question, that the club believes the poem is good, simply because all members of the club have agreed on this position. But despite this, it is quite possible that none of the members actually liked the poem personally! Maybe the first member who claimed to like it just liked the author and wanted to support her, even though she did not actually like this particular poem. Maybe the second supporter of the poem did not really like it either, she just wanted to impress (for whatever reason unrelated to poetry) the first member by agreeing with her. And maybe the rest of the group just did not feel like arguing, and so they simply agreed with the first opinion suggested. In any case, we now have a collective belief that the poem is good (i.e. *a group of people jointly committed to belief as a body that the poem is good*) without any such personal belief.⁶

From this it could look like the group members can “decide” to believe something, because it is not they who in the end believe it, it is a group agent, a plural subject, that they created together – and since they have created this subject, they get to decide its beliefs. But this cannot be right – Gilbert does not think that collectives could create some actual new subject, a new “group mind” ontologically different from the individual members. All the talk about emulating a body is more of a metaphor: the individual members merely act as if they composed an agent.⁷ But if this is the case, if the belief “The poem is good” from our example is not a simple belief of some collective subject but a collective belief of individual members of the poetry club, how is it possible that they can

6 See e.g. Gilbert (1987, pp. 190–194), Gilbert (2004, pp. 168–169).

7 Gilbert (2014, p. 9).

collectively believe that the poem is good and individually believe that it is not good at the same time? Gilbert is willing to accept this problematic conclusion, claiming that this is simply how we use collective beliefs in our everyday talk – and if traditional epistemological understanding of belief is different, then so much the worse for epistemology⁸ – but it probably comes as no surprise that many authors do not agree with this conclusion.⁹

The most common objection against this approach is the claim that the collective belief described by Gilbert is not actually a belief but rather an acceptance. Several differences between belief and acceptance have been suggested.¹⁰ In my opinion the most important are the following:

- (1) Acceptance is voluntary, belief is not. As was already mentioned, groups can, according to Gilbert, decide what they will take for their “belief”. This is not the case for individuals. What individuals can do, however, is to decide to accept something. This happens when we say things like: “Let’s imagine for now that she did it” or “Ok, I will agree on this but it’s just so we can move on”; that is, when we accept a proposition e.g. for the sake of a discussion or to compromise. We do not really believe it, but we merely act as if we believed it for some pragmatic reason. Which gets us to the second point:
- (2) Belief necessarily aims at truth, acceptance does not. As we saw in the poetry club example, the group can decide to “believe” something for pragmatic reasons rather than for its truth value. Gilbert presents more examples like this: take a campus improvement committee that “believes” that a new café could improve the campus. But the reason for this “belief” is that the Chancellor wants the new café on campus and

8 Gilbert (2004, p. 178).

9 For various critiques of Gilbert’s account see e.g. Hakli (2007), Mathiesen (2006), Sheehy (2006, primarily chapter two), Tuomela (2000). See also Tollefsen (2015, primarily chapter one) for a brief comparison of Gilbert’s account with some other possible approaches.

10 But see Schmitt (2014) for some interesting objections against these differences.

nobody in the committee dares to disagree with the Chancellor.¹¹ Again, in the individual case we would not call this a belief – on the contrary, we can quite easily imagine some individual member of this committee saying something like “I don’t think this campus needs a new café, but I can’t say that in front of the Chancellor”. One can only believe something if she considers it true (and vice versa, one can only consider something true if she believes it).¹² But this is not the case for acceptance.

- (3) Belief may come in degrees, but acceptance does not. When talking about belief, we often claim that we consider something possible, probable, almost certain, etc. That is, we can believe or doubt something just to a certain degree. This is usually not the case with acceptance – if we understand acceptance as voluntary action taken for pragmatic reasons, there is no need for doubts. We can accept (or refuse) a “belief” – full stop. In the cases of group beliefs described by Gilbert, it looks like the groups always decide to believe absolutely – the poem is good; the café is needed – there are no “buts” or “maybes” in these cases.

What would it mean to replace belief with acceptance? Gilbert would say not that much – we can have semantic discussions about whether to call the phenomenon in question belief or acceptance all we want, but the important thing – the fact that there are some collective mental states – still remains true.¹³ This would indeed be correct for accounts that refuse group beliefs but allow group acceptances,¹⁴ but I think that once we substitute acceptance (in the sense described here) for belief, we can also get rid of the “collective” part of the equation. To explain

¹¹ Gilbert (2002b, p. 152).

¹² One could point to cases where we say things like: “I can’t believe it’s raining on my wedding day” or “I know my sister is dead, I just can’t believe it.” I consider these cases metaphorical, describing a huge disappointment or sadness about the state of affairs in question. If pressed, I am sure that the speaker of such sentences would admit that she actually does believe that it is raining or that her sister is dead.

¹³ Gilbert (2002b, p. 160).

¹⁴ This is the case in e.g. Tuomela (2000).

this, we need to return to the beginning of this article for a moment. There I said that some apparent group beliefs can be explained as only referring to individual beliefs of all (or the vast majority of) the members – the example was that the sentence “Christians believe in the afterlife” might literally mean that every single Christian indeed does believe in the afterlife. We have called this approach a simple summative account. We have also said that this approach cannot explain all the cases of collective belief – in Gilbert’s example of the poetry club, the statement “The club believes that the poem is good” cannot possibly be explained by the simple summative account since, as we saw, no member of the club actually believed that the poem was good individually. But does the same hold for acceptance? It seems to me that it does not, because we can say about every member of the club individually that she accepted (that is – according to our specification of acceptance – voluntarily decided for pragmatic reasons to treat as definitely true) the proposition “The poem is good”. So, let us now consider whether any case of so-called group “belief” might actually be just a case of individual acceptance.

In both of Gilbert’s examples presented here, the view that collective belief is just individual acceptance holds. The case of the campus improvement committee is very similar to the case of the poetry club (we can similarly say that all the members of the committee have individually, voluntarily accepted, out of fear of the chancellor, that the campus without a doubt needs a new café). Even the examples we started this text with might fit this interpretation. Here, the claim about Christians and the claim about students are interpretable simply by the simple summative account (they might still, of course, be false; it might turn out for example that many students actually do not like to party on Friday nights), and the sentence “The government wants to raise the minimum wage” could be interpreted the same way as the cases of the poetry club or the campus improvement committee. Most of the members might not actually believe that the minimum wage should be raised, but they have accepted this position as a part of some negotiation. But there are other examples that are more complicated than this. We will now consider some of these:

First example: take elections or any other kind of large-scale voting. When the voting is done, commentators might say things like “The citizens of this particular country believe that the elected party can solve the ongoing crisis” or – to give a more specific example – “The Britons wanted to leave the European Union”. Most of the time in such elections it is not the case that a big enough majority voted for the winning party (or for leaving the EU) to explain such cases with the simple summative account – but neither does there seem to be any kind of joint commitment to accept (and definitely not to believe) the outcome of the election – as is shown by occasional after-election demonstrations and frustrations of the losing sides.

Second example: suppose that a government is deciding whether to invest more money into the army. The decision to do this might be based on the government’s belief on whether or not the country is under threat of an imminent attack. Imagine now that the government is divided: some of its members strongly believe that there is no threat of such an attack, others are absolutely certain that the attack is bound to happen. It is possible then that the government decides to invest some sum of money into its military because it has a reason to “believe that the attack is somewhat probable”. In this example, not only do none of the individual members believe or accept that the attack is “somewhat probable” (they collectively agreed on an amount of money they will spend on the army but all still hold to their previous beliefs that the attack is either very or not at all probable), but also this collective belief of the government comes in degrees – *somewhat probable* – and therefore cannot be a case of acceptance as defined here.¹⁵

15 Cf. Mathiesen (2006, p. 162).

Third example: imagine that the chief of the police force gets (from an outside source) some important information about a crime going on. Further, imagine that the perpetrators of this crime at the same time get the information that the chief of the police has been informed about their activity. They can now say something like “Oh no, the police know about us”. This claim, it seems to me, would be quite justified since it is reasonable to expect that the police will take some action. And yet the vast majority of the police officers do not know, nor believe, nor accept that the criminal activity is happening – the chief is the only one who got the information.¹⁶

Not only can these three cases not be explained as referring to the individual acceptances of all the group members, they also cannot be explained by Gilbert’s original plural subject account of collective belief – not even if we substitute collective belief with collective acceptance! And Gilbert cannot just get rid of them by saying that they are merely illusory cases of collective belief, because her own argument for treating collective beliefs as genuine beliefs and not just metaphors relies on the claim that we should respect what we call belief in natural language. So, to justify Gilbert’s approach to collective belief, we would either have to show how to explain the examples above using the plural subject account or suggest a way to distinguish metaphorical collective beliefs from real ones (and such a way would have to eliminate our problematic examples while still allowing Gilbert’s examples and would have to do this without referencing the plural subject account itself to avoid circularity). But I do not think we should go through all this trouble. In the rest of this paper, I will rather try to propose an analysis of the phenomenon that could explain all the examples we have presented so far without needing to invoke any form of plural subjects or collective mental states.

In short, the solution I want to propose is to treat any group that could be said to have mental states independent of the mental states of its members (that is, the cases that could not be easily interpreted by the simple summative account – though I will later argue that we certainly

¹⁶ Cf. Hakli (2007, p. 255).

should use the simple summative account more often than Gilbert seems to suggest) as an intentional system – as an agent whose behaviour can be predicted or interpreted by the so-called intentional strategy proposed by Daniel Dennett.¹⁷ The view I want to defend is that groups, of course, cannot have mental states in the same sense that we generally think humans do – but it is sometimes nevertheless reasonable to treat them as if they did. In what follows, I will first explain in more detail what Dennett means by using the intentional strategy (or, in other words, *assuming an intentional stance*), and then I will try to show how this approach solves the problematic examples presented so far.¹⁸

We are, according to Dennett, “assuming a stance” when we aim to understand the world around us, to interpret it, to predict its future states and actions. But the aforementioned intentional stance is not the only stance that we can assume – there is also what Dennett calls the *physical stance* and the *design stance*. You assume a physical stance when you are trying to predict the behaviour of an object based on the physical constitution of the object in question and the natural laws it obeys. You can assume this stance to easily predict, for example, what a rock thrown out of the window will do: it will fall down (and possibly break something at the end of its fall) due to its weight and the law of gravity. The physical stance can be applied quite universally, but only as long as you know the physical laws binding the future actions you are trying to predict. You can often find yourself in a situation when you want to know how something will act without knowing enough about its inner composition or about the physical laws that govern its behaviour. In such cases you will resort to one of the other two stances. You assume a design stance when you attempt to predict the behaviour of an object based on your knowledge of the design or function of that object. For example, if you guess what will happen when you turn the ignition key in your car, you would probably not think about how the battery, the ignition system or the engine work. You simply do not know enough about the car’s engineering to predict its behaviour from its physical proper-

¹⁷ E.g. Dennett (1971), Dennett (1981).

¹⁸ Dennett did not use the intentional strategy to interpret group behaviour himself – as far as I know, the first one to apply the intentional strategy to groups was Tollefsen (2002), who uses it to interpret intentionality of organizations.

ties (unless you are a professional car mechanic or a car enthusiast, who might do just that in his everyday work). What you most likely do know, however, is how cars are generally designed to work. You know that unless something is broken, a car's engine will start when the ignition key is turned – and therefore, assuming a design stance, you would make exactly that prediction. Last but not least, there is the intentional stance. You assume this stance every time you try to understand the behaviour of something based on its intentions – its beliefs, wishes, desires and other mental states you presume it to have. You of course assume this stance towards other human beings: You can predict that your friend will take an umbrella when she leaves her home not because you try to understand the natural laws binding her every move, nor because you consider her somehow designed to carry an umbrella with her when she goes out, but because you assume that she has a belief that it is raining outside and also a desire not to get wet. But you can, according to Dennett, resort to the intentional strategy in other situations as well. We often use it when we talk about animals (even though to some people ascribing verbalized beliefs or other mental states to animals might be philosophically uncomfortable¹⁹). We can even assume an intentional stance towards machines or programs – for example, the best strategy to predict the moves of a chess-playing program is to treat it as an agent that *wants* to win a game (and also *believes* that you want to win the game, *knows* how the chess pieces move, *intends* to play a certain strategy, etc.). To say that something has mental states means nothing more than saying that you assume an intentional stance towards it, that it is interpretable by an intentional strategy. In this sense, according to Dennett, even animals or objects can have mental states.

It would seem that the physical stance is the most fundamental, the most precise description of reality. Is it not, after all, a goal of many branches of cognitive science to find out how human brains *really* work – to be able to describe thoughts, wishes or intentions just as synaptic transmissions between neurons? Well, that depends on what we are interested in. If we are trying to discover natural laws that human brains

19 See Rowlands (2012, primarily chapter two) for an interesting approach to the problem of ascribing intentional states and propositional content to animals.

abide by, then, sure, we must treat brains as systems interpretable by assuming a physical stance. But if we are interested in how human society works, what human reasoning and decision making actually look like in society, then ignoring the intentional strategy would mean missing a crucial part of the process. I want to argue now that we can assume these three stances not only towards objects, animals or individuals, but also towards groups, and that – as well as in the case of describing human thoughts – while every group is in the end reducible to a more basic level (in this case to the thoughts of its individual members, which could then, of course, be reduced even further to the processes in their brains), we really do use the intentional strategy when talking about groups on a daily basis – and therefore ignoring it would mean ignoring an important part of how human society works.

On this account, assuming a physical stance towards a group would be, as I have already indicated, analogical to the simple summative account approach – we would be predicting the behaviour of a group just by predicting the actions or mental states of all the individual members. When we know all relevant information about all relevant individuals, we can surely predict or explain their behaviour without endorsing the group at all. But just as we often do not know all the laws and mechanisms necessary to assume the physical stance towards objects around us, we often do not have enough information about all the individuals we are interested in. And this is when we resort to intentional strategy – this is why we talk about groups as if they had thoughts and beliefs and desires. Obviously, the government does not *want* to raise the minimum wage in the same sense that I *want* to do something (unlike me, government does not have *mind*), but if we do not know the motivations and intentions of all individual members of the government, treating the whole group as one agent with its own motivations and intentions – that is, assuming an intentional stance towards it – might just be the next best strategy to understand what is going on. And finally, even though we have not really talked about this aspect of group behaviour so far, we could surely also assume a design stance towards some specific kinds of groups: for example, we could predict that if the police encounter criminal activity, they will try to

stop it rather than join in, since it is a purpose of the police to stop crime – police forces are *designed* to do just that.²⁰

Armed with the intentional strategy, we can now try to resolve all the problematic cases that led Gilbert and others to believe that postulating collective mental states is necessary. Let us start with simple situations of small collectives that do not form any kind of social group or institution. Gilbert, for example, presents a case of two friends, Roz and Mark, who believe that defensive war is justified.²¹ If Roz says: “We think that defensive war is justified”, it can mean (in Gilbert’s analysis) one of two things: first, it might be a simple summative account statement claiming that Roz believes that defensive war is justified and Mark just happens to have the same opinion. Second, it could be a case of collective belief – that is, it might mean that Roz and Mark have formed a joint commitment to believe as a body that defensive war is justified. Gilbert’s main argument for the second possible interpretation is that, should Mark now all of a sudden say that he does not believe that defensive war is justified (maybe he just said that before because he wanted Roz to like him), Roz will feel that Mark has offended her somehow (and he will understand that she has the right to feel that way). According to Gilbert, the reason for this feeling has its roots exactly in the violation of their joint commitment, in betraying their collective belief. I disagree – it se-

20 It needs to be said here that my interpretation of Dennett’s intentional strategy is very *instrumentalist*, that is, I am treating the intentional stance mainly as a tool, a useful metaphor. But that is quite different from how Dennett himself intended to understand the strategy. His own approach is in fact *interpretationist*. According to him, there is really nothing more to intentionality than being interpretable by the intentional strategy – and therefore there is no real difference between the intentionality of individual humans and intentionality of machines or groups. Tollefsen (following Dennett) characterizes this position as “mild” or “intermediate” realism: group beliefs (just as any other beliefs) are real, not just metaphors, but their ontological status is different from objects like chairs or tables – they are more like centers of gravity (Tollefsen, 200, pp. 405–406). It seems to me though that there are some important distinctions between the intentionality of groups and the intentionality of individuals that should not be ignored; furthermore, I think that instrumentalist interpretations of the intentional strategy may be easier to understand for the readers that are less familiar with Dennett’s work – and that is why I prefer the instrumentalist approach. But in the end, my main point, that is the claim that the intentional strategy can handle group beliefs better than Gilbert’s joint commitment account, remains the same for both instrumentalist and interpretationist uses of the intentional strategy, and I will therefore leave it to the reader to decide between instrumentalism and interpretationism.

21 Gilbert (2002b, pp. 136–137).

ems to me that there is a much simpler explanation. Roz feels offended simply because Mark lied to her.²² When Roz said “We believe that defensive war is justified”, she did not utter a true sentence about an existing collective belief that Mark (unbeknownst to her) has violated (which is what is happening in Gilbert’s analysis of the situation). Rather, she uttered a false sentence about her and Mark’s mental states – and that sentence surely must have been false, because it was formed on the basis of a false premise (Mark’s deceiving claim that he believes that defensive war is justified). This interpretation, it seems to me, provides satisfying enough answers without the need to postulate new entities like collective beliefs or joint commitments. A simple summative account is all we need here, and therefore, respecting the old and sacred rule of Occam’s razor, it should be preferred.

Gilbert’s other examples – the poetry club and the campus committee – will be more interesting. On the one hand, there might arise a situation in which the analysis would be similar to the case of Roz and Mark. For example, if a member of the poetry club, after the session described before, says something like: “Fine, we all believe that the poem is good”, she is quite clearly claiming that every member of the club does believe so individually – and is wrong in saying that, again, because her claim is based on the lies of other members about their mental states. On the other hand, though, in different situations the ascriptions of mental states might be justified. If the students discussing the new café on campus say things like “Well, the new café will be built because the campus improvement committee wants one”, it would be reasonable to interpret this claim as a case of assuming the intentional stance towards a group that these students cannot interpret otherwise. Notice I say that assuming the intentional stance is *reasonable* rather than that the sentence is true. Talking about the truth or falsity of sentences ge-

²² Or has changed his mind without letting her know. This will in most cases depend on the time that has passed between Mark’s utterance that he believes that defensive war is justified and his denial of this claim. Had it happened within one day, Roz would surely feel that Mark has lied to her. But if a whole year has passed between the two conversations (with no relevant discussion on the topic in the meantime), Roz might just think that Mark has changed his mind. I suppose that her anger towards him would be much less intense in this case – exactly in accordance with the claim that the reason for her anger is Mark’s lying, not violating some sort of binding commitment (which would be as relevant years after as it was the day it was formed).

nerated via the intentional strategy can generally be quite tricky – not only in the case of groups, but even in cases the strategy was originally intended for (are the sentences “The chess program wants to take your queen” or “The cat thinks she owns the place” true or false?). This is not a problem of the intentional strategy though, it is a problem of any sentence in natural language that does not directly describe some state of affairs in the world. Think about sentences like “It’s raining cats and dogs” or “Some unicorns have wings”. Literally, these sentences are, of course, false, but if we understand what role they play in their specific language games, they can sometimes be considered true (in the first case if it is raining heavily when the sentence is uttered; in the second case if it is claimed in the context of some fantasy or mythical story in which there exists a creature that looks like a horse with a horn and wings). Similarly, the sentence “The campus improvement committee wants to build a new café” can be considered true in a discussion in which all the speakers understand that they are treating the committee as an intentional agent for the sake of simplicity or for a better understanding of its actions. But this context can be easily overthrown and the sentence made false if, for example, someone were to say: “They actually don’t even want it, they are just afraid of the chancellor”. By saying this, the intentional stance assumed in the discussion so far is replaced by a physical stance (which is, as we already saw, equivalent to what Gilbert calls a simple summative account and generally describes reality more precisely than the intentional or design stance). Thus, the sentence “The committee wanted a new café” (in this new interpretation meaning that every member of the committee wanted it) can no longer be considered true.

Last, let us take a look at the suggested counterexamples to treating collective beliefs as acceptances. These are all, quite obviously, cases of assuming an intentional stance towards specific social groups. The case of large-scale voting is indeed quite common in our everyday talk, but it is also the most problematic of these examples, since it can very often generate misleading or even false predictions and interpretations. We can again turn to the Brexit example: claiming (immediately after the Brexit referendum) that Britons wanted to leave the EU would probably

be acceptable, but it surely would not be a very precise description of the political situation in the post-Brexit United Kingdom. The second example – the problem of governmental discussion about financing its military – presents a much better case of utilization of the intentional strategy. While in the case of large-scale voting it is usually not difficult to find the actual percentages of votes for all the different sides (and therefore the intentional strategy is not really necessary and a more precise description should be preferred), in the case of the governmental discussion we can imagine that not many people actually know how the conclusion was reached. These people are therefore not usually in the position to say that the final military budget was a compromise between two sides with extremely different opinions about the probability of an attack. Rather, somebody with no information about the discussion might just say “The government decided to invest more money in the military because it believed that the attack was quite probable” – thus providing a good enough explanation of the government’s action without the need to know the mental states of the individual members of the government. The final example moves us from explanations to predictions. The criminals from this (highly hypothetical) situation do not know how the police typically handle such situations or how the chief of police will react to the information he got. And even if they did, it would still be quite complicated to guess and describe what is actually happening at the police station, what information and orders the chief gives to his subordinates, and how this will affect the beliefs of those individual subordinates. But this is not necessary. Assuming an intentional stance towards the police force and treating it as an agent who has a goal to fight crime and who just obtained information about a crime going on will give us a solid prediction of the future actions of this agent, which is all the criminals need (they would probably come to a similar conclusion if they assumed a design stance and treated the police force as a “machine” designed to react to criminal activity in a specific way – the fact that there may be grey areas between individual stances should not surprise us).

To briefly and clearly summarize the argument I have presented in this paper: in our everyday talk about collective mental states, we do not,

as Margaret Gilbert claims, refer to some specific sort of collective agreements, which Gilbert calls joint commitments, nor do we mean by collective belief something different than belief as, for example, acceptance, as some critics of Gilbert have proposed. What we actually do when we talk about the mental states of groups is similar to what we do when we talk about the mental states of objects, animals or even humans – we are using a strategy of interpreting, explaining and predicting the world around us based on our ability to understand intentions, desires, feelings and other mental states. This strategy, originally described by Daniel Dennett, is called the intentional strategy – and while it is not a perfect tool for describing or understanding the behaviour of groups, it is good enough in most everyday situations and, at the same time, is easy and intuitive enough for human beings to use effortlessly, which is why it is used so widely and why it presents such an interesting subject for philosophical study.

References

- Dennett, D. C. (1971): “Intentional Systems.” *The Journal of Philosophy* 68 (4): 87–106.
- Dennett, D. C. (1981): “True Believers: The Intentional Strategy and Why It Works.” In *Mind Design II: Philosophy, Psychology, and Artificial Intelligence*, ed. J. Haugeland, The MIT Press, Cambridge, MA, London, 1997, pp. 57–80.
- Gilbert, M. (1987): “Modelling Collective Belief.” *Synthese* 73 (1): 185–204.
- Gilbert, M. (2002a): “Acting Together.” In *Joint Commitment: How We Make the Social World*, M. Gilbert, Oxford University Press, Oxford, 2014, pp. 23–36.
- Gilbert, M. (2002b): “Belief and Acceptance as Features of Groups.” In *Joint Commitment: How We Make the Social World*, M. Gilbert, Oxford University Press, Oxford, 2014, pp. 131–162.
- Gilbert, M. (2002c): “Collective Guilt and Collective Guilt Feelings.” In *Joint Commitment: How We Make the Social World*, M. Gilbert, Oxford University Press, Oxford, 2014, pp. 229–256.

- Gilbert, M. (2004): “Collective Epistemology.” In *Joint Commitment: How We Make the Social World*, M. Gilbert, Oxford University Press, Oxford, 2014, pp. 163–180.
- Gilbert, M. (2014): *Joint Commitment: How We Make the Social World*. Oxford University Press, Oxford.
- Hakli, R. (2007): “On the Possibility of Group Knowledge without Belief.” *Social Epistemology* 21 (3): 249–266.
- Mathiesen, K. (2006): “The Epistemic Features of Group Belief.” *Episteme* 2 (3): 161–175.
- Rowlands, M. (2012): *Can Animals be Moral?* Oxford University Press, New York.
- Schmitt, F. (2014): “Group Belief and Acceptance.” In *From Individual to Collective Intentionality: New Essays*, eds. S. R. Chant, F. Hindriks & G. Preyer, Oxford University Press, New York, 2014, pp. 61–96.
- Sheehy, P. (2006): *The Reality of Social Groups*. Ashgate, Aldershot.
- Tollefsen, D. P. (2002): “Organizations as True Believers.” *Journal of Social Philosophy* 33 (3): 395–410.
- Tollefsen, D. P. (2015): *Groups as Agents*. Polity Press, Cambridge, UK, Malden, MA.
- Tuomela, R. (2000): “Belief versus acceptance.” *Philosophical Explorations* 3 (2): 122–137.

Abstrakt

Kolektivní přesvědčení a intencionální strategie

Co přesně máme na mysli, když říkáme, že je nějaká skupina o něčem přesvědčená? Myslíme tím pouze to, že jsou o dané věci přesvědčení všichni členové této skupiny, nebo tím snad přiznáváme existenci nějakého kolektivního agenta? Podle Margaret Gilbertové výroky o mentálních stavech skupin popisují specifický typ dohody, kterou nazývá společný závazek (joint commitment) – jinak řečeno, jsme-li o něčem přesvědčení společně, znamená to, že jsme společně s ostatními zavázáni k tomuto přesvědčení. Ve svém článku nejprve podrobněji představím přístup Gilbertové, nicméně následně poukáži na některé s ním spojené problémy, které povedou k jeho odmítnutí. Následně nastíním nejčastější řešení uvedených problémů, které spočívá v nahrazení kolektivního přesvědčení přitakáním (acceptance). Budu ovšem argumentovat, že ani toto řešení není uspokojivé. Ře-

šení, které posléze navrhnu, je založené na tzv. intencionální strategii Daniela Dennetta, což je metoda interpretace a predikce chování věcí kolem nás tím způsobem, že tyto věci chápeme jako racionální agenty s odpovídajícími záměry. Pokusím se ukázat, že veškeré problematické případy kolektivního přesvědčení jsou řešitelné aplikováním intencionální strategie na dotčené skupiny.

Klíčová slova: Daniel Dennett, intencionální strategie, kolektivní přesvědčení, Margaret Gilbertová, společný závazek

Kocourek, D. (2019): „Collective Belief and the Intentional Strategy.“ *Filosofie dnes* 11 (2): 57–75. Dostupné z www.filosofiednes.ff.uhk.cz

10 let LMS centra – historie a současnost

Martin Paleček

Filozofická fakulta
Univerzita Hradec Králové
Rokitanského 62, 500 03
Hradec Králové
martin.palecek@uhk.cz

Matej Drobnák

Filozofická fakulta
Univerzita Hradec Králové
Rokitanského 62, 500 03
Hradec Králové
matej.drobnak@uhk.cz

V roce 2009 bylo při Katedře filozofie a společenských věd Filozofické fakulty Univerzity Hradec Králové založeno LMS centrum (*Language-Mind-Society center*).¹ Letos je to deset let, co LMS existuje a rozvíjí své výzkumné aktivity. Sluší se tedy, abychom k tomuto výročí věnovali několik úvah o jeho minulosti a aktuálním směřování.

Historie LMS sahá do doby nového budování katedry filozofie, respektive přebudování Katedry filozofie a společenských věd z tehdejšího Ústavu filozofie a společenských věd poté, co se od něj odštěpila Katedra sociologie. Stalo se tak krátce po založení Filozofické fakulty UHK v souladu s plány na její rozvoj do plnohodnotné podoby.

Myšlenka založení vědeckého centra byla pozitivně přijata všemi tehdejšími členy katedry. Vzorem pro samotné centrum se stalo obdobné pracoviště na naší partnerské Univerzitě v Oslo. Tam, při katedře filozofie existovalo *Centre for the Study of Mind in Nature* (CSMN), které dokonce získalo statut norské excelence.

LMS bylo budováno tak, aby odpovídalo badatelským záměrům jednotlivých členů katedry a plánované profilaci hradecké filozofie. Proto byl

1 Informace o aktivitách LMS jsou k dispozici na webových stránkách (LMS Center) a na Facebooku (LMS – Language, Mind, Society Center). Záznamy některých přednášek jsou k dispozici na YouTube (LMS (Language, Mind, Society Center)).

zvolen i název *Language-Mind-Society Center*. LMS od samého počátku nemělo pouze dublovat práci samotné katedry. Předpokládali jsme, že centrum bude postupně růst a získá samostatnou pozici na fakultě i univerzitě. Plánem bylo, aby centrum získalo vlastní finanční nezávislost a v budoucnu nabízelo badatelská stipendia, která by bylo možno využít v době sabbaticku jednotlivých akademiků z UHK, ale i mimo univerzitu.

Druhým cílem, který jsme při zakládání centra měli na mysli, byla možnost zvaní přednášejících z jiných, zejména zahraničních univerzit, a tím i možnost rozšířit nabídku přednášek pro posluchače FF UHK.

Třetím cílem mělo být badatelské zapojení studentů, kteří se do práce centra měli zapojit v rámci specifického výzkumu vyhledávaného FF UHK.

Poslední, čtvrtý cíl měl vyřešit problém spojený s růstem KFSV a celé FF. Věděli jsme, že v případě dynamického rozvoje, značného množství výuky a rozvoje badatelství zákonitě klesá i množství příležitostí, kdy jsou jednotliví členové katedry ve vzájemném kontaktu. Proto byly zavedeny tzv. interní semináře, které dodnes slouží k tomu, aby členové katedry informovali kolegy o svém výzkumu a předkládali výsledky své práce kritickým pohledům druhých. Interní semináře měly také sloužit k tomu, abychom se mohli vzájemně informovat o nově publikovaných knihách a článcích v celé šíři svých badatelských záberů.

Na budování centra byl získán pětiletý grant ze Strukturálních fondů ve výši 4 milionů Kč s další pětiletou udržitelností. Vedoucím katedry byla jmenována vědecká rada LMS centra a prvním vedoucím centra se stal dr. Ondřej Švec, v současné době odborný asistent na Univerzitě Karlově.

Bohužel, některé původní záměry se z různých důvodů nepodařilo naplnit. Na svou finanční nezávislost, badatelská stipendia a další rozvoj centrum dosud čeká. Navzdory tomu můžeme v posledních letech pozorovat jistou konsolidaci. Po odchodu Ondřeje Švece převzal vedení centra Ladislav Koreň a úlohy koordinátora aktivit centra se zhostil Matej Drobňák. Pod tímto vedením, a za podpory FF UHK, centrum začalo novou etapu svého fungování a buduje si svou novou tvář.

Centrum se zaměřuje převážně na pořádání přednášek pro studenty a členy FF UHK. Z programu přednášek je již dnes patrný příklon

k internacionalizaci centra. Podstatná část přednášek je totiž vedena v anglickém jazyce. Tento trend je výsledkem dvou faktorů – na jedné straně je to snaha přitáhnout zajímavé a světově známé řečníky na půdu FF UHK, na druhé straně je to výsledek internacionalizace katedry jako takové. Jedním z cílů centra do budoucna je představit studentům řečníky, kteří hýbou současnou filozofií ve světě a které měli studenti doposud možnost poznat jenom zprostředkovaně – ze čtení jejich děl na seminářích. V tomto směru už stojí centrum na pevných základech z minulosti.

Na druhou stranu, badatelský růst katedry vedl k úzké spolupráci s řadou zahraničních výzkumníků a několik z nich našlo na katedře svůj „druhý domov“. To, že jsou interní semináře vedené v anglickém jazyce, tak už dávno není jen výjimkou, která potvrzuje pravidlo. Mezinárodní aktivity členů katedry lákají na katedru i nemalý počet zahraničních badatelů prostřednictvím různých stáží a programů mobilit (zejména Erasmus) – a každý z příchozích je na LMS srdečně vítán.

Mimo to se ale nezapomíná ani na přednášky v českém jazyce, které ocení zejména studenti nižších ročníků. Hlavní snahou je přivést na půdu LMS zajímavá a aktuální témata nejen z filozofie, ale i z příbuzných oborů jako jsou antropologie, psychologie, nebo kognitivní vědy. Růst katedry s sebou také přinesl i zvýšení počtu studentů doktorského studia, kteří zde začínají svou budoucí akademickou kariéru a kteří připravují nemalou část přednášek v češtině. LMS centrum se tak často stává prvním místem, na kterém studenti doktorského programu představují svůj projekt a mohou si tak otestovat nejen své rétorické schopnosti, ale i reakce na dotazy v domácí atmosféře.

LMS tak stále slouží především jako platforma pro pořádání přednášek pro studenty a členy univerzity. Vedle toho jsou stabilně vedeny i interní semináře. Za deset let své existence LMS hostilo více než 150 přednášejících, včetně takových jmen jako je Robert Brandom, Huw Price, nebo Martin Kusch, a stalo se nedílnou součástí života Katedry filozofie a společenských věd. Všichni věříme, že to tak zůstane i nadále a do dalších let života mu přejeme jen to nejlepší.

Informace o aktivitách LMS centra

Informace o aktivitách LMS najdete na webových stránkách:

LMS Center. Centrum pro studium jazyka, mysli a společnosti při Katedře filosofie a společenských věd Filozofické fakulty Univerzity Hradec Králové [online]. Centrum pro studium jazyka, mysli a společnosti, © 2009. Dostupné z: <http://lms.ff.uhk.cz/>

„LMS – Language, Mind, Society Center“ In *Facebook* [online]. 2015-. Dostupné z: <https://www.facebook.com/lmsffuhk/>

Záznamy některých z přednášek pořádaných centrem LMS najdete na Youtube:

„LMS (Language, Mind, Society Center)“ In *YouTube* [online]. 2015-. Dostupné z: <https://www.youtube.com/channel/UCuUqlV4QortVdSAjhinsGig/featured>

Paleček, M., Drobňák, M. (2019): „10 let LMS centra – historie a současnost.“ *Filosofie dnes* 11 (2): 76–79. Dostupné z www.filosofiednes.ff.uhk.cz

ČASOPIS
PRO
SOUČASNOU
FILOSOFII

