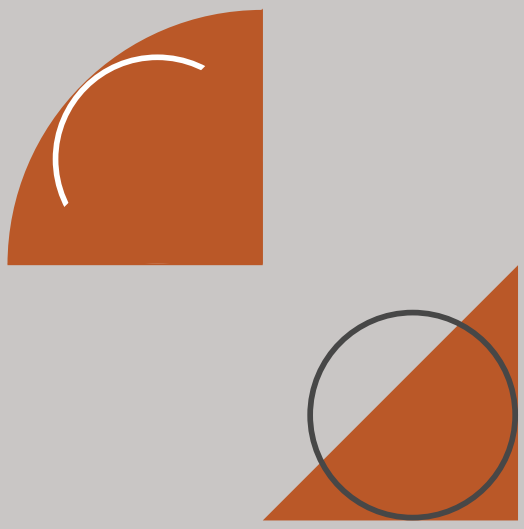
enfa 7

Seventh National Meeting in Analytic Philosophy
Lisbon 13 – 15 September 2018
Book of Abstracts



Sociedade Portuguesa de Filosofia Analítica
Centro de Filosofia da Universidade de Lisboa

book of abstracts

ENFA 7 – Seventh National Meeting in Analytic Philosophy

School of Arts and Humanities, University of Lisbon

13–15 September 2018

Edited by
Diogo Santos
Raimundo Henriques
Ricardo Miguel

Scientific Committee
Adriana S. Graça (UL)
António Lopes (LanCog, CFUL)
Bianca Cepollaro (IFILNOVA, UNL)
Bruno Jacinto (LanCog, CFUL)
Célia Teixeira (LanCog, UFRJ)
Dan Zeman (UVienna)
Daniel Durante (UFRN)
David Yates (LanCog, CFUL)
Dina Mendonça (IFILNOVA, UNL)
Domingos Faria (LanCog, CFUL)
Elia Zardini (LanCog, CFUL)
Fernando Ferreira (UL)
Fiora Salis (CPNSS, LSE)
Gonçalo Santos (CFCUL)
João Alberto Pinto (MLAG, UP)
João Cardoso Rosas (CEPS, UMinho)
João Branquinho (UL)
José Luís Falguera (USC)
Mona Simion (Cardiff University)
Nelson Gomes (UnB)
Pedro Galvão (LanCog, UL)
Ricardo Santos (LanCog, UL)
Rui Silva (LanCog, UAC)
Sofia Miguens (MLAG, UP)
Teresa Marques (LOGOS, UB)
Vítor Moura (CEHUM, UMinho)

Organizing Committee
Adriana Silva Graça (UL)
Célia Teixeira (LanCog, UFRJ)
Diogo Santos (LanCog, UL)
Raimundo Henriques (LanCog, UL)
Ricardo Miguel (LanCog, UL)
Sara Vargas (CFUL)

Graphic Design
Catarina Aguiar

Sponsors
Centre of Philosophy of the University of Lisbon
School of Arts and Humanities
MLAG/IF/UP
Portuguese Society For Analytic Philosophy
University of Lisbon
Fundação para a Ciência e Tecnologia

TABLE OF CONTENTS

PLENARY TALKS

Romantic Love for a Reason (*Disputatio* Lecture)
Berit Brogaard (University of Miami)
13.09, 11:15 – 12:45, Room: Anfiteatro III

In contemporary and historical contributions to the philosophy of love there has been considerable resistance to three claims concerning romantic love. (i) Romantic love is assessable for rationality, (ii) romantic love is love for a reason, and (iii) romantic love is reason-responsive. In this chapter I argue that these three ideas are intimately tied together. I offer justification for all three claims on the basis of more general considerations of the nature of emotions as well as evidence in support of the claim that romantic love is best rendered a complex emotion that *when felt* is truly multimodal. It attributes the property of causing certain internal qualities (e.g. a quickening of the heartbeat) that are experienced through interoception to a person identified through some perceptual or cognitive faculty (e.g., a sight of the beloved). Along the way, I identify some of the main ways in which romantic love differs from other kinds of love, such as friendship love and parental love.

Performing Works of Music Authentically: Another Go
Julian Dodd (University of Manchester)
14.09, 11:15 – 12:45, Room: Anfiteatro III

In this talk I return to the question of the varieties of authenticity governing our practice of performing works of Western art music. Building on, and in some respects revising, the views expressed in my 'Performing works of music authentically' (2015), here is what I now think. Our practice of work performance is governed by two authenticity norms: *score compliance authenticity* (accurately rendering the work›s score into sound); and *interpretive authenticity* (offering a performative interpretation of the work that evinces a penetrating, insightful or profound understanding of it). Both of these varieties of authenticity are valued for their own sakes, thereby placing them in a more rarefied normative space to that inhabited by other performance values. And yet it is interpretive authenticity which is the fundamental norm of the practice and that which grounds the desirability of score compliance.

Legal Duties, Rights, and Liberties: A Framework
Luís Duarte d'Almeida (University of Edinburgh)
15.09, 11:15 – 12:45, Room: Anfiteatro III

I have in previous work suggested we ought to resist W. N. Hohfeld's relational account of legal positions.[1] Hohfeld characterises legal duties, rights, liberties, powers, and so on, as positions in two-party relations. I argued against his approach. I also hinted at a link between Hohfeld's views on relationality and his loose grip on the notion of a liberty. But my arguments were not meant as a

[1] Luís Duarte d'Almeida, "Fundamental Legal Concepts: The Hohfeldian Framework", *Philosophy Compass* 11:10 (2016) 554-569; and "W. N. Hohfeld's 'Fundamental Legal Conceptions'," in M. N. S. Sellers and S. Kirste (eds), *Encyclopedia of the Philosophy of Law and Social Philosophy* (Dordrecht: Springer 2019, forthcoming).

complete defence of my claims; and I provided no non-relational framework of legal positions that could be set out against Hohfeld's own.

My goal in this paper to do just that. Focusing on duties, rights, and liberties, I will begin by drawing some distinctions to help us make progress on the matter. The main distinction is between two classes of statements legal positions. One is that of statements of (what I propose to call) *positive* legal positions: statements of the forms "*A* has a liberty to ϕ", "*A* has a duty to ϕ", "*A* has a duty toward *B* to ϕ" (or, equivalently, "*B* has a right that *A* ϕ"), and so on. The second is the class of statements of *definite* legal positions: statements of the forms "*A* has the liberty to ϕ", "*A* has the duty to ϕ", "*B* has the right that *A* ϕ." The distinction, I argue, is theoretically fruitful, and should form the basis of a framework for the analysis of fundamental legal positions.

I then develop such a framework, putting together a comprehensive picture of how statements of both kinds interrelate logically – both internally to each class, and between the two classes. And as the discussion unfolds, I will also present a fuller case for the general non-relationality of normative positions, and bring out more clearly the difficulties surrounding Hohfeld's (widely adopted, but misguided) take on liberties.

CONTRIBUTED TALKS

The Russell, the John, and the Katherine: The-Defenders of The-Predicativism

Eno Agolli (University of Connecticut, Storrs)

14.09 (15:05 – 15:40), Room: 5.2.

Recently, the philosophical world has witnessed a lively debate about the meaning of proper names. Most philosophers have thus far considered predicative uses of names as deviations from what they regard as their standard uses. In standard uses, they say, names have a kind of meaning radically different from that of predicates – in standard uses, names are singular, referential terms. Here, joining a minority of philosophers of language and linguists, I argue that this is wrong. Names are predicates, and can be argued to be predicates in all of their uses in natural language despite recent objections.

In §I, I briefly outline the two views about names under discussion. The standard view, Referentialism, claims that the main function of a name in language is to refer to an individual (directly or otherwise). As for predicative uses of names, Referentialism postulates a kind of semantic ambiguity. Another view, the one I endorse, Predicativism, denies this. Predicativism claims that names do not refer and are predicates everywhere they occur, thus avoiding the need to postulate semantic ambiguity.

In §II, I give the details of Predicativism. I outline its main motivation, which is that names and common count nouns have exactly the same syntax in natural languages. I then define their meaning as predicates: a name N is satisfied by an individual x iff x is called N. Finally, I outline the linguistic mechanism via which names can be used to refer, as I admit they do, to a thing (rigidly): names refer (rigidly) by being the predicative component of incomplete definite descriptions. The version of Predicativism that embraces this last idea is known as the-Predicativism, and it will be the one I endorse.

In §III, I consider a semantic objection to the-Predicativism: two sentences which, according to Predicativism, have exactly the same semantics, though different syntax, may exhibit a feature

which is not predicted by the-Predicativism semantically: they can make salient either the name of an individual or the individual itself. I argue that this is a pragmatic phenomenon, however, captured and explained by the pragmatic notion of topic or, as I call it, pragmatic stress, which I define as what is relevant or important in the context of utterance of a sentence. The difference in pragmatic stress of such sentences, I argue, is tracked by their different syntax.

In §IV, I use the notion of pragmatic stress to address two crucial objections. According to the-Predicativism, the definite article is used with names according to a certain syntactic rule. This syntactic rule (due to Fara (2015)) is what sustains the main motivation behind the-Predicativism. The syntactic objections are counterexamples to that rule. In response, I treat both of these counterexamples in a uniform manner: I show that both involve pragmatic stress on the name, rather than the individual itself.

I then recommend a modification of the original syntactic rule that accounts for pragmatic stress and thus accommodates the relevant counterexamples.

Explaining Identities

Lorenzo Azzano (SNS Pisa)
Joint work with Massimiliano Carrara (FISPPA, Padua University)
13.09 (17:20 – 17:55), Room: 5.2.

Can identities be explained? An argument has been offered by Kim (2010: 219-220) supporting the a negative answer to this question. We can frame Kim's argument as an argument about the explanation of identity facts, as follows: (schematically)

(P1) If p is a true identity sentence, it is necessary that p.
(P2) If it is necessary that p, nothing makes it the case that p.
(P3) If nothing makes it the case that p, it cannot be explained why p.
(P4) p is a true identity sentence.
(C) Thus, it cannot be explained why p.

After having accepted the necessity of identity in P1, P2 and P3 remain as the two philosophically substantial premises under discussion: according to P2, necessary facts cannot be grounded, and according to P3, ungrounded (or fundamental) facts cannot be explained. Both P2 and P3 are severely problematic. Most crucially, the individuation of a notion of explanation as a primarily epistemic or subject-relative relation, as opposed to a metaphysically substantial and mind-independent relation of grounding will allow us to put into question premise P3.

With this notion of explanation in mind, two sets of cases of explanation of identities are provided; viz. explanation *via* identity criteria and explanation *via* deduction. These cases are more naturally intended as explanations in a primarily epistemic sense, viz. one in which explanation is associated with clarificatory, or context-appropriate because-answers to why-questions. The clarificatory or appropriateness of such answers at least appears to have something to do with their informativeness relatively to their questions.

That being said, some of these cases of explanation of identities will play out as counterexamples of P2, identities that can be grounded, and then, presumably, explained. Others, more interestingly, will play out as counterexamples of P3, viz. identities that, even though they cannot be grounded

(and perhaps they are even fundamental), can still be explained. These cases, more specifically, will drive an important wedge between grounding and explanation. However, whether any case is be to read as a counterexample to P2 rather than P3 crucially depends on specific assumptions concerning whether identity facts can be grounded.

This discussion about the explanation of identities bears overarching consequences. Firstly, it constitutes an opportunity to explore an epistemic notion of explanation, although we do not fully articulate the character of "explanatoriness", the special ingredient that turns factual knowledge into explanatory knowledge.

Secondly, the discussion about the explainability of identity facts, *qua* necessary (and possibly ungrounded) facts, bears overarching consequences in a very recent meta-ontological debate, viz. the debate around the link between grounding –as an objective and mind-independent relation–, and explanation –as an epistemic or otherwise subjective-relative relation (e.g., Thompson 2016, Maurin Forthcoming).

The way we see it, we shouldn't really think about the relation between grounding and explanation as a contrast between two well-defined relations, one objective and mind-independent, and the other subjective, and tied to epistemic and agent-relative constraints. For there is no received view of these relations, especially grounding: on the contrary, grounding lives a dangerous "double life" Raven (2012: 689), as both a metaphysically substantial relation of non-causal determination, and an explanatory relation; and this dual nature is provably difficult to reconcile in an unitary perspective; the recent surge of interest in metaphysically deflated notions of grounding is proof of that. The contrast is rather between those who would take meta-philosophical operators such as "because" and "in virtue of" as expressing something metaphysically substantial, and those who wouldn't.

Is There Any Chance of Biological Randomness?

Augustin Baas (University of Geneva / University Paris-Sorbonne)

13.09 (10:10 – 10:45), Room: Sala de Actos

Stochastic descriptions have always been used in biology, and their usefulness in making predictions has made biologists confdent of their relevance to the feld. Stochasticity is involved in the description of various types of phenomena: the spontaneity of genetic mutations, the randomness of the genetic drift, the unpredictable behavior of animals, stochastic gene expressions, etc. But is there a claim to be made that we are dealing with 'biological randomness'? Obviously the claim assumes that the description makes use both of a notion of randomness and of the vocabulary of biology – not necessarily of biological laws –, such as the spontaneity to be found in genetic mutations. But it is not clear that this is enough to support a substantive claim on biological randomness. Indeed, it might be the case that the description can be formulated (i) within a biological non-stochastic framework (Sober 1981) or (ii) within a non-biological stochastic framework, typically one of physics (Brandon 1986, Stamos 2001, Weber, 2001, Mohseni 2014). And the very fact that neither (i) nor (ii) is presently established for the stochastic description concerned, which surely is the case for a very large majority of biological stochastic descriptions, is obviously not enough to ensure the in-principle impossibility of a reduction to (i) or (ii). This paper aims to assess the possibility of constructing from (i) and (ii) a twofold condition –irreducibility within biology and irreducibility to non-biological randomness– to be satisfed by a biological stochastic description in order to claim biological randomness. I will justify which notion of stochastic description is involved,

why no current biological stochastic description satisfies the two-fold condition, and why it does not follow from it reducibility to non-stochastic biological description. I conclude by promoting a methodological non-reductionism with respect to biological randomness.

The starting point is the clarifcation of the notion of randomness (Earman 1986). Randomness might be directly experimentally detected by the absence of correlations, order, patterns, etc., typically through statistical tests. Randomness refers to different mathematical notions, which each capture a lack of relationships between numbers in a sequence. But it suffers a well-known limitation: its direct evidence cannot be conclusive, statistical tests for example being unable to probe all possible correlations. So, even if the evidence of randomness is able to provide some support to particular descriptions that are framed in terms of randomness, it cannot help in proving neither (i) nor (ii). On the description side, randomness enters into the dynamics usually through probabilities. Randomness then characterizes the process, the different possible outcomes of which follow the probability distribution concerned. Hence, (i) and (ii) are going to be rephrased as (i) irreducibility of probabilities within biology and (ii) irreducibility to physical probabilities.

The debates in philosophy of biology involve, often separately, these two notions of irreducibility that I propose to characterize as follows: (BiR): The biological stochastic description cannot, in principle, be formulated within a biological deterministic theory; and (RiB): The biological stochastic description cannot, in principle, be formulated within a non-biological indeterministic theory. I argue that these conditions are individually necessary and jointly suffcient: to exhibit one biological stochastic description satisfying both (BiR) and (RiB) is enough to claim biological randomness; and any claim of biological randomness assumes at least one biological stochastic description satisfying (BiR) and (RiB).

Debates about biological randomness usually refer to the status of stochastic descriptions in physics. Statistical mechanics and quantum mechanics are then presented as the two paradigmatic cases of reducibility and irreducibility respectively. Against this view, I argue that they both do not differ with respect to the reducibility issue, as there is a quantum theory, namely Bohmian mechanics (Dürr 2009), in which probabilities have essentially the same status as they do in statistical mechanics. This is to say we do not have irreducible randomness for physical stochastic descriptions, thus quantum mechanics does not offer a model for (BiR). Yet, whatever the biological theory of concern, deterministic or not, biological probabilities, which are of the same mathematical nature as the ones in statistical mechanics, are not proved to be irreducible, neither within, nor outside of the field of biology. There is no biological stochastic description satisfying either of the two conditions of biological randomness. This finding has several implications, impacting upon four different debates or proposals in the literature.

First, it does not follow from it that biological stochastic descriptions are reducible within biology, i.e. have a status similar to probabilities within statistical mechanics (Aleksandr 1949). Second, it impacts debates about indeterminism accessed through biology, particularly on the possible percolation of the alleged quantum indeterminism at the biological level (Brandon 1996, Rosenberg 2001) – for, although some quantum effects might be maintained at the biological level (Mohseni 204), the absence of quantum irreducible randomness is a major objection against this idea of biological indeterminism accessed through the epistemic reducibility of biology to physics. Third, the status of the biological stochastic descriptions that we have derived, shows that we have to go beyond the dichotomy presented by the two understandings of probabilities – epistemic probabilities and propensities – as proposed by (Weber, Millstein 2003, Lyon 2011). Finally, combining the claim I defend with the heuristic usefulness of the concept of randomness in biology, which would be hardly

to be dismissed, on can promote a methodological non-reductionism about biological randomness (Van Regenmortel 2004).

An Assessment of Stout's "The Category of Occurrent Continuants"

Riccardo Baratella (University of Salzburg)

14.09, 9:35 – 10:10, Room: 5.2.

Introduction

A continuant continues to exist through time and may change during that period of time; an occurrent is something that occurs. The category of continuant includes physical objects such as Romeo and the Tour Eiffel. The category of occurrent includes events such as Romeo's death and the construction of the Tour Eiffel. Stout (2016) assumes that continuants are things that endure, i.e. persist without being extended in time. Events are occurrents that perdure, i.e. persist by being extended in time. Stout (2016) counters the traditional idea according to which all kinds of occurrent are extended in time. He argues that processes are a kind of occurrents that endure. In this talk, I will argue that Stout's main argument for the thesis that processes endure is not successful.

Stout's Framework

Stout's characterization of the notions of continuant (or enduring entity) and perduring entity concerns the ways in which a thing can have a property. According to Stout, continuants are things that primarily have their properties at times; perduring entities are things that primarily have their properties atemporally. In particular, when we ascribe a property, e.g. being bent, to a continuant, say Socrates, we have to specify the times at which Socrates is *bent*.

Stout's characterization of continuants relates to another feature that Stout takes to be definitional of this category of entities: continuants are things that may change over time. The notion of change assumed by Stout is the following:

(Change) Something changes over time if and only if this thing has a property at one time and at a later time the very same thing does not have that very property.

Then, given (Change), Stout shows that perduring entities cannot change.

According to Stout, a process is a thing that is, was or will be happening. Examples are my writing of this paper, or the explosion that was happening yesterday. Processes are described by expressions whose verb has a progressive aspect. Stout contrasts processes with another kind of occurrents, namely events, which are things that happened or will happen. Examples are Romeo's death, or Cicero's speech. Expressions whose verb has a non-progressive aspect are used to describe events.

Stout's Argument

Stout argues that processes primarily have their properties at a time. Consider a fight that went on outside my house between 11.55 p.m. and 12.05 a.m. last night. That fight was happening at midnight. Stout's argument can be reconstructed as follows:

(1) At first, the fight was quite brutal, but after a few minutes it become gradually less ferocious; meanwhile, it got gradually more noisy until the police arrived and stopped it.
(2) On the face of it, the fight is a thing that continues through time and has different properties at different times.
(3) By (Change), the fight changes.
(4) Since it changes, Stout's framework entails that it endures.
(5) By Stout's characterization, the ongoing fight primarily has its properties at a time.

Resisting Stout's Argument

The inference from (1) to (2) is based on intuitive grounds. I will provide two strategies to block this inference.

Argument from the official line

The inference from (1) to (2) is at odds with those positions according to which an occurrent – an event or a process – cannot change, e.g. (Prior, 1968). The advocates of this positions may help themselves with an account similar to the Kimian official line version of the theory of events as property-instances.

First, they may maintain that the properties which seem to be ascribed to the fight that was happening, such as being brutal, really are property-modifiers which, together with the property of fighting, give rise to other properties, such as the property of fighting brutally, instantiated by the participants to the fight at some time t. Then, they assert that the fight satisfies the predicate "being brutal" because its participants instantiate the property of fighting brutally. Finally, they affirm that the fight does not have the property of being brutal in any derivative sense. Hence, the fight cannot have different properties at different times, contra (2).

The availability of this line of response makes it clear that the inference from (1) to (2) makes substantial metaphysical assumptions that cannot be held without further arguments.

Argument from persistence

There is a model in which (1) is true, but (2) is false. Our ordinary talks about how the same thing is like at different times have to be accounted within a specific theory of persistence (Hawley, 2001). So, the truth-conditions of an ordinary statement like (1) have to be accounted within a specific theory of persistence. If perdurantism is adopted, "the fight" and "it" in (1) refer to the same perduring thing. Moreover, the truth-conditions of (1) are such that (1) is true iff the perduring fight has (atemporally) a t1-temporal part that has (atemporally) the property of being brutal with grade n and a different t2- temporal part that has (atemporally) the property of being brutal with grade m (with $m<n$). Thus, if perdurantism is adopted, these truth-conditions make (1) true. However, (2) is false: e.g., the perduring fight does not have itself the properties of being brutal with grade n at t1.

If Stout wants to deny that processes perdure, he has to reject the perdurance model of (1) and (2). But, he cannot reject this account, and maintains that (2) follows from (1), only on the basis of his intuitions. Since, if he does so, he begs the question against the thesis that processes perdure.

Conclusions

Even though Stout's main argument does not succeed, it doesn't follow that processes perdure. The question of how processes persist has to be decided on the basis of several considerations: e.g., by the examination of the process-features of homogeneity and open-endedness.

Is Interactional Linguistic Relativity a Case In Favour of Enactivism?

Filippo Batisti (Ca' Foscari University of Venice)

14.09, 9:00 – 9:35, Room: Sala de Actos

Do different natural languages affect human thought in different ways? This is the classic formulation of the linguistic relativity (LR) principle (Whorf 2012). In this paper, I argue that, first, a non-representational form of linguistic relativity, namely the interactional one, is possible. Secondly, I explore the possibility that such a form is compatible with an enactivist account of the mind (Hutto and Myn 2017). *Prima facie*, these theses contrast the previously accepted ways of conceiving LR, including the Neo-Whorfian one, *i.e.*, the standard position in the last 25 years.

So far, it has been studied how linguistic categories correlate to non-linguistic mental categories, and if the former affect the latter. Several domains have been investigated, such as color categorization, spatial frames of reference, objects and substances, grammatical gender, etc. (see Everett 2013, Casasanto 2016). The common underlying scheme was the following: certain structural properties of languages, which embody particular "interpretations of reality", may result in "influencing thought about that reality" in the respect of interpretation of experience (Lucy 1997). It has been pointed out that, in principle, "behaviour-centered approaches" could only be successful if they appealed to "concrete everyday situations"; but at the same time, for Lucy (*ibid.*), this kind of analysis was bound to remain vague and untestable.

However, recently, this tenet has been put under further scrutiny. Enfield (2015) acknowledged that due to a common trend in many language-related fields other functions of language beyond reference and representation have been overlooked, and LR made no exception in this respect. In order to fill this gap, Sidnell and Enfield (2012) proposed a third *locus* of LR, namely, social interaction. This view holds that different languages make available different particular "lexicogrammatical resources" as tools to accomplish a basic action in interaction: data from conversation analysis have shown that each language-specific pattern carries "collateral effects" in terms of the subsequent development of the interaction. In other words, there is diversity in how speakers bring about the same action, due to linguistic motivations and *e.g.* not to explicit cultural prescriptions.

I argue that this particular kind of LR has one more peculiarity, that has not been highlighted by its proponents: to describe it there is no need to appeal to mental representations, contrary to the classic relativity models (*inter alios*, Levinson 2003, Boroditsky 2011) which relied on a pictorial/visual metaphor (Zinken 2008a, 2008b). Thus, it can be stated that it is possible to live in different worlds without having different mental representations. Such an account seems at first to be compatible with a radical form of enactivism, a psychological approach that denies the representational nature of cognition (Hutto 1999, Hutto *et al.* 2014). Some objections are raised, but, overall, it seems that (a loose?) compatibility holds: if so, interactional relativity may be a single – but not irrelevant – case in favor of enactivism.

Counsciousness, Concepts and Natural Kinds: on the Prospects for a Consciousness Meter

Tim Bayne (Monash University)

13.09, 15:05 – 15:40, Room: 2.13

We have various everyday measures for identifying the presence of consciousness, such as the capacity for verbal report and the intentional control of behaviour. However, these measures are difficult (if not impossible) to apply when it comes to detecting consciousness in infants, brain-damaged individuals, non-human animals, and machines. There is thus a pressing need to identify new ways of measuring consciousness. In recent years, a number of theorists have suggested that this challenge can be met by treating consciousness as a natural kind (Shea 2012; Shea & Bayne 2010).

Given how successful the natural kind (NK) approach to measurement has been in other areas of science, there is every reason to expect that it can also be successfully applied to consciousness. However, it is also possible that there are good reasons to think that the NK approach isn't appropriate for the science of consciousness. This paper examines two arguments for that claim, both of which focus on the distinctively first-personal way we have of thinking about consciousness, and the fact that our primary means of grasping consciousness is via phenomenal concepts (Balog 2009; Sundström 2011).

Why might phenomenal concepts cause trouble for the NK theorist? One idea is that phenomenal concepts reveal the 'essential nature' of their referents, and that what is revealed is at odds with consciousness being a natural kind. For example, Chalmers describes a phenomenal concept as a concept that picks out its referent 'in terms of its intrinsic nature' (2003: 225); Nida-Rümelin says that our grasp of phenomenal properties goes via a grasp of phenomenal concepts, and that 'to grasp a property is to understand what having that property essentially consists in' (2007: 307); Horgan and Tienson claim that when one thinks of a phenomenal property via a phenomenal concept one thinks of it 'directly, as it is in itself' (2001: 311); and according to Goff, 'phenomenal concepts reveal the essence of the states they denote' (2017: 107). The idea, in short, is that a phenomenal concept provides knowledge of essential truths about its referent. The paper explores various ways in which this claim might be developed, and concludes that none of the analyses suggested by these authors puts pressure on the NK approach, either because the analysis is not itself plausible or because it is consistent with treating consciousness as a natural kind.

A second way in which phenomenal concepts might be thought to cause trouble for the NK theorist concerns the epistemic status that is associated with first-person applications of phenomenal concepts. The idea here is that phenomenal concepts bring with them a peculiar kind of direct and unmediated access to their referents that cannot be accommodated by the NK approach. The worry is that there could be cases where self-application of a phenomenal concept conflicts with a verdict about the subject's state of consciousness made on the basis of the NK approach. In the same way that science can tell us that what we thought was gold is actually fool's gold, so too – the objection runs – the NK approach seems to be committed to the (epistemic) possibility of 'fool's phenomenology'. Putting it another way, the NK approach seems to entail that there could be third-person tests for consciousness that 'trump' the self-application of phenomenal concepts, where a second test trumps the first just in case one is more justified in trusting the second than the first when they conflict. The paper considers various responses to this objection, arguing that it is possible to reconcile the NK approach without jettisoning the epistemic authority of the first-person perspective.

The Identity-Enactment Account of Familial Duties

Saba Bazargan-Forward (UC San Diego)

13.09, 15:40 – 16:15, Room: Mattos Romão

Introduction

Familial duties are agent-centered duties to give defeasible moral priority to the closest members of our family. But where do these duties come from? 'Antireductionists' argue that familial relationships yield *sui generis* duties. Alternatively, 'reductionists' identify familial duties with special duties. Special duties are duties arising from what you have done to others or what others have done to you. For example, reductionists ground parental duties in two special duties: the duty to keep promises, and the duty to care for those we've made vulnerable. If you agree to provide for a child's needs, or you are responsible for creating someone with vital needs which that individual is unable to satisfy alone, then you must provide for those needs, unless that duty is transferred to someone else. By identifying familial duties with special duties, reductionists make sense of the extra weight we must give family by identifying it with the extra weight special duties afford.

The purpose of this paper is twofold. First, I attempt to show that reductionism faces an unnoticed challenge: it has difficulty explaining the *radical priority* of our familial duties. Second, I present an alternative reductionist analysis of familial duties – the 'Identity-Enactment Account' – which not only accommodates the radical priority of our familial duties, but also characterizes our familial duties in an intuitively compelling way. On this account, our strongest familial duties are special duties to *adopt and enact a practical identity in which the duty's beneficiary features prominently*.

Part 1

To see why special duties are not strong enough to ground familial duties, consider the following case.

'Daughter Rescue'

I can save either my daughter from drowning or three strangers, but I can't save all four.

Arguably I have not just a permission but a duty to choose my own child, even though this comes at the cost of allowing three others to die. (This is not a duty to kill – it is instead a duty to refrain from saving the greater number).[2]

Supposing this is correct, how do we account for it? According to reductionism, the parental duty I have toward my child is a special duty to fulfill a promise to care for the child, and/or a special duty to care for the child as a result of having made her vulnerable.

Consider the promissory duty first. It's doubtful that such a duty can warrant choosing one's own child in Daughter Rescue. To see this, consider an impersonalized version of the case:

'Promise Rescue'

I made a promise to you, a stranger, to save your life should it ever be endangered. One day I encounter a situation in which I can either save you from drowning, or three strangers – but I can't save all four.

[2] For an overview of how reductionists identify familial duties in these (and other) special duties, see (Seglow, Jonathan 2013, 'Defending Associative Duties'. New York City: Routledge, p. 8-10).

Presumably, no such promise can ground a permission to rescue you if doing so means failing to save three others. Those circumstances describe the defeaters for such a promise. So if we think that I do indeed have a duty to save my child's life in Daughter Rescue even at the cost of allowing three others to die, then we cannot ground that duty in a promise I made to my daughter. Promises are not strong enough to capture the strength of familial duties.

Recall, though, that reductionists also identify parental duties with the special duty to care for those we have made vulnerable. The problem, though, is that the duties foster-parents have to their adopted children do not include a special duty to protect those whom they have made vulnerable, since by hypothesis foster-parents did not create the adopted children.

The upshot is that familial duties require us to do what no special duty can warrant. This suggests that we cannot use those special duties as a hook upon which to hang our familial duties.

Part 2

I believe reductionists need to radically rethink the content of the special duties in which our familial duties consist. Specifically, our strongest familial duties are not special duties to protect and promote the welfare of those to whom the duties are owed, simply. Instead, according to:

The Identity-Enactment Account

Our strongest familial duties are special duties to protect and promote the welfare of the duty's beneficiary by adopting and enacting a practical identity in which the duty's beneficiary features prominently.

Recall that our strongest familial duties enjoin us to grant radical moral priority to family. A mere promise to protect and promote their welfare falls short. We can grant much greater moral priority by essentially making their welfare ours – i.e., by incorporating them into our practical identity. This yields agent-centered permissions which permit us to give their interests much greater weight than any mere promise could. The special duty to protect and promote their interests then requires that we then act in accordance with those permissions.

In Daughter Rescue, I have an agent centered permission to treat my daughter preferentially, where this permission derives from the role she plays in personal projects constitutive of my practical identity. The same agent-centered permission allowing me to save myself at the cost of failing to save three strangers also allows me to do the same for my daughter, because her well-being is constitutive of my practical identity.

Adverting to our practical identities in this way is a promising strategy in part because, unlike special duties, our reasons to act in accordance with the ends specified in our practical identities possess sufficient reason-giving force to warrant choosing one's own child in cases like Daughter Rescue. Unlike special duties, these agent-centered permissions can do the heavy lifting necessary to meet the Stringency Condition.

The Identity-Enactment account characterizes familial duties in an intuitively appealing way. There are people who can legitimately demand a prominent place in our mental lives for the protection and intimacy that prominence affords. Our familial duties are special duties to accommodate that legitimate demand. We can meet that demand by adopting and enact a practical identity in which the duty's beneficiary features prominently. I argue that incorporating such individuals into our practical identity in this way, and then acting in accordance with the permission doing so affords, provides the duty's beneficiaries with the sort of protection and intimacy intuitively characteristic of familial duties.

Contingent Composition as Identity

Massimiliano Carrara (FISPPA, University of Padua)
Joint work with Giorgio Lando (University of L'Aquila)
14.09, 10:10 – 10:45, Room: 5.2

The debate about Composition as Identity (henceforth CAI) follows a recurring pattern: the question is asked whether composition enjoys a certain feature of identity and whether, as a consequence, composition qualifies – from the viewpoint of that feature – as a form of identity. If composition enjoys this feature, then the friends of CAI mark a point. Ifcomposition lacksthisfeature of identity, then the foes of CAI mark a point.

In this paper, we argue that, when CAI is confronted with the modal features of identity, and in particular with the Necessity of Identity (NI), we should be careful in following this pattern. In particular, it is methodologically desirable to construe the modal notions involved as standardly as possible. This already happens when, e.g., the debate on CAI is focused on the Indiscernibility of Identicals (InId): a standard, unrestricted version of InId is considered, and the question is raised whether composition respects it. The reason why it would be misleading to assume non-standard contentions in the assessment of CAI is not that they are provably false, but simply that they are non-standard and are, for this very reason, unsuitable as a neutral background for the assessment of CAI. If a non-standard background were assumed, the doubt would persist that CAI is refuted or confirmed on the basis of a wrong external assumption.

We will argue that it is equally desirable to adopt standard, neutral contentions when modality is at stake, and in particular when the question is raised whether composition is as necessary as identity. By contrast, in the extant literature (in particular T. Merricks, Composition as Identity, Mereological Essentialism, and Counterpart Theory. AJP 77(2), 1999: 192–195), a non-standard approach to modality is adopted (e.g. counterpart theory), and on this basis CAI is shown to be incompatible or compatible with NI. Our analysis assumes as a standard philosophy of modality (SM) a broadly Kripkean approach, in which inter-world identity is admitted and the extensions of predicates vary from world to world.

Given SM and quite independently of CAI, also the extensions of parthood and composition (as those of any other standard property or relation) should be relativized to possible worlds, in order to avoid violations of InId, when a single entity has different parts at different worlds.

We will show that, once SM and the ensuing relativization of parthood and composition to possible worlds are considered, the picture of the relation between CAI and NI is more complex than the picture given by the current literature. In particular, in order to give a minimal, initial plausibility to CAI's claim that world-relative composition is a kind of identity, also identity should be relativized to worlds. Otherwise, the various relativized forms of composition would collapse on a single, non-relativized form of absolute identity, thereby losing track of the relativization to worlds, which was needed in order to preserve InId.

In this way we obtain a variety of CAI which has not been yet discussed in the literature, namely the thesis that composition is a world-relative, contingent form of identity. We label this new variety of CAI Contingent Composition as Identity (CCAI). If CCAI is adopted, we end up needing two distinct notions of identity: absolute identity is needed in order to express the constitutive thesis of SM, according to which there is inter-world identity; non-absolute, world-relative identity is needed in the formulation of CCAI.

Three main features differentiate them. Standard identity:

a1) is not relative to a possible world;
a2) holds necessarily (NI);
a3) is governed by an unrestricted principle of indiscernibility, InId.

On the other hand, the kind of world-relative identity which characterizes CCAI:

b1) is relative to a possible world;
b2) does not hold necessarily;
b3) is governed by a different principle of indiscernibility, restricted to world-relative properties.

We will show that these differences between these two notions determine – so to say – both the virtue and the limit of CCAI.

The virtue is that CCAI preserves the contingency of composition, without being revisionary about NI. The limit is that world-relative identity is introduced in a suspiciously ad hoc way. Indeed, there is no reason to think that there is a world-relative kind of identity outside of CCAI: the extension of world-relative identity at each world will exclusively include what CCAI is about. As a consequence, to say that any instance of composition is an instance of world-relative identity is not to assimilate composition to an independently characterized relation. CCAI is not the claim that composition is something else, already known. A foe of CAI might even insinuate that "world-relative identity" is simply a new name for world-relative composition.

We will argue that this foe of CAI is too negative. Admittedly, CCAI does not consist in seeing composition as a subcase of a pre-existing notion of identity. As a consequence, CCAI is incompatible with so-called strong CAI, namely the radical version of CAI, according to which any instance of composition is strictly speaking an instance of standard, absolute identity.

Nonetheless, CCAI might be seen as an interesting development of the so-called weak variety of CAI, originally endorsed by Lewis, according to which composition is not strictly speaking identity, but merely analogous to it. CCAI indeed pinpoints a rather precise and interesting analogy between composition and standard identity.

Namely, while standard identity obeys an unrestricted principle of indiscernibility (a3) above), composition – according to CCAI – obeys a restricted, but still quite comprehensive and highly ambitious principle of indiscernibility (b3), which quantifies over every world-relativized property.

The result that composition obeys a quite strong indiscernibility principle might be an additional reason to think that composition is indeed analogous (as weak CAI contends) to standard identity. Thus, the foe of CAI would be overhasty to argue that CCAI is a mere way of rebranding composition as identity, without any philosophical content. By contrast, we will argue that CCAI is an interesting development of weak CAI.

Laws of Nature and Explanatory Circularity

Eduardo Castro (LanCog, Universidade da Beira Interior)

15.09, 9:00 – 9:35, Room: Sala de Actos

Some recent literature (Hicks and Elswyk 2015; Bhogal 2017) has argued that the non-Humean conceptions of laws of nature have a same weakness as the Humean conceptions of laws of nature. Precisely, both conceptions face a problem of explanatory circularity: Humean and non-Humean conceptions of laws of nature agree that the law statements are universal generalisations; thus, both conceptions are vulnerable to an explanatory circularity problem between the laws of nature and their instances. After I distinguish between two explanatory circularity problems, I will defend that Armstrong‘s necessitarian view of laws of nature is invulnerable to these explanatory circularity problems.

In the literature, the terminology “explanatory circularity problem” has been used to designate two slightly different problems. One, it is a full problem of explanatory circularity, hereafter the problem of circularity C. A law of nature is inferred from an observed phenomenon and, thereafter, it is used to explain that same observed phenomena. The other, it is better seen as a problem of self-explanation, hereafter the problem of circularity SE. The problem of circularity SE is a sub-problem of the problem of circularity C. A law of nature explains an observed phenomenon but the law of nature includes that same phenomenon in its content.

Hicks and Elswyk (2015) propose the following argument for the problem of circularity C:

(P1) The natural laws are generalizations. (HUMEANISM)
(P2) The truth of generalizations is (partially) explained by their positive instances. (GENERALIZATION)
(P3) The natural laws explain their instances. (LAWS)
(P4) If A (partially) explains B and B (partially) explains C, then A (partially) explains C. (TRANSITIVITY)
(C1) The natural laws are (partially) explained by their positive instances. (P1 & P2)
(C2) The instances of laws explain themselves. (P3, P4, & C1) (Hicks and Elswyk 2015, 435)

They claim that this argument also applies to the non-Humean conceptions of laws of nature: “Humeans and anti-Humeans should agree that law statements are universal generalizations (...) If we‘re right about this much, anti-Humeans are vulnerable to a *tu quoque*.” (Hicks and Elswyk 2015, 435, my italics)

Armstrong‘s necessitarian view of laws of nature is a non-Humean conception. Laws of nature govern the events. They confer order to the observable world. They are additional entities above the regularities of the Humean mosaic. In light of this view, there are several troubles in the argument above. Let us suppose that the terms of the premise (P1) “natural laws” refer law-statements. That is, the natural laws are regularities, “all Fs are Gs”. Premise (P1) is true, because it is widely assumed that most law-statements are generalisations such as “all Fs are Gs”. However, in light of Armstrong conception of laws of nature, premise (P2) is false.

Armstrong claims that laws of nature are states of affairs. A second-order relation, called N, binds first-order universals F and G (i.e. N (F, G)). This second-order relation, N (F, G), entails and explains the regularity, “all Fs are Gs” (Armstrong 1983, 41; Armstrong 1988, 225; Armstrong 1993, 422).

However, “all Fs are Gs” is not a law of nature. The generalisation “all Fs are Gs” is a law-statement such as Kepler‘s laws, Newton‘s laws and so on. For example, the law-statement “all ravens are black” is entailed and explained by the law N(F,G), where F is the universal ravenhood and G is the universal blackness. A particular raven a is black because it instantiates the universals ravenhood and blackness and these universals are necessary related by N. In sum, (P2) is false because the truthmaker of the generalisation “all Fs are Gs” is the state of affairs, N (F, G) (Armstrong 1991, 507). The falsity of (P2) is enough to block the problem of circularity C.

The argument above can be reframed to underpin the problem of semantic circularity SE.

(P1) The natural laws are generalizations (HUMEANISM)
(P2)* If natural laws are generalizations, then the natural laws are constituted (in part) by their instances.
(P3) The natural laws explain their instances. (LAWS)
(C2) The instances of laws explain themselves.

Premise (P3) is false. The law statement “All Fs are Gs” does not explain their instances. It is again the universal N (F, G) that explains their instances.

In my communication, I intend to develop these necessitarian replies to the problems of explanatory circularity C and SE.

The Phenomenal-Less Theory of Mood

Daniele Cassaghi (University of Milan)
13.09, 14:30-15:05, Room: 2.13

Among the various kinds of mental states, moods strike us because of their undirectedness. When we are in a certain mood, it seems that we undergo to a distinctive experience but we report this experience not to be directed to anything in particular. This makes moods different from other states like emotions. In addition to undirectedness, the literature about moods describe two other features which are typical of moods and emotions lack. The first one is their peculiar *role in folk-psychological explanations.* In other words, in our folk-psychological explanations we consider moods as having a cause but not a reason. On the contrary, emotions can be viewed as having reasons in the right circumstances. Moreover, we do not treat mood as evaluative states. The opposite is true for emotions. (Lormand 1985, DeLancey 2006)

The second mood feature is the *pervasive effect* moods have to our overall cognition (Lormand 1985, DeLancey 2006). A mood is usually considered responsible for the occurrence of a wide array of mental states: a depressive state, for example, is responsible for the generation of various negative beliefs and desires. Emotions seem not to have the same impact. A theory of moods should be able to account for all these peculiar features: undirectedness, folk-psychological role and pervasivity.

The undirectedness of moods is the main sources of difficulties for philosophers with intentionalist inclinations (Tye 1995, Crane 1998). According to intentionalism indeed, the phenomenal character of any experience is nothing over and above the content of the experience itself. Moods provide an obvious counterexample to the intentionalism: they are mental states with a distinctive phenomenal character and no content. Intentionalists reply by denying moods’ undirectedness. They claim that

moods are about a very general object, namely the world. (Crane 1998).

This strategy gives up the most salient feature distinguishing moods from emotions. Arguably, an intentionalist may even try to conflate moods and emotions by making two moves. If moods have content, they may be evaluative states like emotions. If their content is a general one, this explains pervasivity. So, from the intentionalist strategy endowing moods with contents, one should come to give up *all the distinctive* features of moods. I shall argue that it is possible to save the both undirectedness of moods and intentionalism and offer an account of moods that vindicates their salient features (pervasiveness and their folk-psychological role), while keeping them distinct from emotions.

I shall endorse a functional account of the moods (Lormand 1985, Griffiths 1997, Sizer 2000) and argue that moods are parameters, whose primary function is to generate biased cognition. For example, anxiety has two main functions: 1) favoring the generation certain mental states (beliefs regarding threats in the environment, emotions like anger and fear) and 2) hampering positive emotions like joy and happiness and beliefs about positive things. Following a suggestion by Price (2006), I contend that moods evolved as having these functional roles, in order to make us readier to react to the surrounding environment. This appeal to the functional role of moods is tailor-made to account for both pervasivity and their folk-psychological role.

Unlike Lormand, Sizer and Griffiths I shall argue that moods do not have a content and, in accordance with the intentionalist doctrine, they do not have a phenomenal character either. I shall argue that by giving up the assumption that moods are endowed with a phenomenal character, we can capture all the features commonly associated with them (undirectedness, pervasivity and their folk-psychological role) in a way compatible with intentionalism. This also secures the distinction in kind between moods and emotions. In the final part of the paper I shall try to locate the origins of the mistaken view that moods have a distinctive phenomenal character. I will make three moves. The first move is to recognize that, in accordance with the intentionalist doctrine, emotions have a genuine phenomenal character because they are intentional states. The second move is to maintain that moods cause the occurrence of certain kinds of emotions (as predicted by the functionalist theory of moods). The third move is to claim that we erroneously attribute to the mood the phenomenal character belonging to the emotions generated by the mood itself. We tend to make this mistake by attending to the phenomenal character of the emotions without paying attention to their content. We do not make this mistake, if we also attend to the emotions content.

In Which Sense Do We Obey the Criminal Law?

Olivier Chassaing (Sophiapol, Paris Nanterre University)

15.09, 9:35 – 10:10, Room: Mattos Romão

Punishment is a paradoxical institution of contemporary democratic societies: the abuses and over-penalisation consequences of security policies are criticized, but the impunity of certain crimes remains scandalous and people urge for justice; the abolition of specific kinds of punishments (like prison), or the introduction of alternative forms of regulation (for instance practices of restorative justice) are claimed, but the difficulty of enforcing law without deterrent sanctions seems inextricable.

Facing these dilemmas, one can advocate that punishment bears several roles in contemporary liberal democracies: it has first a *coercive* function, since it is used as an afflictive instrument (Bentham, 1823); but it also gives a vehicle to collective indignation and resentment, following its *expressive* function (Feinberg, 1965); finally, and given its legal nature, punishment also assumes a

normative function, which is manifested through various phenomena: criminal laws affirm the fundamental character of certain prohibitions and take part in recognizing wrongs suffered by individuals, sometimes against the dominant morality (Hart, 1963); it renews social norms and distinguishes offenses that matter to the state from those that will stay invisible; processes of prosecuting and sentencing shape the content of social conflictuality by raising demands for justice in front of a public authority.

Nevertheless, the thesis of this last function needs to be thought through: it assumes that punishment is fundamentally a retributive practice, and that this conceptual feature is a descriptive and not a normative one.

To address this thesis, I would, in a *first* place, face the problem of whether the functions of criminal law can be described without being justified at the same time. In a *second* place, following the discussions of legal positivism to properly describe the relations between legal norms and the constraints provided to enforce them (Hans Kelsen, 1945; H.L.A. Hart, 1961; Joseph Raz, 1979; Duarte d'Almeida, 2007), I would like to show that penal coercion has a normative power *sui generis*: through punishment, the state lays down practical references and evaluation *criteria* for individuals, despite the generality of legal statements and the disagreement that may affect officials' judgements. In a *third* place, I would then examine how the norms issued by most criminal systems in liberal democraties, namely the sanction- and prohibition-norms, are submitted to various interpretations: inferred from the *a priori* threat of punishment, prohibitions are shaped in a way that determines the type of acceptance and obedience of which they can be the object, *a fortiori* when individuals only recognize the force of law by its punitive side (Schauer, 2015).

What the Values of 2nd Order Variables Could Be?

Demetra Christopoulou (University of Athens)

14.09, 17:55 – 18:30, Room: Sala de Actos

This paper raises the question about what kind of entities the values of 2nd order variables are. For example, ∃FF(a) is a second order formula in which the quantifier bounds a 2nd order variable 'F'. Hence 2nd order quantifiers apply to positions of predicates. An extensional reading of predicates raises the question W. V. Quine has settled about the ontological commitments of 2nd order logic. He assimilated it with a "wolf in sheep's clothing" (set theory rather than logic) to the extend that he was suspicious about its ontological innocence. According to Quine, to put a predicate symbol 'F' under a quantifier is to treat positions of predicates as positions of names and deal with predicates as if they were names of entities of some kind. The threat is that higher order entities are denoted if common characteristics of individual objects are conceived in a way that hypostasiases them (e.g. redness, triangularity etc.) So, if an intensional reading of predicates is endorsed, then 2nd order quantification is ontologically committed to properties.

Ante rem realists take properties to be abstract entities which lie outside the spatio-temporal realm and independently of their instances whereas *in re* realists take them to exist *in* their instances. If one pursues a deflationary account of properties she has to dismiss the metaphysical status of them as entities. Here, it would be useful to think of an asymmetry that occurs in case of the relation between names and their references on the one hand and predicates and properties on the other hand. Names refer to objects however predicates should not be taken to refer to properties, for example, the predicate 'x is red' should not be taken to refer to the property *redness*. The role of a

predicate is to designate a property or to *ascribe* a property to certain individual objects. Hence, the relation among a predicate and a property might be that of *ascription* rather than that of reference. So a property is the ascription of a predicate. If properties are not referents of predicates then they are not necessarily regarded as entities. This account is consistent with an antirealistic approach to properties. In this case, 2nd order quantification does not necessarily commit to any abstract entities. Besides, we should think that not every predicate ascribes a property, for example negative predicates, disjunctive predicates etc. do not ascribe properties at all. Moreover the ontological status of properties as entities may be doubted on the grounds that they do not possess definite identity criteria. Frege himself has noted (in Grundlagen §68), that properties may have the same extensions but different intensions, so to state a criterion of identity in this case faces an ambiguity. For example, "being equal sided" and "being equiangular" are distinct but co-extensional properties. If we identify coextensive properties then we ignore their intensions.

Deflationary accounts of properties are often offered by construing them as concepts. Concepts and universals have a similarity in that they apply to more than one object. They express generality since both express features in common. In a nominalistic view, universals can be taken to be concepts. Contemporary conceptualism deflates universals as well. Generality is situated in thought alone so concepts are taken to be mental entities. Empiricists and naturalists deal with them as mental or psychological entities. If concepts are taken to be mental constructions and 2nd order variables take concepts as their values then 2nd order quantification is committed to concepts. Then there is no ontological blame to impute. Besides, according to E. J. Lowe, concepts are distinguished from universals since the latter are purely extralinguistic entities but concepts are mental. Concepts should be taken to be ways of thinking, hence we have more than one way of thinking of the same entity. So, ontological commitment to concepts appears to be faint. Nonetheless, if concepts are to be construed as *functions* on the basis of a Fregean account then 2nd order quantification is ontologically committed just to functions. A Fregean concept is a function whose values are truth-values (for example, 'x is a satellite of Jupiter' is a Fregean concept). What is the essence of a function is *unsaturatedness*. Hence, Fregean concepts need not be construed as universals or other abstract entities. In conclusion, if 2nd order variables get Fregean concepts, i.e. functions, as their values then ontological commitments appear to be eliminated.

Nevertheless if ontological commitments of 2nd order variables have to be ruled out, intensions of predicates should not be construed as abstract entities of any kind. Any account of such intensions that construes them as abstract objects reinforces ontological commitments of 2nd order logic.

Abstraction and Contradiction: the Role of Basic Law V in Russell's Paradox

Ludovica Conti (Università di Pavia)

15.09, 10:10 – 10:45, Room: 5.2

As is well known, Frege wanted to prove as logical theorems the axioms of arithmetic. No less notoriously, Russell has however shown that the logic, by means of which Frege wanted to carry out this ambitious program, allows the derivation of a contradiction, and so is unfit to the purpose. This is what has come to be known as Russell's paradox. In this paper I purport to compare with one another three different "explanations" of Russell's paradox, briefly rehearse the reasons why the first (Cantorian) explanation can be neglected, and give some final reasons for preferring a combination of the second (Predicativist) and the third (Extensionalist).

A paradox is an argument in which "an apparently unacceptable conclusion (is) derived by apparently acceptable reasoning from apparently acceptable premises" (Sainsbury 1995). To give an "explanation" of a paradox is to offer a diagnosis that is preliminary to the therapy, namely it is to identify the non-obvious flaw that makes the premises or the reasoning erroneously acceptable or that makes the conclusion erroneously unacceptable.

More in particular, by "explaining" a paradox that – as Russell's paradox – ends with a contradiction, I mean, first and foremost ("syntactic thesis"), giving a derivation of the contradiction and identifying which premises are necessary conditions of that conclusion, namely which axioms or theorems it is sufficient to abandon for the purpose of restoring consistency. Different explanations of the same paradox so diverge from one another in terms of the theorems and axioms they pick out as premises for the purpose of the derivation. As I will attempt to show, the proposal of a given explanation is however susceptible of being motivated also by semantic considerations ("semantic thesis"). So, apart from its syntactic features, the explanations of Russell's paradox I shall review – Cantorian, Predicativist and Extensionalist – can also be assessed in terms of their alleged semantic underpinnings.

The "syntactic thesis" of the Cantorian explanation consists in identifying, as necessary condition of the paradox, the conditional axiom ($\forall F \forall G$ ($est(F) = est(G) \rightarrow \forall x\ (Fx \leftrightarrow Gx)$)) which is the left- to-right conditional contained in Basic Law V: the existential generalization of this principle ($\exists \iota \forall F \forall G\ (\iota F = \iota G \rightarrow \forall x(Fx \leftrightarrow Gx))$), constituting the negation of Cantor's Theorem, determines an original contradiction in the fregean system. On the contrary, we can reject this thesis showing that (Paseau 2015) the same contradiction can also be derived from a weaker principle (*Definable*-BLV: $\forall F \forall G\ (\forall x\ (Fx \leftrightarrow \phi x) \rightarrow (estF = estG \leftrightarrow \forall x\ (Fx \leftrightarrow Gx))))$ that is consistent with Cantor's theorem.

The "semantic thesis" of the Cantorian explanation consist of interpreting the contradiction between Basic Law V and Cantor's Theorem as the unsatisfiable conjunction of two incompatible requests about the cardinality of the domains (according to Cantor's theorem, the cardinality of the first order domain must be less than the cardinality of the second order domain; according to the reading from left-to-right of Basic Law V, the cardinality of the first and second order domains must be the same). However, we can refuse this thesis in its full generality beacause it is true only in standard models.

The "syntactic thesis" of the Predicativist explanation consists in identifying, as necessary condition of the paradox, an istance of impredicative comprehension's axiom ($\exists F \forall x\ (Fx \leftrightarrow \exists G\ (x= estG \wedge \sim Gx))$). By "impredicative" comprehension axiom, I mean a schematic comprehension axiom whose comprehension formula contains bounded second-order variables (in the scope of a quantifier or of an abstraction operator). The "syntactic thesis" of the Predicativist explanation is correct, altough incomplete because second order logic (with impredicative comprehension axiom and without Basic Law V) is consistent.

Also the "semantic thesis" of Predicativist explanation is partially confirmed: the contradiction is false in each model, so it is obviously false in each model in which the russellian istance of comprehension axiom is false.

Then an istance of impredicative comprehension axiom is one but not the only one necessary condition for the purpose of derivation.

The "syntactic thesis" of the Extensionalist explanation consists in identifying, as (implicit) necessary condition of the paradox, a theorem about the Existence of Extensions (Zalta 1998) ($\forall F \exists x\ (x= estF)$) which states that the correlation between concepts and objects (extensions) is elsewhere defined in the domain of the concepts. This theorem can be derived in any system in which we

introduce a term-forming operator (i.e. extensional operator – est -) into second order logic with identity. In Frege's system this theorem is implicated from Basic Law V and from a general Principle of Extensionality ($\forall F \forall G$ (est(F) = est(G) $\rightarrow$ ($\forall x$ (Fx $\leftrightarrow$ Gx) $\rightarrow$F = G))) which constitutes a necessary premise for the derivation of Russell's contradiction. The "syntactic thesis" of the Extensionalist explanation is correct, altough incomplete because the derivation of Russell's paradox is impossible without an istance of impredicative comprehension axiom.

The "semantic thesis" of the Extensionalist explanation says that the russellian contradiction is due to an unsatisfiable correlation between entities of the second and first order domains. Also this thesis seems to be incomplete because this request is unsatisfiable only in conjunction with an impredicative istance of the comprehension axiom.

In conclusion, the best "explanation" of Russell's Paradox identifies as necessary premises both the impredicative comprehension axiom and Basic Law V. Neverthless, this last principle is problematic not (as in the Cantorian Explanation) because of the iniectivity of the extensional correlation between concepts and objects or the cardinality request over the domains but because of the functional feature of that correlation.

Internal/External Logical Pluralism

Bogdan Dicher (LanCog, Universidade de Lisboa)

14.09, 16:45 – 17:20, Room: Sala de Actos

Logical pluralism is the view that there are several equally correct accounts of logical consequence.

In this paper I discuss a thus far neglected route into logical pluralism, grounded in a familiar phenomenon in substructural logics: the possible mismatch between the internal and external consequence relations of logic.

According to Tarski, a consequence relation (hereafter CR) is a reflexive, transitive, monotonic relation between sets of formulae (the premises) and a single formula (the conclusion).[3]

There are many well known departures from this standard Tarskian definition; these are best presented in the context of Gentzen's sequent calculus. Using sequents, the properties of logical consequence can be internalised by way of structural rules.[4] For instance, the reflexivity of a CR is represented within the calculus by the zero-premise rule Identity. Monotonicity is expressed by Weakening and transitivity by Cut.

Given a sequent calculus, there are two ways to associate a CR to it.[5] One is based on reading sequents as consequence claims. Thus a formula A is a consequence of a collection of formulae X iff the sequent X: A is derivable in the system. This is the internal CR of the calculus. Since the structural rules and properties of sequent calculi can diverge from one logic to another, we can get deviations from the Tarskian concept of consequence. For instance, without Contraction, it matters how many occurrences a formula has in the sequent. By dropping Weakening monotonicity is affected.[6] Cut-less CRs are nontransitive.[7]

[3] Alfred Tarski. "On the Concept of Following Logically". In: History and Philosophy of Logic 23.3 (2002). Originally published in 1936, pp. 155–196.
[4] A. S. Troelstra and H. Schwichtenberg. Basic Proof Theory. 2nd ed. Cambridge University Press, 2000.
[5] Arnon Avron. "The semantics and proof theory of linear logic". In: Theoretical Computer Science 57 (1988), pp. 161–184; A. S. Troelstra. Lectures on Linear Logic. csli Publications, 1992.
[6] J. Michael Dunn and Greg Restall. "Relevance Logic". In: Handbook of Philosophical Logic. Ed. by Dov M. Gabbay. Second. Vol. 6. Kluwer Academic Publishers, 2002, pp. 1–136; Francesco Paoli. "Implicational Paradoxes and the Meaning of Logical Constants". In: Australasian Journal of Philosophy 85.4 (2007), pp. 553–579.
[7] David Ripley. "Conservatively Extending Classical Logic with Transparent Truth". In: Re- view of Symbolic Logic 5.2

The other way of associating a CR with a sequent calculus has it that A is a consequence of X if and only if the sequent (: A) is derivable in the calculus together with the extra axioms (: Bi), for each Bi ∈ X (parentheses added for readability). This is the external CR of the calculus. Unlike the internal CR, this is (usually taken to be) property invariant: it is always reflexive, monotone and transitive. The choice between them is of little consequence for say, classical or intuitionist logic, where these two relations are (extensionally) equivalent. But it is quite important for other logics, such as linear logic[8] or the non-transitive logic ST,[9] where they come apart.

In this paper I will argue that, when it occurs, the mismatch between the internal and external consequence relations of sequent calculi is indicative of a non-trivial form of logical pluralism.

In the final part of the paper I will compare this form of pluralism about logical consequence with other pluralist tenets developed in the literature, in particular the meaning-invariant pluralism of Beall and Restall and the intra-theoretical pluralism of defended by Hjortland.[10]

A Defense of Robust Virtue Epistemology

Pedro Dinis (LanCog, Universidade de Lisboa)

13.09, 16:45 – 17:20, Room: Sala de Actos

Robust Virtue Epistemology (RVE) is a version of Virtue Epistemology (VE) that maintains that knowledge is primarily explained by agent's cognitive ability. According to RVE, agent S knows that p if and only if S's cognitive ability plays a primordial role (or a sufficiently relevant role) obtaining S's true belief that p. This primordial role played by S's cognitive ability makes it a sufficient condition for knowledge.

Several objections to RVE attacked its sufficiency and necessity for knowledge. Pritchard has been one of those critics, suggesting a moderate or weak version of VE, according to which S's cognitive ability is a necessary, but not sufficient, condition for knowledge. Pritchard adds a safety condition as necessary for knowledge, which has a modal character: in most possible near worlds to the actual world in which S forms her belief in the same way, S would continue to obtain a true belief. This definition could be simplified to: it is not the case that S would easily believe in the falsity of p.

One of the objections to the sufficiency of RVE for knowledge is the well-known BARN FAÇADE CASE: S has the true belief that a barn is in front of her. But unbeknownst to her, she is in a place of false barn façades, so that she can't distinguish a true barn from a false one. In this case, S's perceptual ability apparently plays a sufficiently relevant role in obtaining S's true belief that p. This is because S's cognitive success seems to be attributed to the exercise of her perceptive capacities just as it would be in a situation without false barns all around. However, we do not say that S knows that p. And that poses a problem to RVE.

The main attacks on RVE necessity for knowledge are cases of knowledge by testimony (see for example CHICAGO VISITOR by Lackey (2007); Pritchard (2010, 2012) and by extended cognition (see for example SISSI CASE by Vaesen (2011)). It is argued in these cases that S knows that p, but S's cognitive ability to obtain the true belief that p did not play a sufficiently relevant role. This role

(2012), pp. 354–378.

[8] Jean-Yves Girard. "Linear Logic". In: Theoretical Computer Science 50 (1987), pp. 1–101. 7

[9] Ripley, "Conservatively Extending Classical Logic with Transparent Truth".

[10] Ole Thomassen Hjortland. "Logical Pluralism, Meaning-Variance, and Verbal Disputes". In: Australasian Journal of Philosophy 91.2 (2013), pp. 355–373.

was instead played by those who transmitted the information.

For those cases that attack RVE sufficiency for knowledge I propose an evaluation of cases in terms of the reliability of S's cognitive ability in the environment in which S obtains p. By the reliability of S's cognitive ability in environment x I understand the following: S's cognitive ability is reliable in environment x if and only if S's true belief that p depends more or to a greater degree on S's cognitive ability than environmental (or interventional) luck. If S's true belief depends more on her cognitive ability than on luck, S doesn't know that p. In the BARN FAÇADE CASE, luck plays a more relevant role than S's cognitive ability obtaining S's true belief that p. And this is so because in the possible near world to the actual world S would easily believe in the falsity of p. By this way, RVE could accommodate the safety condition without having to require it as a necessary condition for knowledge. It is sufficient that S's true belief depends more on her cognitive ability than luck.

For cases that attack the necessity of RVE for knowledge I propose the idea of shared or jointly ability. Cases of knowledge by testimony and extended cognition are cases of social knowledge, cases in which knowledge obtained does not depends only on S, but depends also on other(s) agent(s). My proposal is that in these kind of cases S's cognitive performance cannot be evaluated in isolation, but as part of a shared or jointly ability.

An ability y is shared if and only if S's true belief that p depends more on actor's cognitive ability than from luck. A shared ability is one that plays a sufficiently relevant role in S's cognitive success. Each actor belonging to the shared ability make an appropriate contribution to S' cognitive success, even if a given contribution viewed in isolation may seem a poor or insufficiently relevant performance. But because it is part of y, that is, because it is also responsible for meeting S's cognitive success, I understand each one of these contributions as adequate (for that purpose).

The case presented by Lackey, CHICAGO VISITOR, is a case of shared competence. Morri's cognitive success depends more on actors cognitive ability than on luck. The same applies to the SISSI CASE presented by Vaesen, in which there is also shared competence. Cases in which S obtains jointly cognitive success about p, but where there is no shared ability, are cases in which S does not know that p. In these cases, luck primarily determined S's cognitive success.

From my proposal, I conclude that RVE could give a good reply to the objections attacking RVE sufficiency and necessity for knowledge.

Transcendental Arguments and Pritchard's Insight

Casey Doyle (University of Oxford)

14.09, 10:10 – 10:45, Room: Mattos Romão

This paper proposes that two quite different responses to scepticism in a particular domain are, surprisingly, made for each other. The basic idea is that each style of response can seem unsatisfying, by its own lights, and that that this dissatisfaction can be removed by conjoining the two responses.

In "Does Moral Philosophy Rest on a Mistake?", H.A. Prichard argued that a certain form of scepticism about moral obligation, one that finds expression in the question "why be moral?" is confused. The question demands a universally acceptable and non-question-begging reason for obeying moral requirements. Such a reason, it would seem, would have to come from outside of morality. Prichard charged that this is a mistake: the normativity of morality does not come from outside of it.

Even those who are sympathetic with Prichard's point might find themselves unsatisfied with leaving things here. That is, even if you agree that it is a mistake to ask for a non-moral reason to follow moral requirements, one might reasonably ask for some form of understanding our commitment to those requirements. Even if it is illegitimate, the sceptical question grips us, and at least one reason for this is that we are committed to the claim that one ought to act morally. And one might reasonably ask for some understanding of that commitment.

A transcendental argument seeks to vindicate a certain basic commitment (such as that there is a mind-independent world, that there are other minds, that there are moral facts, etc.) by showing that it is a necessary precondition on a form of cognition which it is safe to take for granted. Since Stroud's seminal work, it has been widely accepted that while such an argument can show that creatures like us necessarily or inescapably have those commitments, it cannot prove that those commitments constitute knowledge. It might be that we can't but take there to be other minds, say, but that doesn't show that we are justified in believing that there are other minds. Here too we have an anti-skeptical strategy that leaves us wanting more.

My suggestion is that a transcendental argument can provide the understanding that Prichard's move leaves us wanting. Such an argument can reveal the source of a particular commitment as grounded in something other than a reason that is external to the relevant domain. And Prichard's diagnosis of where the skeptic goes wrong reveals that the request for proof which the transcendental argument fails to deliver may not be needed.

Transcendental arguments and versions of Pritchard's response can arise in response to sceptical challenges in different areas, The paper's modest aim is to gesture toward the advantages of marrying these two responses in the way outlined. However, I believe the response is especially helpful in thinking about response to the "why be rational?" question that has been central in the literature on normativity. After outlining the general proposal, the paper will consider how it applies to that case.

Cognitive Propositions and Kripkean A Posteriori Necessity: Prioritizing Linguistic Meaning over Semantic Content

Ainhoa Fernández (University of the Basque Country)

14.09, 10:10 – 10:45, Room: Sala de Actos

One of the functions of the propositions is to be the meanings of sentences. This implies that, depending on which notion of proposition is being defended; different analyses can be made of the necessary or contingent character of certain statements, as well as offering different explanations as to whether we can know their semantic content a priori or a posteriori. Kripke (1971, 1972/1980) distinguishes four types of statements: those necessary a priori (analytical statements), those necessary a posteriori, those contingent a priori and those contingent a posteriori. Thus, the concepts of the necessary, the a priori and the analytical are not extensionally equivalent as upheld by logical positivists and Quine (1960). So, according to the Kripkean classification, there are necessary statements which can be a posteriori. The canonical case presented by Kripke is the statement (1)

(1) Hesperus is Phosphorus

Soames (2006) argues that (1) is not a genuine case of necessary a posteriori, showing how its semantic content is knowable a priori. However, according to Soames, this argument does not affect other statements of a posteriori necessity or statements that include natural kind terms, such as (2)

(2) Water is H2O

In Soames' argument against (1) as a genuine case, a rudimentary, neutral notion of proposition is employed. More recently, Soames (2010, 2014, 2015) defended a theory of propositions whereby propositions are abstractions of types of cognitive acts. In this new position, some propositions have incorporated modes of presentation which are not Fregean but Millian; as they are not present in all the propositions nor are they representatively relevant, although they do cause the propositions to which they are associated to be cognitively different to one another. Incorporating this non descriptive notion of mode of presentation, Soames has offered a different explanation of statements such as (1) and (2). I shall argue that the new explanation offered by Soames is unsatisfactory and shall suggest a different solution that prioritizes linguistic meaning over semantic content, allowing us to adequately account for the necessary a posteriori character of this sort of statements.

Let me explain the new Soames' arguments. In the case of (1), considering the following statements

(1) Hesperus is Phosphorus
(1') Hesperus is Hesperus

'Hesperus' and 'Phosphorus' are names denoting the same object, namely Venus. If we consider an utterance of (1) in a context in which two speakers share the presupposition that both understand the names, the speaker asserts not only the bare singular proposition expressed by (1), but also the linguistically enhanced proposition entertainable only by those who identify the predication targets via the two names. The proposition semantically expressed by (1) and (1') is the same since the two proper names stand for the same object and the only contribution that a proper name makes to the sentence in which it appears is the object itself. This proposition is necessary but knowable only a priori. However, the linguistically enhanced proposition through modes of presentation is necessary but only knowable a posteriori.[11]

But, if we ask whether (1) and (2) are necessary a posteriori statements, what answer do we get from the cognitivist theory of propositions that Soames defends? If we ask whether the proposition semantically expressed by (1) and (2), respectively, is necessary a posteriori, the answer is no. The answer is no, the proposition semantically expressed by (1) and (2), respectively, is not necessary a posteriori, although another proposition associated with it, the linguistically enhanced one, is indeed. This answer offered by Soames is unsatisfactory because it is not a direct answer to the question. How can we get a straightforward answer to the question of whether (1) and (2) are necessary a posteriori? Soames' own theory, with some changes, can provide that answer.

In Soames' theory, propositions are abstractions of cognitive types of predication, in this case linguistic acts. When we perform an act of predication, any act of speech, we refer to the object or objective of predication in a certain way and, if propositions are what Soames says they are, and ultimately depend on the cognitive capacity of the agents performing such acts, then the mode of linguistic presentation must be included in the abstraction of the act, in the corresponding proposition. That is to say, since the mode of presentation provides the distinctive element of a

[11] The same argument applies to the case of the statement (2).

given proposition where modes of presentation intervene, it is crucial that it be reflected in the extracted proposition. What Soames calls the semantically expressed proposition can be extracted from different statements, where different modes of presentation intervene and therefore make a difference between one proposition and another. This semantically expressed proposition takes up the classical notion of semantic content, but not linguistic meaning, which is at stake when it comes to modes of presentation. Linguistic meaning is what imposes conditions for the successful communication and understanding of speakers. Semantic content only serves, in this analysis, to justify the a priori character of a proposition, which is unnecessary if the notion of linguistic meaning is prioritized over that of semantic content, since the semantically expressed proposition contains the modes of presentation of the object. So, if what we want to explain is the successful communication and understanding of speakers, we must prioritize linguistic meaning over semantic content in the analysis of propositions involving modes of presentation. And, if in cases such as (1) and (2), we prioritize the notion of linguistic meaning and not the notion of semantic content, since it is the former that determines the a posteriori necessity of both statements, we can give a direct answer to the question at hand. Thus, the semantically expressed proposition must be what Soames qualifies as "linguistically enhanced". In this way, the proposition drawn in the first instance from the act of predication, in this case asserting (1) and (2) respectively, is necessary a posteriori.

Alethic Pluralism and the Value of Truth

Filippo Ferrari (University of Bonn)

13.09, 15:05 – 15:40, Room: 5.2

Is truth valuable? Some philosophers (e.g. Gibbard 2003, Horwich 2013, James 1942, Lynch 2009, Wright 1992) have thought that there is a direct connection between truth and value: in many cases, having true beliefs is a good thing and, overall, a better thing than having false beliefs. This seems hardly disputable. However, as it is customary for many philosophical topics, the question whether truth is valuable requires some clarifications.

First, philosophers have expressed the thought that truth is valuable in a variety of ways. Some have claimed that truth is what we aim at in pursuing enquiry (e.g. Lynch 2009). Others maintained that truth is what we ought to believe (e.g. Horwich 2013). Following James (1942) some philosophers have claimed that truth is the good in the way of belief. Still others have taken truth to provide the mark of correctness for beliefs (e.g. Gibbard). As I have argued elsewhere, these are questions concerning distinct aspects of truth's normative profile. Here, I will be exclusively concerned with what I call the axiological aspect of truth's normative profile – i.e. the question whether truth is a kind of good and whether believing what's truth is valuable. Second, some clarifications concerning the meaning of expressions such as 'truth is good' and 'believing what is true is valuable' are called for. In particular, we should specify the type of value involved in axiological principles: is it instrumental, or final? Is it intrinsic or extrinsic? Is it conditional or unconditional? (see @, Wrenn 2014). Philosophers have given different answers to these questions. Most of them were, however, assuming a monistic approach to truth – i.e. the idea that truth is one and has a uniform nature regardless of variations in subject matter. The investigation gains an additional level of complexity if, as some philosophers do, we drop this monistic assumption and endorse the pluralist thesis that there is a varieties of truth properties, each of which with its own distinctive nature.

Thus, the aim of this paper is to address the question: how should we conceive of the value of

truth within a pluralistic framework? After having introduced and motivated the aforementioned distinctions in the axiology of truth, the paper discusses whether and to what extent the two main pluralist models about truth – moderate (e.g. Edwards 2018, Lynch 2009, Pedersen & Wright 20013) and strong pluralism (e.g. Wright 1992) – are able to account for the thought that truth is valuable.

I argue that truth should not be conceived as having a uniform and invariable axiological nature. My argument for this conclusion relies on the thesis, defended elsewhere, that the occurrence of disagreement has a variable normative significance in relation to the subject matter of disagreement. Call this the normative variability conjecture. The explanation of this variability concerns the fact that truth is assigned with a variable axiological significance in relation to the subject matter at issue. The thought is that the degree of rational tolerance which is, arguably, appropriate to have vis-à-vis the occurrence of disagreement about topic X depends on how deeply subjects typically care about the project of getting to the truth about X. Typically, subjects care much less about getting to the truth of things about, e.g., basic taste – about whether oysters are delicious – than about getting to the truth of things concerning, e.g., fundamental morality – about whether death-penalty is morally admissible – or about political economy – about whether austerity policies stimulates economic growth. On the basis of these considerations I conclude that if truth has an axiological nature, it shouldn't be conceived as uniform and invariant across all subject matters.

I then argue that all current forms of moderate truth pluralism are structurally unable to allow for the required variability in truth's axiological nature. This is because of two core commitments that all forms of moderate pluralism on the market share. The first commitment is that truth is metaphysically both one and many, meaning that there is a generic truth property that all true propositions have in virtue of possessing a domain-specific truth property. The second commitment is the idea that truth is a normative property with an axiological nature. Regardless of whether we conceive of truth's axiological nature in 'light' terms or 'heavy' terms, moderate pluralists won't be able to provide a satisfactory explanation of the normative variability conjecture. This gives us a principled reason to reject all the models of moderate pluralism currently on the market.

In the last part of the paper I argue that strong pluralists about truth don't face the problem affecting moderate pluralism due to the fact that the metaphysical structure of strong pluralism allows for much more flexibility than its moderate counterpart. I then sketch two models for accounting for the value of truth that are fully compatible with strong pluralism. Both models take truth to be finally and conditionally valuable. However, the first model takes the source of the value of truth to be intrinsic to the nature of the various admissible truth properties with the proviso that each truth property has a different axiological nature so that to account for the normative variability conjecture. The second model takes the value of truth to be grounded entirely in factors that are extrinsic to the nature of truth – factors that have to do with non-truth-theoretic aspects of our investigative practices. This would make truth's nature free of any axiological feature without, however, depriving it of its axiological significance. The variability conjecture would then be accounted for by the fact that our investigative practices substantially differ in relation to the subject matter at the core of our investigation. Among these differences there's also a variation in the degree in which we care about getting to the truth of things – a variation that goes in tandem with variation in the nature of the subject matter.

I conclude by arguing that the second axiological model should be preferred to the first model because compatible with a form of truth pluralism that encompasses among the various admissible truth properties also a non-normative truth property – I have argued for the possibility of a normatively deflated truth property elsewhere.

Revising Logic: Anti-Exceptionalism and Circularity

Maria Paola Sforza Fogliani (University School for Advanced Studies, IUSS Pavia)

14.09, 17:20 – 17:55, Room: Sala de Actos

According to anti-exceptionalism (AE) about logic, (*i*) logical laws do not possess the series of properties that have historically been thought to make the discipline to some degree 'special' – *viz., aprioricity* (or *unrevisability*), *analyticity*, *necessity* and *normativity*; rather, logic is in many senses continuous with science. Also, (*ii*) theory-choice in logic is to be carried out by means of an abductive methodology; that is, logical theories are justified, revised and com- pared with respect to a set of traditional criteria – *e.g.*, *adequacy to the data*, *simplicity*, *strength*, *elegance* and *fruitfulness*.

We'll first try to clarify claim (*i*), by reviewing which properties AEs think logical laws should be deprived of. It will turn out that there is a substantial disagreement on what logic (allegedly) cannot be, the only agreed upon feature being non-apriorism; furthermore, it seems that AEs use '*a priori / a posteriori'* in an unspecific sense – in that they do not make reference to empirical kinds of evidence, but rather equate non-aprioricity with revisability.

We'll then move on to (*ii*), and try to unpack the abductive methodology and its criteria. In order to do this, we will first review the main implementations of the AE model – namely, Priest's [2016] and Williamson's [2017].

Priest provides a formal model for theory-choice in logic. Consider a set of criteria $\{c_1, ..., c_n\}$ and a set of theories $\{T_1, ..., T_n\}$. Let an evaluation scale of a theory T with respect to a criterion c be the set $X = \{x \hat{I} \mathbb{R} | -10 < x < +10\}$; then for every criterion and theory there is a measure function $\mu_c(T) \hat{I} X$. Since not all criteria might be equally important, we assign a weight to each of them – i.e. $w_c \hat{I} X$. The rationality index of a theory T is thus defined as:

$$p(T) = w_{c1}\, \mu_{c1}\, (T) + ... + w_{cn}\, \mu_{cn}\, (T)$$

Among a set of competing theories, the one displaying the highest rationality index will be the rationally preferable.

Williamson's account focuses on consequence operators (*Cn*). Let Γ be the set of sentences of a well-confirmed theory (e.g., solid laws of physics), and let a consequence operator Cn_1 for a consequence relation $\vDash_1$ and theory Γ be defined as:

$$Cn_1 (\Gamma) = \{A \mid \Gamma \vDash_1 A\}$$

Cn_1 takes sentences of Γ as inputs, and outputs Cn_1 (Γ) – i.e., the set of sentences that follow from Γ via $\vDash_1$. The model compares consequences relations with respect to how well they do when applied to well-confirmed theories; e.g., a consequence operator yielding, from premises we have grounds to believe, a sentence we have grounds to reject, will count as evidence against that consequence operator.

Secondly, we'll take a closer look at the abductive criteria, and in particular at the prominent among them – namely, adequacy to data. The so far most detailed account has been presented by Priest, and has it that data for the assessment of a logical theory are provided by our pre-theoretical intuitions about the validity of natural language inferences.

Despite providing a sensible account of theory-change in logic, I will argue that AE has to face a host of objections that stem from a well-know argument in the philosophy of logic – the Centrality

Argument (CA; e.g., Putnam [1978]) – which, in a nutshell, runs as follows: logical laws are so central in every rational reasoning that any attempt either to revise or to justify them ends up using those laws themselves and, so, winds up being circular or otherwise illegitimate. We can build versions of CA that are specifically targeted against the AE accounts; moreover, since logical laws are extremely pervasive in reasoning, we expect circularity issues to arise at several different levels of the abductive computation.

Let us start with justification, and consider Priest's implementation of the adequacy to data criterion:

P1. If we have favorable intuitions for a logical law *l*, then *l* is justified;
P2. we have favorable intuitions for *l*;
– – – – – – – – – –
C. *l* is justified

Obviously, this is a *modus ponens*; so, if l happened to be *modus ponens* itself, our justification would be circular. As for revision, let us apply CA to Williamson's *Cn*-account:

P1. If Cn_1 (Γ) includes false sentences, *l* has to be revised;
P2. Cn_1 (Γ) includes false sentences;
– – – – – – – – – –
C. *l* has to be revised.

Suppose, moreover, that we could tell the issue with Cn_1 (Γ) was specifically caused by *modus ponens*; again, our revision would be circular. Furthermore, we can device a dull version of the argument that attaches – at a higher level – to the general AE inference:

P1. If $p(L_1) > p(L_2), p(L_3) \ldots p(L_n)$, then L_1 is the best theory among the set of logics L
P2. $p(L_1) > p(L_2), p(L_3) \ldots p(L_n)$
– – – – – – – – – –
C. L_1 is the best theory among the set of logics L_1, L_2 ... L_n.

In case L_1 does not validate *modus ponens*, while its rivals do, we might be in for a clash.

Thus, we appear to be left with two conflicting cases about the revisability of logic: logical theories have indeed often been revised on the basis of an abductive methodology, which has taken into account their adequacy to our intuitions, their strength and their simplicity; however, the AE account faces some serious threats of circularity.

I will end by proposing a way of reconciling these seemingly opposing in- tuitions, which appeals to Priest's [2014] distinction between *logica docens* and *logica ens* – that is, between what logicians claim about logic, and what is actually valid. I will argue that AEs seem to submit only that *logica docens* is revisable, while remaining silent on *logica ens*' fate; on the other hand, a minimal version of CA shows only that we can neither revise nor justify the laws of the correct logic – i.e., of *logica ens* – whatever this logic is. I'll thus submit that some compatibility can be worked out, at least between modest versions of the two opposing positions.

Self-Knowledge through Bypass? Jordi Fernández on the Transparency of Mind

Martin Fricke (Universidad Nacional Autónoma de México)

14.09, 14:30 – 15:05, Room: Sala de Actos

In his recent book Transparent Minds (2013), Jordi Fernández aims to give an account of self-knowledge regarding beliefs and desires which is inspired by Gareth Evans's dictum that "in making a self-ascription of belief, one's eyes are, so to speak, or occasionally literally, directed outward – upon the world. [...] I get myself in a position to answer the question whether I believe that p by putting into operation whatever procedure I have for answering the question whether p." (Evans 1982: 225) According to Fernández, Evans's observation rings true because we form our second-order beliefs not on the basis of the first-order beliefs to be ascribed but, rather, by attending to the mental states that ground those first-order beliefs. Fernández calls this claim "the bypass view":

> For any proposition P and subject S:
> Normally, if S believes that she believes that P, then there is a state E such that
> (a) S's (higher-order) belief has been formed on the basis of E.
> (b) E constitutes grounds for the belief that P in S. (Fernández 2013: 49)

When I am trying to answer the question of whether I believe that there is an apple in front of me I might look at what is in front of me or try to remember whether there was one the last time I looked or listen to the testimony of other people about what is in front of me (cf. Fernández 2013: 50). In each case I attend to the mental states that might ground the belief that there is an apple in front of me. If I find myself in such a mental state (not the first-order belief itself), I can directly form the higher-order belief which ascribes that first-order belief. The formation of the second-order belief bypasses the first-order belief.

Fernández further argues that Bypass is a good procedure because we form all our beliefs on the basis of other mental states that regularly cause them as grounds. Fernández summarises these regularities in various "production-of-belief principles", such as "If S apparently perceives that P, then S comes to believe that P" or "If S believes that S* is providing her with the information that P, then S comes to believe that P" (Fernández 2013: 46). Every such principle describes an antecedent mental state that regularly leads to the production of a corresponding belief. Fernández argues that it is in virtue of such causal relations that we can use the Bypass procedure and self-ascribe a belief simply on the basis of the mental state that normally causes it, thereby bypassing the belief itself.

In my talk I shall discuss four objections to Fernández's theory. These can be divided into two objections that concern the normative aspect of the theory and two that deal with the metaphysical aspect.

Regarding the normative aspect, Matthew Boyle's criticism of a different version of transparency theory (that due to Alex Byrne [2005]) seems relevant to Fernández's theory as well. Boyle argues that an inference from a fact about the world ("p") to a fact about one's beliefs ("I believe that p") is mad; and if I find myself making such an inference I should discard the second-order belief produced by it because it is not well-founded. Although Fernández denies that Bypass describes an inference, the objection still seems relevant. On the basis of a mental state that is about the world (e.g. seeing that p), I am supposed to form the belief that I believe that p. But, Boyle claims, "a modicum of rational insight will inform me that, even if it is true that P, this by itself has no tendency to show that I believe it" (Boyle 2011: 230). The seriousness of this objection depends on how far we wish to go along with the internalism inherent in Boyle's criticism.

The second objection concerning the normative aspect of Fernández's theory considers cases in which the above-mentioned production-of-belief principles break down because the subject makes an error. If the subject has a habit of making this type of error, there is still a regular (irrational) causal connection between the mental state and the first-order belief formed on its basis and Fernández claims that this means following Bypass is justified. – However, what about a fluke error? In this case, the subject normally avoids the relevant fallacy, but only on one occasion she does not. This means that there is no causal regularity connecting the mental state with the belief arrived at. In consequence, the self-ascription of belief resulting from following Bypass could not be justified by such a regularity. Fernández endorses this conclusion (cf. Fernández 2015: 150); but it is unclear whether this is plausible.

Regarding the metaphysical aspect of the theory, it might be doubted that Bypass can reveal our pre-existing beliefs. As Brie Gertler (2011) has argued, by considering what reasons we have for some first-order belief, we might first come to form a belief about the matter or we might come to change pre-existing beliefs. The resulting belief-ascription might be true, but it would not reveal what belief we had anyway. Fernández can probably answer this objection by pointing out that Bypass can also self-ascribe a belief on the basis of a simple memory that p.

A second objection to do with the mechanics of the Bypass procedure concerns changes in the way we form our first-order beliefs. Can the procedure detect such changes? Suppose today I affirm the consequent, while tomorrow I refuse to do so. (This might be a permanent change of reasoning or a fluke change.) It seems that for the Bypass procedure to be sensitive to such changes it has to be sensitive not only to the mental state itself that grounds the first-order belief, but sensitive to the direction in which the mental state pulls the subject to form a first-order belief. It is not enough, for example, for the procedure to "take note" of (to base the self-ascription on) the fact that the subject believes that if p then q and that q. It must also "take note" of (base the self-ascription on) the fact that by these beliefs the subject is pulled towards the belief that p (if it is of a fallacious mind) or not (if it has reformed its reasoning). In my talk, I shall argue that this means that the procedure becomes rather similar to a procedure that bases self-ascriptions directly on the first-order beliefs to be ascribed, such as Byrne's rule BEL (cf. Byrne 2005).

Vulgarity, Pejoration, and Slurs

Aldo Frigerio and Maria Paola Tenchini (Università Cattolica del S. Cuore di Milano)

13.09, 14:30 – 15:05, Room: 5.2

Recently, the semantic status of slurs has been much and vividly discussed in analytic philosophy of language and linguistics. Obviously, the interest in this issue has been partly determined by the social and political significance of slurs in everyday life because the use of such expressions seems to involve the speaker's adhesion to discriminatory practices towards minority groups. The primary interest in slurs has contributed to obscure other expressions whose semantics is similar to slurs', which nonetheless are not slurs. This paper aims to clarify the semantics of such expressions in order to take apart two components which are usually confused, namely the derogatory dimension on the one side and the vulgar dimension on the other side. This step is necessary because there are pejoratives (and slurs) that are not vulgar and vulgar expressions that are not pejorative (and, thus, not slurs either). The need of this distinction is clear if we consider Hom's analysis of an expression like fucking. In fact, he treats this expression like a pejorative, i.e. like an expression belonging to

“the class of expressions that are meant to insult or disparage” (Hom 2012). Nonetheless, it has been noticed that fucking can be associated with absolutely positive evaluations, such as

(1) This is fucking good
(2) He is a fucking genius

It is quite hard to believe that, at least in sentences like (1) and (2), fucking serves to insult or disparage somebody or something. The same applies to different swear/bad words:

(3) Oh shit, I’ve lost my keys
(4) Oh fuck, this is an amazing party

Again, the swear words in (3) and (4) do not seem to disparage anybody. Even though we would define the expressions in (1)-(4) as vulgar, not every derogatory expression is vulgar, for example:

(5) George is completely stupid

(5) denigrates George but actually it is not a vulgar sentence. The same is true for many slurs. In particular, it seems to be true for the word that is considered the slur par excellence (“nigger”). Certainly, this is a strongly derogative word; nonetheless, in our opinion, it is not a vulgar word. Many of these expressions are often defined as “expressive”. For instance, Pott (2007) calls “expressives” words like “damn” and “bastard” and claims that “a speaker’s expressives indicate that she is in a heightened emotional state”. There is no doubt that one of the functions of these words is the expression of a heightened emotional state. Yet, this does not capture the whole semantics of such words, as the speaker can express her upset state without using pejorative or swear words. Once again, the current analyses fail to take into account the vulgarity component of the meaning of these expressions. Our proposal is to distinguish between denigration and vulgarity. Derogatory expressions convey a negative judgement or attitude by the speaker against somebody or something. Some of these expressions are also vulgar, but, as we have seen, not all of them are. The following derogatory expressions are also vulgar:

(6) Paul is an asshole
(7) John is a bastard

However, not every derogatory expression is vulgar, as it is demonstrated by (5). Slurs are a particular sub- class of derogatory expressions, in which the negative judgement about the target person is grounded on her membership to a particular social group and, thus, the whole social group is involved in the negative judgement. Slurs too can be vulgar or not. For instance, the following slurs are vulgar:

(8) Ann is a bitch
(9) John is a faggot

On the other hand, we have underlined that many slurs are not vulgar, in particular, those that have nothing to do with sex, but exclusively with the color of the skin or with ethnicity. On the basis of

these distinctions, we will put forward the following points:

• Vulgarity is always projected. Even though the vulgar term is embedded into operators or contexts of various kinds, the sentence that contains it is still vulgar.

• Vulgarity engenders expressivity and, therefore, when a vulgar pejorative is used, this intensifies the force of the insult directed at a person (to say "asshole" to a person is more insulting than to say "idiot").

• However, expressivity is only a consequence of vulgarity, not an independent dimension. Vulgar pejoratives are expressive because they are vulgar (not the other way around).

• Swear words are vulgar terms. When used alone ("shit!", "fuck!"), they are expressive because of their vulgarity. They can be offensive not because they denigrate somebody, but because they can damage negative politeness (cf. Brown and Levinson, Politeness, 1987). Obviously, the damage depends on the context and on interlocutors. For example, it is more inappropriate and annoying to use these words in a formal context than in an informal one.

• In light of this, in dealing with these expressions, it is worth taking apart two kinds of offence that are rarely distinguished: 1) the offence arising from the use of a pejorative directed towards a person ("idiot", "bastard"). In this case, offence arises from denigration; 2) the offence arinsing from a damage to negative politeness, when swear words are used ("oh fuck!", "shit!").

• Vulgarity has to do with the reference to certain taboo domains (sex, defecation) and, thus, with negative politeness. However, not every reference to these domains is vulgar because there are non- vulgar terms referring to them: cf. "penis" vs. "dick"; "vagina" vs. "pussy". It is a particular lexical choice that is vulgar. Our aim is to inquire this distinction between vulgar and non-vulgar terms referring to these domains, so to define vulgarity in a precise way.

Singular Reference in Fictional Discourse

Manuel García-Carpintero (LOGOS-BIAP, University of Barcelona)

14.09, 9:35 – 10:10, Room: Sala de Actos

Do ordinary proper names behave when they occur in uses of simple declarative sentences in fictional discourse exactly as they do when they occur in default assertoric uses of the corresponding sentences? Are they *rigid*, assuming with Kripke that they are in the latter case? In this paper I argue that – be they empty or non-empty – they are not, at least when we consider the core case of discourse constituting works of fiction, and also the case, dependent on it, of discourse stating the content of fictions. Kot'átko (2010, 94) argues for the opposite view: "proper names remain *rigid designators* in Kripke's sense ... even if transferred from "everyday" communication to literary texts". Friend (2000, 2011, 2014) has also argued that only such a view can properly capture a certain "object-directness" intuition that we have regarding the content of fictions. They both support the claim I want to confront here, which Kroon (1994) – in a persuasive critical discussion of views of this kind, providing arguments complementary to mine – states thus:

(R) Occurrences in fictional contexts of real proper names like 'London', 'Baker Street', 'Napoleon', and so on, are purely referential and take their usual reference.

I'll outline the main considerations against (R) that I want to develop in the paper. My view adopts from referentialists such as Braun (1993, 2005) the idea that assertions of 'Vulcan is smaller than Mars' have gappy contents. Assertions of atomic sentences with these contents are untrue, false according to the free logics that Braun opts for, neither true nor false according to the one that I prefer. This captures the fact that these assertions are wrong, with respect to a dimension of evaluation (truth) essential to assertions; corresponding remarks could be made about questions or orders with such contents, vis-à-vis their constitutive norms. However, there is absolutely nothing wrong about the acts of fiction-makers who use empty names; there is, for instance, no appearance of "imaginative resistance" on the part of appreciators of such fictions. By placing features accounting for differences in "cognitive significance" between 'Hesperus is smaller than Mars' and 'Phosphorus is smaller than Mars', or 'today is Tuesday' and 'tomorrow was Tuesday' (with the respective contexts of utterance coordinated so that indexicals and tenses have the same referents) at a different level than that of the asserted content – the "ways of believing" of referentialists such as Salmon and Braun, or the presuppositional level my own view posits – we capture the intuitive commonalities in "what is said" among utterances made by people otherwise with very different perspectives on what they talk about, explain communicative success (cf. Perry 2001, 5, 19), and, importantly, account for our reflective intuitions about the objectivity of many subject-matters for our representational acts (Schroeter 2008). A good case can be made that these commonalities extend to straightforward assertions of 'Marlowe is a clever detective' and 'Holmes is a clever detective' by confused speakers who have taken fictional stories for factual ones; the manifest differences in cognitive significance between such utterances would be accounted for in the usual ways.

However, as Lamarque & Olsen (1994) also emphasize, nothing of this sort can be said about the contents that fictions intend proper appreciators to imagine. While the mode of thinking through which we think of Venus when we assert 'Hesperus is smaller than Mars' is intuitively and theoretically irrelevant to what we assert, in that many other modes of thinking about it may do as well, the corresponding modes of thinking "about" Marlowe and Holmes provided by the relevant fictions are essential to their contents: no proper appreciation can ignore them; no proper appreciation can do without building the corresponding files, starting with 'object picked out by the relevant 'Marlowe' naming practice', and stacking into it all the information about the character derived from the fiction. All of this applies equally well to non-empty singular terms occurring in fictions, such as 'Napoleon' in *War and Peace* or 'London' in *1984*.

To sum up, I'll argue that there are no good reasons to support the claim that either 'Bloom' or 'Dublin' behaves like a rigid designator with respect to the content of the utterances constituting Joyce's Ulysses. In the first place, the descriptive content associated with those *names (in particular, person named 'Bloom'/city named 'Dublin' in relation to tokens used in* Ulysses) is *not* intuitively irrelevant with respect to that content; in the second place, it is not intuitively the case that, when we consider counterfactual circumstances to establish whether or not they constitute the contents the fiction ask us to imagine, we just consider how things are with a single Bloom/Dublin.

Ontological Fundamentality

Joaquim Giannotti (University of Glasgow)

14.09, 16:45 – 17:20, Room: 5.2

The notion of fundamentality aims to capture the idea that there is something basic, or primitive in the reality that we inhabit. Metaphorically speaking, the fundamental entities of a possible world are the building blocks of that world. More precisely, the notion of fundamentality captures two ideas: first, that some entities serve as a foundation for everything else that exists; second, that reality has a hierarchical structure and the fundamental entities are those that occupy its bottom. However, there is no consensus on how define fundamentality in more precise terms. A promising view is to conceive of fundamentality in terms of ontological independence. A merit of this view is to capture a natural connection between the notions of fundamentality and ontological dependence. However, it has been recently argued that it is possible that there are fundamental and yet ontologically dependent entities. If so, we should not characterise the fundamental in terms of ontological independence. My aim is to show that the previous possibility does not represent a serious threat for a conception of fundamentality as ontological independence. I will make my case by offering a definition of equifundamentality and showing that ontologically dependent entities can be treated as equifundamental. Let me elaborate these claims in what follows.

There is a natural connection between fundamentality and ontological independence (Lowe 1994; Correia 2008; Koslicki 2013; Tahko and Lowe 2015). At first approximation, to say that an entity x is ontologically independent from another entity y means that x's existence does not require y's existence. For example, if Mary is ontologically independent from Miriam, then Mary can exist in absence of Miriam. My focus will be on accounts that embrace the following biconditional or something in the vicinity: a type-entity x is fundamental if and only if (1) there is no other type-entity y such that x ontologically depends on y, and (2) x is not identical with y. I will call this conception "ontological fundamentality" for short. At heart, the underlying idea is that a type of entity is fundamental just in case does not depend on any other types of entity. A merit of ontological fundamentality is to capture the natural connection between fundamentality and ontological dependence. However, it has been recently argued that there are fundamental and yet ontologically dependent entities (Wilson 2014; 2016: Tahko forthcoming). Such a possibility threatens the plausibility of ontological fundamentality. My aim is to show that it is possible to accommodate the possibility of fundamental and yet ontologically dependent entities in terms of equifundamentality.

I will begin by illustrating the relevance of the notion of fundamentality with respect to philosophical inquiry and scientific theorizing. I will then point out that every satisfactory account should capture an absolute sense of fundamentality and a relative one. Namely, it should regiment respectively the conditions under which an entity can be said to be fundamental full stop and the conditions under which one entity can be said to be more or less fundamental or equifundamental as compared to another one.

The next step is to elucidate the notion of ontological dependence. One might worry that this notion faces a number of objections. Inevitably, some of these objections carry over ontological fundamentality. However, it is possible to invoke the strategies to resist them to defend the tenability of ontological fundamentality. As such, difficulties related to ontological dependence do not hinder the prospects for an account of ontological fundamentality.

I will then offer three definitions that tie together fundamentality and ontological dependence: one for absolute fundamentality, one for relative fundamentality, and another one for equifundamentality.

While these definitions do not help us adjudicating certain metaphysical disputes, they improve the precision of fundamentality claims. After discussing its merits of and providing few examples, I will move on discussing the main objection against an account of ontological fundamentality: the possibility of fundamental and yet ontologically dependent entities. I will argue that the possibility of fundamental and yet ontologically dependent entities can be treated as a case of equifundamentality. Since the possibility of equifundamental entities is unproblematic for ontological fundamentality, the objection is ineffective. Interestingly, there is no need to concoct a far-fetched metaphysical scenario to appreciate the force of this objection.

Tuomas Tahko (forthcoming) illustrates the possibility of fundamental and yet ontologically dependent entities by considering the Standard Model of particle physics. According to it, quarks are types of fundamental particles. For all we know, however, quarks in nature do not exist in isolation. Rather they exist in groups of three: the so- called hadrons such as protons and neutrons. This creates an evident problem for ontological fundamentality: if quarks are ontologically dependent on each other, they then cannot be fundamental. But according to the Standard Model quarks are fundamental!

An obvious strategy would be to reject the fundamentality of quarks. However, this option leaves the conceptual problem unscathed. It is still possible that there are fundamental and yet ontologically dependent entities. Shall we therefore reject ontological fundamentality? The answer, I will contend, is “no”.

I will suggest that the case of quarks is problematic insofar one neglects the possibility of equifundamental entities. I acknowledge that the notion of equifundamentality is somewhat obscure. Yet the proposed account of ontological fundamentality is equipped with a serviceable definition. By appealing to it, I will show that quarks turn out to be equifundamental entities. Of course, the strategy generalises. We can treat cases of entities that are dependent on each other as equifundamental. This requires us to abandon the idea that such entities are fundamental in an absolute sense. But this is cost that is worth paying: in fact, it allows us to preserve an account of ontological fundamentality. Here I acknowledge that it is desirable to preserve the intuition that some equifundamental entities are more fundamental than other ones. So I will conclude by sketching a strategy to preserve such an intuition. Overall, an account of ontological fundamental remains an attractive option.

Mental Action and Control in the Case of Remembering

Seth Goldwasser (University of Pittsburgh)

14.09, 16:45 – 17:20, Room: 2.13

Remembering is an ubiquitous activity. Much of adult life consists in activities and actions which necessitate recalling the right information at the right times for a given purpose or project. In our everyday activities we must remember any number of names, dates, words, digits, techniques etc. to continue a conversation, write, mail, purchase, call, efficaciously tackle tasks requiring skill, and so on. Moreover, the mnemonic strategies implemented in and the phenomenology of remembering varies with types of information recalled and across contexts. Many times we can recite a series of digits immediately and with near unshakable certainty. Other times, despite some initial difficulty, we may provide, say, the name of someone that a friend is searching for during a moment of what Felipe De Brigard calls ‘joint-reminiscing’ (2018). Still other times we are at a complete loss for, say, a date or are totally unaware of some fact we should be remembering until we are given salient

contextual information. Finally, sometimes we just fail to remember. We try every mnemonic trick in the book and appeal to others, but to no avail – we just can't remember. All of these situations of remembering (or of failing to remember) and their phenomenology are familiar.

Cognitive psychologists and philosophers of mind have made strides in seeking to understand the nature of memory. The greatest of these, at least in the case of episodic memory, is the reconceptualization of the cognitive processes underlying memory from passive storage of information and retrieval later to active encoding of experienced content and imaginative reconstruction later. Thinking of memory as paradigmatically active, imaginative reconstruction of content has invited further theorizing about the nature of memory which places agency at its core. Philosophers and psychologists have not hesitated to answer that call. Several theories of memory have recently cropped up that all claim to explain how it is that remembering is an activity or mental action but do so in radically different ways. One major hurdle each account faces is finding a place for the subject to control the process of recollection. Given an assumed automaticity-control dichotomy in the cognitive psychology literature (Bargh & Chartrand, 1999; Koch & Crick, 2001; Schneider & Shiffrin, 1977; Bargh, 1994; Wegner & Bargh, 1998; Palmeri, 2002; Schneider, 2001) and since Galen Strawson's pivotal (2003) paper, onus has been placed on proponents to show how a candidate cognitive process is controlled by the subject in order for it to count as a mental action. Thus, philosophers who argue that we should understand remembering as an activity or mental action must provide some foothold for the subject to guide and control the process of recollection. Yet, there is serious disagreement about how exactly a subject controls her memory. What is more, proponents of each account have empirical support from cognitive psychology and neurology which they leverage in favor of their pet theory.

In this paper, I attempt to consolidate and organize the variety of accounts of remembering as a mental action. In section 2, I survey and classify accounts according to how subjects exert control over the process of recollection. There are accounts which claim that control is exerted by system-1 metacognition, accounts which claim that control is exerted by system 2 metacognition, and accounts which claim that control is exerted externally and supra-personally. I argue that each class describes some means by which a subject controls the process of recollection but each nonetheless fails to capture remembering as a mental action. In section 3, I, in agreement with Wayne Wu, reject the automaticity-control dichotomy as Strawson presents it. I then draw on recent findings from neurodevelopmental and cognitive psychology to propose that control exerted by subjects over the process of recollection is primarily the kind of control we exert over habitual actions. I thus claim that the external/supra-personal accounts describe how it is that the cognitive process of recollection is first controlled by the subject – by habituation in the context of joint activity. Nonetheless, as stated, all three types of account adequately describe other ways that the subject plays a role in the process of recollection. In section 4, I conclude that the ecumenical developmental account of remembering as a mental action that emerges from §§2 and 3 has all the strengths of each type of account while avoiding their limited scopes.

Are Delusions Beliefs or Experiences?

Jorge Gonçalves (IFILNOVA)

15.09, 9:35 – 10:10, Room: 2.13

The aim of this article is to determine what kind of mental state are delusions whether they are beliefs or experiences. In psychiatry delusions are defined as beliefs: ""A false belief based on incorrect inference about external reality that is firmly held despite what almost everyone else believes and despite what constitutes incontrovertible and obvious proof or evidence to the contrary" (DSM5). However, several researchers (Gallagher, Sass and others) argue that delusions are not beliefs. The arguments are of two types. In the first place there are those who maintain that beliefs change when the evidence of the facts so imposes, and this does not happen with delusions. They resist any argument or factual evidence. The second type of argument is that a belief to be truly a belief must have consequences in action, must guide action, and at least some delusions do not seem to have coherent consequences in action. For example, a patient claims that someone is poisoning his food and yet he does not hesitate to eat it (without showing any signs of intention to commit suicide). These situations are known as *double bookkeeping*. These two kinds of arguments would prove that delusions are a kind of mental state different from beliefs. So, these researchers think that delusions are a different kind of mental state and phenomenologists think they are experiences. For example, Gallagher (2009) consider that delusions are an experience of a reality alternative to everyday consciousness. If delusions are experiences the referred paradoxes about delusions would be understandable. For example, Sass (1994) affirm that sometimes we can see some irony in schizophrenics as they known that their delusions are not real. A counterargument that could be given (Bortolotti, 2012) is that the referred properties of delusions (incoherence and resistance to evidence) are also present in normal mental states that one call usually "beliefs". So, not all beliefs are rational, some beliefs we have in everyday life are irrational. If this is true there is a continuity between delusions and normality. Which of these two views is right? I think they are not necessarily incompatible. I think that beliefs may exist within delusional reality. They will not be beliefs about the world of everyday reality, but beliefs within delusions themselves (although the world of everyday reality may also enter the delusion). A delirious patient who believes herself to be a Messiah has this belief in an alternative way different from everyday reality. However, since this alternative reality can be the dominant universe of the psychological life of the subject, from the external point of view of society, it is as if the patient has beliefs about everyday reality. What is lived inwardly as an intense experience may be regarded by the external observer as a very irrational belief. Nonetheless, I think that the element that can tilt the balance in favour of belief is that some delusions are motivated. They have the function of protect psychologically a subject of greater troubles. For example, a pianist who had an accident and whose girlfriend had leave him, made a delusion where he was going to married her. The delusion protected him of falling completely in depression. So, on my view at least these kinds of delusions could be experiences, but the component of belief is also very important, because it is needed to have a motivated function. It is a case of research whether one can generalize for all delusions that they have some kind of function, and in this case, it seems to me that the component "belief" is very important.

Predictive Processing and the Content of Consciousness: a Fundamental Limitation?

Steven S. Gouveia (Universidade do Minho)
Joint work with Georg Northoff (Institute of Mental Health Research, University of Ottawa)
13.09, 15:40 – 16:15, Room: 2.13

Predictive processing is a fresh and new framework in cognitive and computational neuroscience that have been influenced by several disciplines, including artificial intelligence, philosophy, psychology, and so on. One of its main ideas is to see to the brain as a prediction machine: its goal is to anticipate the incoming sensory data (that is predicted) with the actual sensory data (real). The PP framework has been applied to several distinct functions of the brain including action, perception, attention, cognition, etc. Most recently, PP has also been suggested to serve as framework for consciousness.

The main focus in this paper is on whether PP can properly explain consciousness. Consciousness can be characterized by content, level/state, and form. Based on various lines of empirical data, we argue that PP can well account for the content of consciousness. In contrast, PP remains insufficient when it comes to the level/state and especially the form of consciousness including the subjective experience of the contents of consciousness as characterized by various phenomenal features. Hence, we conclude that PP remains limited in explaining the association of content with consciousness.

Therefore, PP needs to be complemented by a wider and different framework which, as based on the recent temporo-spatial theory of consciousness (TTC), may be spatiotemporal.

An Alternative to the Orthodox View of Non-Existence and Meinongian Approach

Zhiwei Gu (Central European University)
14.09, 9:00 – 9:35, Room: 5.2

I try to answer the question why some statements about fictional characters are true. Peter Van Inwagen (1977) makes the distinction between a statement within a fiction and a statement about fictional characters but out of a fiction. This distinction is helpful in tackling the question but it seems incomplete. The following three statements illustrate the distinction and what I meant by “incomplete”:

(1) Small, podgy, and at best middle-aged, he [Mr. George Smiley] was by appearance one of London’s meek who do not inherit the earth. His legs were short, his gait anything but agile, his dress costly, ill-fitting, and extremely wet. (Tinker, Tailor, Soldier, Spy, Chapter 2)
(2) Mr. George Smiley is more famous than Sergei Skripal.
(3) Mr. George Smiley is a small, podgy and middle-aged spy.

(1) is typically a statement within a fiction, since I quoted it from Chapter 2 of John Le Carre’s spy novel “Tinker, Tailor, Soldier, Spy”. While (2) is not, since Sergei Skripal is a real spy--a double agent who was recently almost killed in England. (3) is not within a fiction because there is no such sentence in John Le Carre’s novels, whereas (3) can be inferred from (1). I think a statement like (3) belongs to a third category beyond Van Inwagen’s distinction. I will mainly focus on statements like (3), whose truth-value is most controversial. Many philosophers (e.g. Tim Crane and Peter Van

Inwagen) think that (3) is false. I argue that both (2) and (3) are true, but they are true because of different reasons. I also argue that statements such as (1) settle down the rule of using proper names of characters. They are not asserting truth or falsity.

I think that Tim Crane's (2013, p. 124) reductive explanation is on the right track in dealing with the problem of non-existence. The basic idea of reductive explanation is to explain truths about non-existing things in terms of truths about existing things. I agree that reductive explanation can deal with truths like (2). But I also think that the methodology of reductive explanation can be applied to explaining why (3) is true but with a minor revision: some truths about non-existing things are true because they can be reductively explained by some rules.

According to Van Inwagen, (1) is not about Mr. George Smiley and it is not used by Le Carre as a vehicles of assertion. He invents the jargon "ascription" to depict the relation between Mr. George Smiley and smallness, podginess and so on. I agree with Van Inwagen that (3) is not used to asserting truth or falsity. However, (1) is surely *about* Mr. George Smiley, the creature of a series of Le Carre's spy novels. Then what is the role of (1)? Why (1) is not the vehicle of truth-value? When John Le Carre was writing (1), he was creating a character. There are nothing about the character to be asserted before an author wrote down something descriptive. But there will be something to assert after an author's creation. We will know who Mr. George Smiley is once he was created. We know Mr. George Smiley is not tall or handsome like James Bond. But what is to create a character? It surely is not to create a real person or something. To create a character is to settle down how to use the character's name thereafter and is to put constraints on a character. So those statements within fictions are like rules in an instruction. They are more like orders than descriptions. As Wittgenstein talked about the standard meter in Paris, "But this [the standard meter is 1 meter long] is, of course, not to ascribe any remarkable property to it, but only to mark its peculiar role in the game of measuring with a meter-rule"(§50, 29e, PI). (1) is also not ascribing any remarkable property to a character, but only to settle down the role of the proper name "Mr. George Smiley" in relevant contexts. Therefore, we had better say that a statement like (1) is not asserting truth or falsity.

If (1) is conceived as a rule of using the proper name 'Mr. George Smiley' or a constraint on depiction, then it seems sensible to claim that (3) is true. The reason is simple, because (3) is derived from (1). (1) explains why (3) is true. One might contend that if (1) itself is not true how it could explain the truth of (3). Crane's reductive explanation requires the explanans to be a truth. Consider an instruction of using a Microwave. There may be an instruction like this: don't put any metallic material into the microwave. This is only an order. Now if you follow this order and do not heat food with a metallic utensil then what you do is *correct*. Similarly, (3) is true because it is accordance with (1).

If (3) is true, does it mean Mr. George Smiley as a character has the properties of smallness, podginess and so on? Both Van Inwagen and Crane reject this. Van Inwagen thinks (1) is not really predication. It is an ascription instead. Crane thinks that smallness, podginess are not really representation-dependent properties and only representation-dependent properties can be predicated to a character, such as being created by John Le Carre, being a character, being recently portrayed by Gary Oldman, etc. I like Crane's distinction between representation-dependent properties and non-representation-dependent properties. But whether a property is representation-dependent is indexed to contexts. Once a character rather than a real individual is involved, the properties talked automatically become representation-dependent. So (3) involves representation-dependent properties. (3) is true because it can be reductively explained by (1). This reductive explanation is not essentially different from an explanation of a true statement such as (4) Mr. George Smiley is a character in Crane's work.

Why Fit-Based Buck-Passing Fails

Tanner Burke Hammond (Boston University)

13.09, 15:05 – 15:40, Room: Mattos Romão

Some recent attempts to rescue the buck-passing analysis of value[12] from the so-called "wrong kind of reason problem"[13] have tried to marry it to the notion of a *fitting* attitude, specifically by appeal to the Brentanian notion of *emotional correctness*.[14] On these accounts, "goodmaking" right-kind reasons for an attitude towards an object are to be understood as facts about an object that make a pro-attitude towards it *correct* or *fitting*.[15] In response to the "evil demon" articulation of the WKR problem, the 4 fit-based buck passer will argue that while it may be true that we have *some* reason to have a pro-attitude towards an evil demon who threatens to harm us unless we do so, it also seems clear that this attitude is not *fitting* with respect to its object, and thus we can avoid the problematic result that the demon is "good." In this paper, I argue that the putative appeal of the fit-based BPA rests upon the equivocal use of an unclarified notion of emotional correctness. In order to properly assess the prospects of emotional correctness for the BPA, I distinguish and clarify three different and often conflated ways in which this notion can be understood:

C1. *Adequation*: Emotional attitudes are "correct" in virtue of an agreement between some essential feature of a given emotion and some corresponding feature of the object of that emotion.
C2. *Deontic Satisfaction*: Emotional attitudes are "correct" just insofar as they are required or permitted by some broader normative criteria that set deontic constraints on what ought to be in general – e.g. Maximization, Universalizability, etc.
C3. *Unanalyzability*: "Correctness" is a primitive, unanalyzable feature of emotional attitudes. On this account, the fact that an object is correctly loved, and thus good, means nothing more than the fact that loving it is correct, sans phrase. Correctness is thereby a normatively self-imposing feature of emotional attitudes.

---

[12] According to the traditional "reductive" buck-passing analysis of value (henceforth BPA), x's being "good" (or valuable) is not a substantive property, but the merely formal property of having other lower-order (non-evaluative) properties that give us reasons to favor x. The locus classicus of buck-passing analyses can be found in Scanlon's What We Owe Each Other p. 97: "A being good, or valuable, is not a property that itself provides a reason to respond to a thing in certain ways. Rather, to be good or valuable is to have other properties that constitute such reasons." Crisp (2008) argues that we can actually distinguish between two distinct buck-passing theses in Scanlon's account that are in principle separable: BPA(Negative) : Being good is not itself a reason-providing property; BPA (Positive) : Being good is merely the higher-order property of having lower-order (non-evaluative) properties that provide reasons to respond in particular ways. As Crisp and others have pointed out, only BPA (Positive) entails a reductive analysis of value. While Scanlon (2002) has since come to distance himself from the positive buck-passing thesis and the reductive analysis of value it entails, such reductive accounts continue to enjoy wide support, and are often taken to have the explanatory edge over value-based axiologies due to Mackie-style concerns over the putative ontological and epistemological queerness – and general explanatory profligacy – of substantive value properties.

[13] A classic problem facing the BPA – namely the Wrong Kind of Reason Problem (WKR) – is that the presence of reasons to favor x appears insufficient to determine the goodness (or other evaluative status) of x, as there are cases in which we have reasons for a pro-attitude toward an object despite the fact that the object is clearly not good – in which case, the reasons to favor x are of the "wrong kind" to ground the goodness of x. The most well-known articulation of this problem appeals to a scenario in which an evil demon will inflict severe pain on me (or impose some other disastrous consequence) unless I have a pro-attitude towards a saucer of mud. Insofar as my having the attitude will shield me from pain, it appears I have a reason to favor the saucer of mud, and thus we are lead the unpalatable conclusion that the saucer of mud is good. Objections of this kind are varied and can be applied to a range of different evaluative concepts, but they all turn upon the basic idea that a reductive buck-passing analysis suffers from a problem of extensional implausibility, as it leads to the result that the extension of some evaluative concept includes things that contradict our basic platitudes about what things have that value.

[14] See Brentano 1952, p. 144: "If we call [nennen] an object good... we do not thereby want to add a further determination [Bestimmung] to the determinations of the thing in question... If we call certain objects good, and others bad, we say no more than that whoever loves this, hates that, is correct to do so [verhalte sich richtig]."

[15] See Olson and Danielsson (2007)

I argue that C1 and C2 will not help the BPA escape the WKR problem, and furthermore that the commitments needed to defeat these challenges will pressure the fit-based buck passer to appeal to C3. However, I argue that C3 turns upon theoretical posits that incur the same charges of explanatory profligacy leveled against value-based axiologies.

Why C1 Fails to Save BPA: The most intuitive way to account for emotional correctness is to posit a feature of the object that somehow corresponds to the essential character of the attitude in question: for instance, an attitude of admiration would be correct only insofar as we find some essentially admirable correlate property on the object-side of the emotion. However, I argue that all such accounts of correctness either (a) analyze values by appeal to further lower-order evaluative terms, and so are not available options for a buck-passing analysis of value, or else (b) in eliminating evaluative properties, also thereby forfeit criteria for excluding those reason-giving properties that would generate WKR problems. Take for instance the demon's property of being such that admiring him will avert disaster – a property which gives us reason to admire him, but not the right kind to ground the demon's goodness. We need some criteria for excluding such properties from candidacy in order to avoid the WKR problem. However, I argue that without object-side evaluative properties to serve as normative constraints on correct emotional attitudes, there appears to be no reason in principle why those properties underlying wrong-kind reasons could not serve as fitting object-side correlates for a given pro-attitude.

Why C2 Fails to Save BPA: The problem with analyzing a correct emotion as one that satisfies some general deontic consideration is that it makes emotional correctness an essentially instrumental and relational property, leading to cases where we ought to have pro-attitudes towards objects that we would not consider good. For example, suppose that the evil demon threatens to destroy the planet unless we admire him. If we assume further that there is some general normative schema according to which such a large-scale catastrophe ought to be avoided whenever doing so is, say, optimific, then admiring the demon would be "correct" insofar as it satisfies the general normative requirement of realizing optimal outcomes. The evil demon would thereby constitute the object of an attitude that "ought to be." Applying a buck passing analysis of values, we are lead once again to the problematic result that the evil demon is good.[16]

Why C3 Fails to Save BPA: If the buck passing analysis of value is motivated in part by concerns over the explanatory profligacy and queerness of value terms, then appealing to an unanalyzable, primitive feature of mental states will incur the same problems. In Mackie-esque terms, making correctness an unanalyzable feature of attitudes would entail that correct emotion somehow has "to-be-ness" built into the attitude itself. In endorsing C3, we thereby avoid adding a substantive property to objects we call good only at the expense of adding one elsewhere – specifically, to the emotional attitude itself now considered as a psychological entity. Thus while "goodness" would remain a merely formal or "syncategorematic" term, "fittingness" must now be given its own ontological purchase. The question, then, of how goodness is ontologically grounded remains just as mysterious on a fit-based analysis of value. Furthermore, and as a result of these ontological concerns, the fit-based analysis also thereby incurs the attendant charge of epistemic queerness often levelled against value-based axiologies. Unless we are to forfeit our claim to moral knowledge, emotional correctness is a feature we must be aware of in order for the object of the emotion to be judged good. Upon C3, however, the correctness of an attitude is not given by the awareness of

[16] Of course, we could try to avoid this kind of problem under C2 by stipulating that correct attitudes must also bear an appropriate relationship with their objects, but I argue that all options for doing so will ultimately lead us back to the challenges faced in C1.

some correlate property, nor by the awareness that some external normative criteria are satisfied. Rather, correctness is supposed to be an unanalyzable, irreducibly normative feature, which raises the question of how exactly we are supposed to have epistemic access to it. On both ontological and epistemic grounds, then, appealing to C3 forfeits any presumed advantage of explanatory parsimony over value-based axiologies.

Coherence of Character and the Undefeatability of Safety

Jaakko Hirvelä (University of Helsinki)

13.09, 17:20 – 17:55, Room: Sala de Actos

Knowledge defeat seems to be a genuine phenomenon. Sometimes a subject knows that *p*, gains overwhelming misleading evidence that her belief doesn't amount to knowledge and thereby loses her knowledge, even if she continues to hold onto her belief. Or so defeatists claim. The defeatists hold that higher-order evidence can destroy one's knowledge without destroying the underlying belief. If knowledge defeat is a genuine phenomenon, then theories of knowledge that cannot accommodate it are in trouble. Safety-based theories of knowledge belong to this unfortunate lot.

According to the safety condition a subject S true belief that *p* amounts to knowledge only if in all nearby possible worlds where S believes that *p* via the method of belief-formation that she uses in the actual world, *p* is true. While relatively few epistemologists think that the safety condition gives the necessary and sufficient conditions for knowledge, here we will examine the prospects of the idea that knowledge is, essentially, a safe belief but leave it as an open question whether knowledge can be analyzed in terms of safety.

The safety condition cannot be defeated by misleading evidence. If S's belief that *p* is safe at t^1 and S acquires misleading evidence that her belief that *p* isn't safe at t^2, and she ignores the misleading evidence holding dogmatically onto her belief, her belief that *p* can continue to be safe as long as she doesn't re-base her belief. Lasonen-Aarnio (2010) has argued that the strategies that the safety theorist could pursue in explaining knowledge defeat don't work. If Lasonen-Aarnio is right, then the safety theorists must hold that knowledge defeat isn't a genuine phenomenon.

That said, safety theorists have at least two promising explanations of knowledge defeat at their disposal. Lasonen-Aarnio (2010) has argued that intuitions of knowledge defeat track the manifestation of knowledge-conducive dispositions and that one's belief can be safe even if it's gained via bad epistemic dispositions. Baker-Hytch and Benton (2015, p. 56) have argued that externalists should accept the knowledge norm of belief, according to which one must not believe that *p* if one doesn't know that *p*, which motivates a guidance norm, that states that one must refrain from believing *p* if one comes to believe that one's belief that *p* isn't knowledge. According to Baker-Hytch and Benton our defeatist intuitions track the violations of the guidance norm.

While both of these strategies of explaining our intuitions of knowledge defeat entail that knowledge can be preserved in the face of misleading evidence, there is logical space for a new kind of explanation of our defeatist intuitions that denies the significance of misleading higher-order evidence. This new solution is similar to the earlier solutions in holding that misleading evidence doesn't necessarily have the power to destroy knowledge. It differs from them in claiming that in paradigmatic cases of knowledge defeat where the subject holds onto her belief knowledge was never acquired by the subject to begin with. I will argue that such a solution presents itself once we relativize the safety condition to virtuous methods of belief-formation.

The idea that knowledge is always gained through the exercise of our cognitive abilities is quite plausible and widely held among contemporary epistemologists.[17] Virtue epistemologists are the foremost proponents of this idea, since they think that our epistemic competences or virtues must be manifested in our cognitive success in order for that success to count as knowledge. For our purposes it's enough that knowledge is always gained through epistemic virtues. Assuming that knowledge is always gained through virtue, the safety theorist should be allowed to relativize the safety condition to virtuous methods of belief-formation. After all, if we have independent reasons for thinking that all beliefs that amount to knowledge are virtuously formed, then by relativizing the safety condition to virtuous methods of belief-formation we don't risk excluding any beliefs that amount to knowledge from the extension of the safety condition.

Following Sosa (1991, p. 284) epistemic virtues can be conceived as dispositions of the agent to believe correctly. Epistemic virtues differ from reliable dispositions is that an epistemic virtue must be integrated to the subject's cognitive character (Pritchard, 2012, p. 262). The stable belief-forming dispositions that are integrated to our cognitive character act in concert rather than be in tension with each other. The subject's belief-forming dispositions can act in concert without the agent being aware of this. No reflective endorsement of the reliability of the dispositions is required. The minimal requirement of cognitive integration is that the dispositions silently confirm the outputs of each other as our perceptual faculties normally do.

The virtue relativized safety condition isn't satisfied if the subject believes that p and doesn't suspend her judgement upon encountering and accepting the "misleading" evidence. In such a case the subject's original belief cannot have been formed via virtuous method of belief-formation, since that belief-forming disposition is in clear tension with the other belief-forming dispositions of the subject. Therefore, the subject's original belief that p was not virtuously formed, and she cannot satisfy the virtue-relativized safety condition. If, on the other hand, the subject would suspend her judgement when she encounters the misleading evidence it's entirely possible that her original method of belief-formation was a virtuous one, since that disposition is acting in concert with the subject's other belief-forming dispositions that make up her cognitive character. In such cases the subject's belief might have qualified as knowledge prior to encountering the misleading higher-order evidence, but doesn't count as knowledge afterwards, since she has suspended her judgement.

While both kinds of cases might seem to involve knowledge defeat, neither actually does. In the former cases knowledge was never had and in the latter cases knowledge is lost due to loss of belief. Knowledge defeat is, then, an illusory phenomenon.

[17] This idea is most prominently challenged by Lasonen-Aarnio (2010).

There are no Different Kinds of Imagining

Madeleine Hyde (Stockholm University)

14.09, 17:20 – 17:55, Room: 2.13

Imagining is an everyday propositional attitude – i.e. a mental state with a propositional content – that we can employ for various purposes. I can imagine my dream holiday, what a centaur might look like, imagine a new kind of chair or use my imagination to figure out whether it is possible to fit a sofa through my front door. What we have here looks like at least four different uses of our imagination, respectively: (i) picturing how things would be if the world better fitted my desires, (ii) entering into a fiction, (iii) creatively coming up with something novel and (iv) working out what is possible.

Some authors have treated cases like the above as different *kinds of* imagining (see e.g. (Currie & Ravenscroft, 2002) (Dorsch, 2016) (Williamson, 2016), and (Kind, 2016) for an overview). I argue that such a view is mistaken: that all of these candidate types of imaginative episode are structurally similar enough that we need not say that there is *this* or *that* kind of imagining, but merely observe that these are different uses of a single propositional attitude. What tells them apart, as will become apparent, only concerns the individual aim of the imaginative exercise.

My aim will be to scrutinize these four cases of imagining piecemeal, in order to see why they have been variously treated as unique kinds. In doing so, it will also come to light what these cases have in common. In having propositional content, our imaginative episode can build up a picture of a possible world via the object(s) and properties that the content gives us. In each case, our imaginative episode draws its mental imagery from past experience and 'recycles' it to come up with something new (i.e. currently non-actual).

This can otherwise be described as a shift in perspective, from the actual to the non-actual: which can focus on just an object, as when we creatively try to imagine something new, or in going from a whole actual scenario or world to a particular possible one: this happens when we imaginatively engage with a work of fiction, and also when we try to imagine what is possible, given our actual circumstances e.g. whether I can climb the tree in front of me. The difference between the last two examples is with respect to how close the imagined possible world is to the relevant actual world (i.e. the one that the imagining subject is situated within). Again, this difference depends upon the aim of the imaginative episode. Thus we find, by trying to pick these uses of the imagination apart, that all that really separates them, at bottom, is the aim behind them.

Uncertainty in Science: a Study on the Role of Non-Cognitive Values in the Assessment of Inductive Risk

Silvia Ivani (Tilburg University)
Joint work with Matteo Colombo (Tilburg University) and Leandra Bucher (University of Wuppertal)

14.09, 14:30 – 15:05, Room: Mattos Romão

Women are frequently excluded from clinical trials because of high costs (women are expensive to test because of hormone fluctuations) and medical concern (clinical trials are potentially dangerous for women's fertility). Many scientists question the reliability of these clinical trials (see Simon 2005): because of sex differences in drug reactions, both women and men must be tested. This prevents health risks for both sexes. Clinical trials involving both sexes are now often conducted.

Still, uncertainty remains about the right methodological choice to take.

Scientific research involves uncertainty. Scientists have to take decisions about methodologies and hypotheses and each one of these decisions involves uncertainty. Lack of sufficient evidence and disagreements about methodologies – as the example above shows – are sources of uncertainty that can introduce error in scientific reasoning. One kind of error is associated with the notion of inductive risk, i.e., the chance of taking wrong decisions, such as accepting a hypothesis that is in fact false. Philosophers argue that inductive risk challenges the ideal of value-free science, i.e., the idea that non-cognitive values (e.g. moral and economic values) do not influence research, and it shows their actual beneficial role in science (Hempel 1965; Douglas 2000). Specifically, considering non-cognitive values is beneficial when taking wrong decisions may involve non-cognitive consequences, such as harming women's health.

Our study aims at investigating the relation between non-cognitive values and inductive risk. We present the results of an experimental study clarifying the psychological impact of political values and personal features like one's race and sex on the acceptance (or rejection) of scientific hypotheses in the face of inductive risk. Our hypotheses was that political and personal identity features reliably predict people's sensitivity to scientific errors. Specifically, people are less likely to accept hypotheses that they perceive as clashing with their political ideology and identity.

In our study, participants were asked to read and evaluate three vignettes, where scientists disagree about the adequacy of a specific test and take decisions about hypotheses involving sexual or racial differences. In each vignette, the consequences of a mistaken decision could harm a group of people (either women, men, black or white people). One of the vignettes concerned the exclusion of women from clinical trials. In this vignette, scientists decided to introduce a new drug tested on a group including only men into the market. Participants were asked to express how certain they were that the decision taken was a good one. Our hypothesis was that conservative men were more likely than women to see that decision as a good decision. At the end of the survey, information about political ideology, race, and sex was collected.

Our results provide us with a more nuanced understanding of the bearing of non-cognitive values on the psychology of inductive risk. Though philosophers of science have drawn on several historical case-studies to clarify the notion of inductive risk, little attention has been paid to how people actually reason about inductive risks. In this paper, we set out to begin filling this gap in the philosophical literature by investigating the relationship between reasoning, inductive risk, and non-cognitive values.

Climate Change and Harm to Collectives

Tvrtko Jolić (Institute of Philosophy Zagreb)

14.09, 9:00 – 9:35, Room: Mattos Romão

With the rising concern that the climate change is going to turn the life prospects of the future generations for the worse, philosophers try to answer who exactly are we harming by our environmentally reckless behaviour. It is obvious that the most dramatic consequences of the climate change will be felt by the generations that will not come into existence during our lifetime. As Parfit noted, our choice between different policies affect who will later be born (Parfit 1987). We can for instance choose a risky public policy that leads to 3C temperature rise 200 years from now: that will create harsh life conditions for the future humans, but on the other hand, these people would

never have been born had we chosen some other policy. This is the famous non-identity problem. This reasoning led some to the conclusion that these future people are not harmed by our reckless behaviour that created harsh life conditions as long as they are provided with basic requirements for leading minimally decent life. Many find this conclusion counterintuitive: they believe that some act can be wrong even though they don't harm future persons they cause to both exist and suffer. But who is then wronged/harmed? One way to deal with this problem, suggested by Page (2006), is to claim that the harm is done not to the individuals but to the collectives. Even though no particular individual will exist in the future where risky policy is adopted who would have existed had it not been, various collectives will.

In my presentation I take up on this suggestion, and I will try to provide explanation how collectives can be harmed. The problem I want to address is how can a collective be harmed by an action when no single member of that collective is harmed by it (in above mentioned sense). The idea of collective rights, where collective rights are understood to be simply aggregation of the rights of its members, cannot answer the question. This is because the change of membership in the generations causes changes in the aggregated collective rights. The non-identity problem kicks in again! Instead I suggest we should adopt a more robust idea of collectives, the one based on the idea of group agency. As noted by List and Pettit, group agents "are distinguished by the fact that they can enter a system of obligations recognized in common with others, and limit their influence on one another to that permitted within the terms of that system" (List and Pettit 2011: 178). States are prime example of the collectives that are also the group agents, and I will claim that they can be harmed by our environmentally reckless behaviour. I take it that states have a duty to secure basic human rights of their own citizens and of non-citizens. Hence, states have corresponding right not to be subverted in performing this duty. Other people (and other states) violate this collective right by adopting environmentally reckless behaviour that leads to the condition where state cannot perform its basic function.

Towards Constitutive Normativity of Meaning

Bartosz Kaluziński (Adam Mickiewicz University in Poznań)

13.09, 15:40 – 16:15, Room: 5.2

The thesis that "meaning is a normative notion" became very influential in contemporary analytic philosophy due to the famous book by Saul A. Kripke titled Wittgenstein on rules and private language (1982). Although the slogan "meaning is normative" is very catchy and many philosophers support it (cf. Boghossian 1989; Buleandra 2008; Brandom 1994; Ebbs 1997; McDowell 1984; Millar 2004; Peregrin 2012; Whiting 2007, 2016), it is not entirely clear how that thesis should be understood.

Usually, normativists about meaning assume that the meaning of a given linguistic expression produces pragmatic obligations as how we ought to use that expression. To put things in a quasi-formal way, normativists about meaning tend to assume that following semantic norm:

$$\text{(SN) } t \text{ means } F \rightarrow \forall x \text{ (} t \text{ applies correctly to } x \rightarrow x \text{ is } f\text{)}$$

produces pragmatic obligation:

$$\text{(PN) } S \text{ means } F \text{ by } t \rightarrow \forall x \text{ (} S \text{ ought to apply } t \text{ to } x \rightarrow x \text{ is } f\text{)}$$

where t is a term, F is the phrase stating its meaning, x is a variable and f is a feature by virtue of which t relates to x and S is a speaker.

Abovementioned interpretation of the normativity of meaning thesis brought serious criticism (cf. Bilgrami 1993; Glüer 1999; Hattiangadi 2007; Wikforss 2001; Boghossian 2005):

A. It has very counterintuitive consequences when it comes to lying or using irony (we always ought to speak the truth and when we apply the term "horse" to cow in order to make an ironic statement we to not ascribe to that term its standard meaning).
B. It is doubtful whether (SN) can produce (PN) "just like that". Usually, prescriptivists claim that the notion of correctness in (SN) must be normative, hence it produces (PN). The anti-normativists reply that "correct" can be used in both normative and non-normative way, and they ask for an argument supporting the claim that in (SN) it is indeed used in a normative way.

In my presentation I intend to outline an alternative reading of the normativity of meaning thesis along the lines of constitutive normativity that will prevent these difficulties from arising.

I will briefly recall the classical distinction between regulative and constitutive rules made by Searle (1969). But the orthodox formal account of constitutive rules, that seems to assume that they have specific form (X counts as Y in C) is problematic (cf. Randsell 1971; Hindriks 2009), so I will examine complicated team games that are considered as a paradigm examples of rule-constituted practices to make few adjustments. I will argue that (i) rules form a system and can have multiple forms (from definitions/specifications to prescriptions); (ii) rules are interconnected; (iii) rules play two roles simultaneously – they define some types of actions (offside), and govern the conduct within the practice (determine which moves are allowed and forbidden).

The revised account of constitutive rules (inspired by Ransdell 1971; Hindriks 2009) will been applied to the issue of the normativity of meaning. I will try to show that system of rules: (i) specifies the meaning of a term and determine which uses of that term are allowed/forbidden, (ii) should not be confused with "rules of efficiency" that are recommendations of as how we should speak to communicate fluently. This is to say, the system of rules specify the meaning of terms "bachelor," "widower" and "divorcee" and determine which "moves" are correct (allowed) – referring to bachelors by using the term "bachelor" – and which "moves" are incorrect (not allowed) – referring to bachelors by using term "widower." But that does not entail the conclusion that we always ought to speak the truth, irrespective of any other considerations (as it was in case of orthodox accounts of the normativity of meaning thesis that are mentioned above). Every practice has its ultimate goal (cf. Marmor 2007) (obtaining a favourable result/fluent communication) and sometimes "rules of efficiency" that tell us how to achieve that goal, override duties that constitutive rules put upon participants of the practice. So, it appears that norms constitutive for meaning can be conceived as imposing prima facie duties.

Summarising, I will argue that it is possible to conceive norms determining meaning as constitutive. And such an account can be seen as superior over orthodox accounts of the normativity of meaning thesis, because it is intuitive (many people have a strong intuition that language is somehow "made by us" and that language is somehow analogous to games) and it prevents issues A. and B. from arising.

A Mechanistic Conception of Metaphysical Grounding

Thomas Kivatinos (The City University of New York)

14.09, 17:20 – 17:55, Room: 5.2

Although the topic of grounding has received overwhelming attention in the past decade, it has been overlooked that the notion of grounding is extremely similar to another notion: the notion mechanistic dependence. As I demonstrate, these notions are so similar that they serve the same theoretical roles for elucidating the natural world and they pick out the same features of the natural world (rather than the domain of logic or abstract entities, which I leave aside). So, the notions can be plausibly treated as the same notion. I thus develop a "mechanistic conception of grounding" that identifies the notion of grounding with the notion of mechanistic dependence. This conception situates the notion of grounding in the mechanistic theoretical framework that aims to delineate the ontology and methodology of science. By doing so, this conception suggests that grounding belongs to the ontology of science and the notion of grounding is to be tailored to, and constrained by, the scientific practices of mechanistic modelling and mechanistic explanation.

The notions of grounding and mechanistic dependence are both notions of synchronic, non-causal dependence, meaning: if Y is grounded by X or Y mechanistically depends upon X, then Y synchronically and non-causally depends upon X. Mechanistic dependence, but not grounding, is understood in terms of the notion of mechanisms. A mechanism M is a collection of entities whose interactions and organization produce some phenomenon P, in which case P mechanistically depends upon M. For example, say that P is the formation of an agent's long-term memory. The mechanism that produces P, which P thus mechanistically depends upon, is the interactions and organization of the neurons in the agent's brain that are responsible for the agent forming a long-term memory.

As I argue, the notions of grounding and mechanistic dependence play the same theoretical role in the sense that each notion is employed for portraying phenomena in the same way: both notions are employed for portraying how "higher-level," less fundamental phenomena depend upon and arise out of "lower-level," more fundamental phenomena. For example, just as mechanistic dependence is employed for portraying how psychological processes depend upon and arise out of neural processes, the notion of grounding is used for doing so as well. For although the psychological is sometimes characterized as mechanistically depending upon the neural, the psychological is sometimes characterized as being grounded by the neural. Standard treatments of grounding imply that there are differences in what these characterizations capture about how higher-level phenomena arise out of lower-level phenomena. So, these treatments imply that there are differences between what a grounding-based account and a mechanistic-dependence-based account reveals about higher-to- lower-level connections between phenomena. But as I argue, those differences are superficial. That is, regardless of whether grounding or mechanistic dependence is used to theoretically portray how higher-level phenomena connect to lower-level phenomena, what is captured by such portrayals is the very same. In each case, what is captured is an asymmetry between some phenomena X and Y such that Y not only depends upon X, but Y is explained by X. And, crucially, regardless of whether Y is thought to be grounded by X or Y mechanistically dependent upon X, the form of explanation by which X explains Y is the same. Thus, the notions of grounding and mechanistic dependence capture not only the same way that natural phenomena are ontologically connected, but also the same way that they are explanatorily or epistemically connected. With this as my basis, I identify the notions and thereby construct a mechanistic conception of grounding.

Have We Lost Spacetime on the Way:
Narrowing the Gap Between General Relativity and Quantum Gravity

Baptiste Le Bihan (University of Geneva)

14.09, 17:55 – 18:30, Room: 5.2

In the recent literature much has been written on the emergence of space- time and spatio-temporal theories from non-spatio-temporal theories (cf. for instance Huggett and Wüthrich (2013), Crowther (2016)). In this talk, we discuss the extent to which this so-called emergence of spacetime is as radical as it has been claimed. We will argue in favor of two claims. First, contrary to what is often said, namely that spacetime emerges from a non-spatio- temporal structure, many of these cases might be described as the *emergence of non-spatio-temporal structures* from distinct non-spatio-temporal structures because spacetime, in a substantive sense, is already missing in General Relativity (GR hereafter). Second, we argue that in many programs in Quantum Gravity (QG from now on), a local distinction between space and time can still be implemented. Taken together, these two claims entail that spacetime may be regarded as already partially missing in GR, and as still partially present in QG, narrowing the explanatory gap between the two theories and questioning the relevance of describing the relation between GR and QG as entailing that GR-spacetime emerges from a non-spatio-temporal QG structure.

What do we mean exactly by "explanatory gap"? Following Le Bihan (2018), it is useful to introduce a distinction between *the hard problem of spacetime emergence* and *the easy problem of spacetime emergence*. The hard problem is to understand the ontological status of the two structures at play (the *fundamental* and the *derivative* structures) and the nature of the *relation* obtaining between the two structures: Are these two structures real, one of them being more fundamental than the other, or is the derivative structure a pure illusion, an approximate description of the fundamental structure without an ontological counterpart? The easy problem consists in finding the formal derivation of the derivative theory by using mathematical tools and bridge principles between the primitive notions involved in both theories.[18] We then define the "explanatory gap" between the two theories as the conceptual discrepancy between the primitive notions involved in the two theories. For instance, admitting that the easy problem was solved in the case of QG, one could ask how we should interpret the primitive notions of the derivative theory (GR), allegedly relying on the existence of spacetime, if these are very different from the primitive notions of the fundamental theory (QG), which are potentially non-spatio-temporal.

Note that in this talk we do not take stock on whether there *genuinely is* a hard problem of spacetime. One may indeed argue that there is no hard problem of spacetime because the derivative theory (GR) does not describe a genuinely existing structure, just as one may deny that there is a hard problem of consciousness by denying the existence of consciousness. However in this scenario one still has to explain *why* it is the case that there is no hard problem of spacetime (see e.g. Chalmers (2018) for a defense of this point in the context of the philosophy of mind). In a nutshell, if one argues that the physical system described by GR is not real in such a way that there is no hard problem of spacetime, the *ontological hard problem of spacetime* will be deflated. Nonetheless, the more general hard problem of spacetime will not be solved because we still have to understand why this approximation (the structure described by GR), rather than another one, describes the world. Compare this situation with the one we find in philosophy of mind: It is not because we solve the

[18] We keep the name "easy" in analogy with the terminology in the philosophy of mind, with a touch of humor. The easy problem of spacetime relies on finding a theory of quantum gravity and relating it to GR, which arguably is one of the most difficult challenge ever met by physicists.

hard problem of consciousness by stipulating that there are no mental states that the problem is fully solved. We still have to explain why we do have the illusion that there are mental states, and therefore, to explain why we are lured into believing that there is a hard problem of consciousness in the first place.

As we will go on to explain, we may narrow the conceptual discrepancy by endorsing the dynamical approach to GR *and/or* by acknowledging that the fundamental structure described by the various theories of QG already relies on a distinction between structures that maps to the spacetime split present in GR. Indeed, the two moves presented in order to narrow the explanatory gap stand independently. For instance, the reader who does not accept that the dynamical approach is consistent may accept the point about quantum gravity, and vice versa.

Here is a more precise description of the two moves. First, spacetime may be construed as already emerging in the framework of GR, entailing that what we (only) need to recover from QG is not the standard spatio-temporal in- terpretation of GR, but a non-standard non-spatio-temporal interpretation of GR. Indeed, as it is often pointed out in the literature, strictly speaking we do not need to recover GR from QG, but only something close enough to GR in order to explain its predictive success.[19] If this is true, recovering any particular interpretation of GR is enough to account for its predictive success. In particular, recovering a dynamical interpretation of GR in which "spacetime" is just a field systematically coupled to other fields – the so- called matter fields – is enough to explain its predictive success. Second, the property of (local) Lorentz symmetry shared by all relativistic fields corre- sponds to a crucial feature of spacetime, namely a split between space and time. And, as we shall see, this notion of a split between space and time still obtains in the fundamental structures described by QG approaches either in the form of (representations of) the Lorentz symmetry, or along another principle playing the same role in the theory.

Expressivist Account of Self-Deception

Jesús López-Campillo (University of Murcia)

15.09, 9:00 – 9:35, Room: 2.13

Self-deception is a species of motivated irrationality in which subjects manifests a characteristic conflict (Davidson 1985, Fernández 2013, Funkhouser 2005) between what they sincerely say (i.e.: "I believe I'm healthy") and how they act (i.e.: overstating their alleged health condition, avoiding going to the doctor, writing a will in spite of being too young, etc.). This conflict shows that self-deceived subjects suffer both from lack of knowledge and from lack of first-person authority. On the one hand, they suffer from lack of knowledge because the conflicting attitude that they have towards the object of their self-deception (i.e.: their health condition) seems incompatible with true warranted belief. On the other hand, they suffer from lack of first-person authority because self-deceived subjects are unconscious of being self-deceived (one could say: "I'm 100% sure that I'm healthy") so that it is easier to recognize cases of self-deception in others than in oneself (Lazar 1999).

There are three main accounts of self-deception in the literature: intentionalism (Davidson 1985, Demos 1960, Foss 1980, Pears 1984), *epistemicism* (Fernandez 2013, Funkhouser 2005, Holton 2001, Patten 2010, Sandford 1988) and *motivationalism* (Barnes 1997, Johnston 1988, Lauria et al. 2016, Lazar 1999, Mele 2001, Nelkin 2002, Szabados 1974, V.Leeuwen 2007). Since motivationalist accounts are dominant in the contemporary literature, in this paper I will focus on them.

[19] Cf. for instance Lam and Wüthrich (2017).

Motivationalism claims that self-deception is a biased belief-forming process unintentionally caused by a motivational state. Self-deceived subjects unintentionally weight the available evidence about p in a biased way (according to their own epistemic norms) because of the influence of their motivational states (i.e.: desires, anxiety, fear...).

Usually, intentionalists claim that their account is able to explain the characteristic conflict of self- deception (Mele 2010). On the one hand, subjects sincerely say "I'm healthy" because they have the biased belief that they are healthy. On the other hand, they act in a conflicting way because avoiding going to the doctor (where they could find unwelcome evidence) or overstating their alleged healthy condition (making thus more vivid the evidence supporting the belief that they are healthy) are typical ways of biasing the assessment of the evidence (Mele 2001). Therefore, self-deceived subjects suffer from lack of knowledge because they have unwarranted (biased) beliefs about p and from lack of first- person authority because their biases are unintentional so that they are not conscious of being biased.

I think motivationalism has not taken seriously the task of explaining the complexity of the self-deceived subject's behaviour hitherto. Motivationalism can explain the subject's sincere utterance and part of her actions (as we saw in the last paragraph). But it cannot explain, for instance, why the subject is also writing a will in spite of being too young. This piece of behaviour cannot be the result of her biased belief that she is healthy nor the result of a bias mechanism (since there's no way that writing a will biases the subject into thinking that she is healthy). Then, motivationalism would need to posit that the subject has contradictory beliefs (i.e.: the belief that she is healthy and ill) to explain all her actions (at least in some cases of self-deception). As a result, motivationalism would fall into the static paradox, one of the main criticisms that motivationalism itself makes to intentionalism (Mele 2001).

Therefore, I propose a *neo-expressivist account* of self-deception. This neo-expressivist account claims the difficulty to explain the behaviour of self-deceived subjects stems from the fact that self-deception is a *sui generis* mental state. Let's see the argument. (1) According to the neo-expressivist account that I propose, mental states are identical to *expressive patterns* of behaviour (Wittgenstein, 1953). An expressive pattern of behaviour is a series of *expressive episodes* (utterances, gestures, facial expressions, silences, pretenses, actions...) manifested by the subject over a certain period of time (minutes, hours, days, months..., depending on the kind of mental state). (2) Self-deceived subjects exhibit a particular expressive pattern according to which they cannot linguistically express *with first- person authority* the mental state in which they are. Indeed, when someone says "I believe it is raining", the belief itself is not normally questioned but, at most, the content of the belief (if it is raining or not). However, it is not unusual to interrogate someone who exhibits the conflicting behaviour of self-deception (i.e.: she sincerely says "I believe I'm healthy", but she avoids doctors and she is writing a will) about whether she really believes to be healthy and about the real reasons of some of her actions. (3) Conscious and unconscious mental states have different *kinds* of expressive patterns. Conscious mental states are those that can be linguistically expressed by the subject with first-person authority, while unconscious mental states are those that cannot be linguistically expressed with first-person authority (self-deception is an example of this).

The conclusion of the argument is that self-deception is a sui generis mental state because it has a *sui generis* expressive pattern, namely: the one constitutive of unconscious mental states. Self-deception and belief are two different psychological kinds because self-deceived subjects cannot linguistically express with first-person authority what they are self-deceived about, while believers can linguistically express with first-person authority what they believe. Self-deceptions are

unconscious mental states, while beliefs are conscious mental states. The virtue of the neo-expressivist account of self-deception that I propose here is that, unlike motivationalism, is able to explain the whole conflicting behaviour of self-deceptive subjects without falling into the static paradox. The actions of writing a will, avoiding doctors or saying "I believe I'm healthy" are not explained by positing contradictory beliefs, but by one single mental state: they are all expressive episodes of the *sui generis* mental state of self- deception. Therefore, on the one hand, self-deceived subjects suffer from lack of knowledge because belief is a necessary condition of knowledge and they have another kind of cognitive mental state (self- deception). On the other hand, self-deceived subjects suffer from lack of first-person authority because the expressive pattern of self-deception is the one constitutive of unconscious mental states.

Bifurcated Moral Status and Post-Humans

Giuliana Manca (King's College London)

15.09, 10:10 – 10:45, Room: Mattos Romão

The concept of moral status provides a foundation for human rights and the proper treatment of persons in social, political, and legal spheres. When theorists delineate moral status, they often appeal to moral considerability, or how one is or should be treated, considered, and respected; rarely do theorists consider moral responsibility, or how an individual is or should be responsible, accountable, and answerable. This paper argues that our concept of moral status is intuitively bifurcated, in that it contains a notion of moral considerability as well as a notion of moral responsibility. Examination of our practical conception of moral status, specifically how we treat children and young adults, shows that our concept of moral status includes both a threshold notion of moral considerability and a scalar notion of moral responsibility. Upon this conceptual foundation, it is suggested posthumans might have a higher moral status than human persons in the scalar notion of moral responsibility, but would have equal moral status in the threshold notion of moral considerability. With their elevated moral responsibility status, posthumans might be considered postpersons: individuals with a personhood status above that of contemporary persons. Possible future interactions between postpersons and persons are explored; the interactions between contemporary human adults and human teenagers – where the former enact a paternalistic relationship with the latter – is argued to be analogous to potential relations between postpersons and persons. As contemporary adults possess greater civil liberties than human adolescents, so too would postpersons hold greater legal, political, and social freedoms in comparison to persons. This paper concludes by reasoning that the autonomy, rights, and freedoms of future persons may be limited in comparison to those enjoyed by contemporary persons. Thus, human persons may be less free in a society in which postpersons and persons coexists.

Measurement-Theoretic Account of Propositional Attitudes and Intentional Realism

Maria Matuszkiewicz (University of Warsaw)

14.09, 17:55 – 18:30, Room: Mattos Romão

Number of philosophers (Churchland 1979, Davidson 1989, Dennett 1987, Stalnaker 1984, 1999, 2008) have expressed the thought that propositional attitudes predicates are analogous to measurement predicates. On this view, which was recently given a systematic formulation in Matthews (2007) propositional attitudes are not relations between thinkers and objects (e.g. propositions), but monadic properties of the agents. When we say that John weights 80 kilograms, the number 80 is a representation of the quantity of physical magnitude, and not something that John is related to. In the very same way a that clause that we use in belief report does not name the object that the thinker is related to and which is the object of her propositional attitude, but rather it names an abstract object, which is a representation of the thinker's dispositional property. The measurement-theoretic approach aims to provide a satisfactory account of salient features of propositional attitudes (semantic evaluability, causal efficacy, productivity, systematicity), as well as some features of our practice of attitude attribution (notably context-dependence).

In my paper I shall consider two versions of the measurement-theoretic approach to propositional attitudes: its systematic exposition due to Matthews (2007) and a more sketchy proposal by Stalnaker (1984, 1999, 2008), focusing on their relation to intentional realism and content realism. It may be argued (Matthews) that since contents are properties of abstract objects that we use to ascribe propositional attitudes, and not intrinsic properties of propositional attitudes qua dispositional states, the view is incompatible with intentional realism and content realism. However, I will argue that the theory is realist with regards to contents, as long as for there is a mapping which assigns to each propositional attitude qua dispositional state one object from the representing domain (or a subset of objects from the representing domain). I will consider two potential threats to realism: (i) context- dependence of belief attributions, (ii) situations in which a thinker is ignorant of an identity statement and thus there is no fact of the matter which proposition best represents her belief (Stalnaker 1999, 2008). I will argue that because the measurement-theoretic approach to content requires that there exists a structure-preserving mapping between two relational structures: the empirical structure of propositional attitudes and a representing structure of abstract objects that we use to characterize attitudes (propositions or some complex linguistic objects), the view is realist in the minimal sense stated above. I will also argue that for the very same reason the view satisfies a minimal internalist condition: that is although contents are specified in terms of objects external to the agent, it is a matter of her internal dispositional state, which classes of contents can be used to characterize her mental state.

Film as philosophy: a Defence of an Analytic Account

Tom McClelland (University of Warwick) and Diana Neiva (University of Porto)

14.09, 17:20 – 17:55, Room: Mattos Romão

The idea of "film as philosophy" (FAP) has its origins in the seventies and eighties with Stanley Cavell and Gilles Deleuze. That idea turned into a large debate with Stephen Mulhall's *On Film*.

In the contemporary contributions to FAP, the great majority of books, papers, and conferences come from what is called the "film-philosophy" approach to the idea. According to some its main protagonists, such as Robert Sinnerbrink, David Martin-Jones, or David Sorfa, this approach seeks to avoid the dichotomy between continental and analytic philosophy.

However, Sorfa admits that researchers working in this approach arrived at film-philosophy "through structuralism, post-structuralism and the liberationist critiques promised by Michel Foucault, Jacques Lacan, Jacques Derrida, Gilles Deleuze and Félix Guattari." (Sorfa 2016: 4). This is not the only indication that film-philosophy is bound to continental philosophy. In his whole defence of film-philosophy, Sinnerbrink argues for the existence of two ways of relating film and philosophy: what Noël Carroll and David Bordwell designate "Grand Theory" with continental influence (Bordwell and Carroll 1996), and what Sinnerbrink calls the "Analytic-Cognitivist," "Cognitivist-Naturalist," or "Anglo-American" tradition (Sinnerbrink 2011: 2-5).

Sinnerbrink identifies the "Analytic-Cognitivist" approach with "philosophy of film", accusing it of disenfranchising the philosophical role film could have. Such "philosophical disenfranchisement of film" turns film into an epistemically inferior endeavour (Sinnerbrink 2011: 128). Thus, instead of making a division between analytic and continental philosophy of film, he makes a division between film-philosophy and philosophy of film. The former would not disenfranchise film, whereas the latter would analyse, study and theorize about film. Furthermore, he argues that film-philosophy allows film and philosophy to mutually transform each other, asserting that film can be seen as engaging in philosophy in a "distinctively *cinematic* kind of thinking." (Sinnerbrink 2011: 7). In such an account, some films are aesthetically revealing in such a way that that they resist any traditional philosophical analysis, transforming philosophy itself.

In this line of argument, film-philosophers argue that philosophy of film (that they equate with analytic philosophy of film, despite resisting the dichotomy between continental and analytic philosophy) is unable to appreciate film's aesthetic particularities. So, Sinnerbrink's argument for film-philosophy generally derives from a false dichotomy he makes between film-philosophy and philosophy of film: whether we accept film-philosophy and do not disenfranchise film and respect its aesthetic properties, or we do philosophy of film and disenfranchise film by taking no interest in its aesthetic properties. Even, according to Martin-Jones, the division between film-philosophy and analytic approaches to film presupposes that the former explores film as philosophy, and the latter is philosophy of film. (Martin-Jones 2016: 7). For Sorfa, the film-philosophy approach affirms that films can do philosophy and, only after that, does it discuss whether that is the case for all films, or only a few of them, and the quality of the philosophy in such films (Sorfa 2016: 3). Film-philosophers, thus, see film-philosophy as the only approach able to defend the claim that films can philosophize.

However, this seems not to be the case. The debate around FAP started with Mulhall's defence of the idea, and at that point FAP was not questioned and debated in extension. Several books and papers were being produced with philosophical accounts of film, without properly questioning film's ability to philosophize. Philosophy of film (or analytic philosophy of film) questioned why film would do philosophy. Film as an art form, and philosophy are traditionally in different fields and have

traditionally different interests: why overlap them? It is thus fair to say that in analytic philosophy of film there is more scepticism regarding FAP, mainly represented by Paisley Livingston, Bruce Russell and Murray Smith. However, there are analytic philosophers of film that defend FAP, mainly Thomas Wartenberg and Noël Carroll. Sinnerbrink's accusations are not accurate since there are analytic philosophers of film who are very interested in film's *medium* and how film itself contributes to philosophy in a particular way.

We want to argue that the main difference between film-philosophy (that could be mostly seen as continental FAP), and analytic FAP, is not that the latter disenfranchises film. This assumes that all analytic FAP argues for the same position, as a particular theory only (an accusation usually made against analytic philosophy in general). On the contrary, we argue that the major differences are the methods and assumptions in both approaches. Film-philosophers assume film can philosophize in exclusively cinematic ways, and thus that their work is analysing particular films; analytics philosophers of film question whether this possibility exists, and try to argue why it does, both confronting what film's *medium* has to offer, and making meta-philosophical arguments against scepticism towards FAP. Their positions are, thus, not by default, as in film-philosophy, but, quite the contrary, a defensive position against scepticism. Such scepticism is usually a reaction against film-philosophy for their methods and presuppositions. It is also the position by default, since philosophy and film are, indeed, usually different endeavours with different priorities. This is the main reason why advocates for FAP in the analytic philosophy of film are accused of being too conservative, since they go from what philosophy is/does, to how films can do such things. We want to argue that this conservative position, which could be seen instead as a moderate one on the FAP spectrum, is the best position to argue for the idea that films can philosophize without falling into unargued assumptions.

Rigidity and the Status of Theoretical Identities

Luís Fernández Moreno (Complutense University of Madrid)

13.09, 17:55 – 18:30, Room: 5.2

Kripke holds the view that theoretical identities are necessary identity statements whose necessity follows from the rigidity of the natural kind terms contained in them. In my talk, I will firstly contend that the thesis can be held that natural kind terms are rigid designators whose referents are universals but, secondly, that the conception of natural kind terms as rigid designators of universals hinders the establishment of the truth of theoretical identities and hence of their necessity.

In the first two lectures of *Naming and Necessity* – henceforth (1980) – Kripke mainly deals with proper names, while in the third he focusses on natural kind terms. He claims that there are certain *similarities* between those terms and proper names, one of them being that both sorts of expressions appear in identity statements that, if true, are necessary, i.e., true with respect to every possible world. I will allude to those identity statements containing natural kind terms as "theoretical identities", one of the expressions used by Kripke. There are different ways in which theoretical identities have been understood, but I will assume Kripke's view, according to which theoretical identities are *identity statements* and natural kind terms are *general terms*. Kripke exemplifies theoretical identities by the statements "Water is H_2O" and "Gold is the element with the atomic number 79". Nevertheless, Kripke claims that the said similarity follows from another one, namely that proper names and natural kind terms are *rigid designators*.

Kripke stated later that the notion of a rigid designator (for singular terms) he intended in (1980) is the following:

> "[A] designator *d* of an object *x* is *rigid*, if it designates *x* with respect to all possible worlds where *x* exists, and *never designates an object other than x with respect to any possible world*." (Quoted in Kaplan "Afterthoughts", 1989: 569).

Kripke's definition of rigidity for singular terms leaves two options open, namely that a rigid designator designates the same object with respect to every possible world or that it designates the same object with respect to every possible world in which the object exists and has no reference with respect to every other possible world. However, the definition of rigid designation for *proper names* actually intended by Kripke is the first one, because in (1980: 21, n. 21) he asserts that he considers proper names as rigid *de jure*. A designator is rigid *de jure* if at fixing its reference it is *stipulated* that its referent is the same with respect to every possible world.

Having reached this point, I should undertake two tasks. First, since Kripke has not explicitly presented a definition of rigidity for natural kind terms, I will propose a definition of rigid designation for them. Second, I will put forward a view of *natural kinds* that is consistent with Kripke's few remarks about them.

Regarding the first task, I will extend the definition of rigid designation for singular terms intended by Kripke to natural kind terms. The most literal extension, and the only one I will take as a basis for my considerations, is the following:

> A designator *d* of a natural kind *k* is *rigid*, if it designates *k* with respect to all possible worlds where *k* exists, and *never designates a kind other than k with respect to any possible world*.

This characterization of rigidity for natural kind terms leaves *two* options open for natural kind terms. However, I will assume the first option as in the definition of rigidity applicable to proper names.

The view of natural kinds that is more in keeping with Kripke's few remarks is to conceive them as a type of universal instantiated in particular entities; this is precisely the dominant view among the authors who accept the rigidity of natural kind terms. Now, since theoretical identities express the identity of natural kinds, I will assume this *necessary condition* for the identity of natural kinds: if two natural kinds are identical, the instances of each kind are the same in all possible worlds. In the following I will centre my considerations on the theoretical identity "Water is H_2O".

As already indicated, I have adopted a view of natural kinds, according to which the substances (natural kinds) water and H_2O are universals - this makes it possible for the corresponding natural kind terms to be rigid designators. Thus, the statement "Water is H_2O" will be *true* if and only if the substances water and H_2O are identical. However, according to the necessary condition for the identity of natural kinds and hence of substances proposed above, this will hold *only* if the instances of both substances are the same in all possible worlds, i.e., only if the terms "water" and "H_2O" are coextensive in all of them. Notwithstanding, even if we concede that these terms are coextensive in the actual world, from the rigidity of the terms "water" and "H_2O" and their coextensiveness at the actual world it does not follow their *coextensiveness* in *all possible worlds*, which is the necessary condition to be satisfied for the truth - and the necessity - of the statement "Water is H_2O". In other words, even if we accept that the terms "water" and "H_2O" are rigid designators and that the extension of those terms is the same in the actual world, this does not lead to the conclusion that their

referents - the universals designated by them - are identical, i.e., that the theoretical identity “Water is H_2O”, conceived as expressing an identity between substances (universals), is true, although if it were true, it would also be necessary.

The conclusion to be drawn from the foregoing considerations is that, although it can be held that natural kind terms are *rigid designators*, the view of natural kinds as *universals*, which allows for natural kind terms to be rigid designators, hinders the establishment of the *truth* of theoretical identities, and given their truth, of their *necessity*.

The Denotation of Copredicative Nouns: Truth Conditions and the Counting Puzzle

Marina Ortega-Andrés (University of the Basque Country)

14.09, 15:40 – 16:15, Room: 5.2

Copredication occurs when the same polysemous nominal expression comes along with simultaneous predications selecting for two different meanings or senses in a sentence. For example:

(1) The books are heavy and interesting.
(2) The school caught fire and was celebrating 4th of July when the fire started.

In (1) ‘book’ refers the physical object, but also the information that it transmits. In (2), the word ‘school’ refers the building and the people inside the building. One of the issues about these nouns is that it is not clear what their denotations actually are. This fact has been used as an argument against semantic externalism (Pietroski 2005; Chomsky 2000; Collins 2017): semantic externalism says that the truth conditions of a statement rely on a correspondence between the content of the words of the statement and real world entities. That means that the truth conditions of (2) rely on the existence of an entity that is a building that caught fire and a group of people that celebrate 4th of July. Because there is no such entity in the world, semantic externalism must be incorrect.

One response to the argument is that nouns that allow copredication denote complex entities that have two (or more) different parts or aspects (see Arapinis and Vieu 2015; Gotham 2016). This mereological hypothesis generates some metaphysical puzzles. For example, the persistence conditions of the alleged complex entities are not clear. Consider the following sentence:

(3) London is so unhappy [people], ugly [architecture], and polluted [area] that it should be destroyed [architecture] and rebuilt 100 miles away (Chomsky, 2000).

If we suppose that the buildings and streets are destroyed and the population decides not to move to the new London (100 miles away), then the following statements could be true:

(4)
a. London [old architecture] has been destroyed.
b. London [new architecture] is still ugly.
c. London [population] refused to move to its [new architecture] new location and [population] ended up settling down in a different place.

Someone that maintains that denotations of copredicative nouns are complex objects has to

commit to the view that the whole would persist even when some of its co-constitutive parts are pulled apart. The complex entity would be able to persist in several different entities, giving rise to too many Londons.

My proposal is that the truth conditions of copredicative sentences derive from a process of sense de-compilation and the assignment of each predicate to its respective denotation. Different senses that de-compilate are related by realization/explicative relations, that is, they are different ways in which our representation or prototype of a school is presented to us in the real world. If we think that the prototype of a school is an educative institution, then the relatedness between the aspects is defined in terms of what an educative institution will normally require for its being in the world. For example: a physical space (the building) and the people that participate in the education process.

The NP 'school' is a compilatory term that binds various different aspects of the concept of school. In (1) 'book' has two different denotations that are in a relation of realization: the readable content and its physical realization (a set of written pages or a volume). 'School' in (2) refers a building that is the physical realization of an institution and a group of people that are located in the building. According to this idea, sentences (1) and (2) could be paraphrased as follows:

(1') The books [physical realizations of content-books] are heavy and the books [the contents-books] are interesting.
(2') The school [physical realization of institution-school] caught fire and the school [group of occupants [of the physical realization]] was celebrating 4th of July when the fire started.

One puzzle that copredication generates is the counting problem (see Asher 2011; Gotham 2016). Consider (5) in context C:

(5) There are three interesting and heavy books.
C: I have three copies of a volume of three different books (informational contents) from Tolstoi; the three contents are interesting and the three copies are heavy.

Intuitively, (5) should be true in C: there are three books that are interesting and heavy. However, if 'book' in (5) denotes two different parts of a complex entity, then, there are three physical-books and three informational-books. That means that there are six books (or parts of books) instead of three and that (5) would be false in C. However, if we consider that the word 'book' in (5) is a compilatory term that has two different denotations –as it is proposed here-, then (5) could be paraphrased as follows:

(5'): There are three interesting books [informative contents] and the books [physical realizations [of those informative-books]] are heavy.

The books (informative contents) are physically realized by the physical object or volume that is heavy. Therefore, the sense of 'book' as object stands for the physical realizations, while the sense of 'book' as content stands for the three informational-contents. The truth conditions of (5) are the following:

There is a plurality of three books [informative contents], such that:
(a) they are interesting;
(b) their physical realizations are heavy.

In this case, we count three book-contents, whose physical realizations are heavy, so (5) is true in C.

In conclusion, the proposal explains some problematic issues from the semantic and philosophical discussion about copredication: it gives a response to the argument against semantic externalism without having to postulate the existence of complex objects and it solves some puzzles like how we count books in copredicative sentences.

Content Variation in Epistemic Conditions

Niall Paterson (University of Bristol)

14.09, 15:40 – 16:15, Room: Sala de Actos

A number of contemporary epistemologists (Bird 2007, Pritchard 2012) endorse epistemic subjunctive conditionals. For example, on Bird's acccount, *S* is justified in judging that *p* only if there is a possible world, holding fixed *S*'s mental states, in which *S* forms a corresponding judgement which yields knowledge of *p*. On Pritchard's account (amongst many others), one knows that *p* only if one's belief is *safe*, that is one could not easily falsely believe that *p*.

Such conditionals are notoriously susceptible to counterexamples involving necessary propositions. For example, there seem to be cases where one could be justified in believing a necessary falsehood. Nonetheless, there is no possible world, holding fixed one's mental states, in which one forms a corresponding judgment that yields knowledge of *p*, for knowledge is factive, and *p* is necessarily false. Furthermore, the theoretical role *safety* plays (namely, the elimination of beliefs that are true by luck) seems to require that one could unsafely believe a necessary truth. But one cannot falsely believe a necessary truth, so one's belief in a necessary truth must be safe as defined.

In response, both authors allow for content variation in their conditionals. That is to say, they allow the target proposition to vary slightly. One can be justified in judging a necessary falsehood, for there is some proposition that differs slightly in content that one would form a corresponding judgement of, that yields knowledge. Further still, one could unsafely believe a necessary truth, as one could easily believe some other proposition that differs slightly in its content, where the content of the belief is false.

In this paper, I argue that this strategy fails. Even if one allows for content variation, there are analogous cases that serve as counterexamples to the accounts. The reason for this is fairly simple: in both cases, we can construct a case in which the belief/judgement formation is sensitive to the content across certain classes of possible worlds. I focus on cases which involve what I call a *hunch*. A *hunch* is a kind of dispositional profile – one has a *hunch* iff. (1) one is disposed to believe/judge *p* under some conditions C, but (2) not to believe/judge p^* under C, where $p^* \neq p$. I will show that, using *hunches*, one can construct counterexamples to the modified accounts. I finish by suggesting a different strategy for each condition to avoid the problematic counterexamples.

Fit is no Good

Alain Pe-Curto (Yale University)

13.09, 14:30 – 15:05, Room: Mattos Romão

In the *Principia Ethica*, Moore (1903) introduced the principle of organic unities, according to which we should not assume that the intrinsic value of a whole is equal or proportional to the sum of the intrinsic values of its parts.[20] Like Nozick (1981), he took this principle, to reveal something both essential to and paradoxical about value.

I argue that one of the most compelling additivist responses to Moore's and Nozick's conception of intrinsic value as essentially organic, namely the value atomist response[21], conflates values and the distinct normative feature of fit. The problem for such an approach is that fit, I claim, is not a value property and, therefore, it is inapt to play the role the thick atomist needs it to assume in her rejection of organic complexes.[22] My claim, of course, needs to be unpacked, which I undertake in the paper. Nevertheless, this preliminary summary of my argument already hints at how the opposition between Mooreans and value atomists regarding organic unities pertains to the very nature of value.

I consider more precisely on what I call "Thick atomism" (Oddie 2001a/b, 2005). Thick atomism is a form of value atomism because, in short, it takes the parts of a whole to behave like separable atoms, as far as intrinsic value is concerned. Indeed, it defends that we can decompose a complex whole into parts such that (1) they have intrinsic values and (2) the sum of their intrinsic values is equal to the value of the whole. I should note that atomism shares Moore's understanding of intrinsic value as locally supervenient (e.g. Moore 1922), so that an intrinsic value-bearer has its intrinsic value solely supervene on its intrinsic nature and can therefore separate from the whole it belongs to without its intrinsic value being affected (on separability, e.g. Broome 1991, Oddie 2001a/b,).

By contrast, the Moorean theory of organic unities is a kind of value holism. Value holism can be defined as the negation of (1), (2), or both.[23] Moore denies (2) specifically. Moore is doubtless in good company,[24] but one should not forget the advantages of particularly powerful forms of atomism, such as Thick atomism.

I call the latter "thick" for the following reason. This form of value atomism directs its attention to the thick evaluative properties – in other words, "substantive values", such as moral goodness, beauty, and pleasantness –[25] that are identified in the typical presentation of alleged organic complexes. It then assumes the additivity of intrinsic value as a matter of "regulative ideal", and, according to my interpretation, proceeds to select which thick intrinsic values are to be acknowledged within our substantive axiology on the basis of this additive assumption. Thus, Thick atomism assumes additivity in order to identify substantive values, instead of starting from intuitions regarding what substantive values there are in order to assess the truth or falsity of the additivity of intrinsic value. An epistemological slogan for Thick atomism could therefore be "formal axiology first, substantive axiology second".

[20] For a recent overview, Orsi (2015, Sect. 5.3), cf. Carlson (1997).
[21] E.g. Oddie (2001a/b), (2005), Zimmerman (1983), (1999), (2001), (2014); cf. Feldman (2000), (2004).
[22] On fit, e.g. Howard (forthcoming).
[23] E.g. Brown (2007) for other kinds of holisms.
[24] E.g. Brown (2007), Lemos (1994: 32-48), (2010), (2015), Chisholm (1968), (1986: 69-102), Hurka (1998), (1998), (1994) and Nozick (1981: 403-450). Cf., for contextualist value holism, Dancy (2003), (2004), (2005), Fletcher (2010), Kagan (1988, 1998).
[25] On thick and thin, e.g. Smith (2013), Tappolet (2004), (Williams, 1985).

I put into question the legitimacy of such an approach with three specific objections and one general explanation for the defectiveness of the account, namely that it takes fit to be a value.

First, I point out that thick atomism distributes intrinsic value across the mereological structures of wholes in an unsuitable way. The value bonus or malus (compared to the sum of the intrinsic values of the parts) that the theory of organic unities attributes to the whole is relocated to one of its parts by the thick atomist. The account consequently yields bizarre understandings of specific wholes, as I emphasize in detail in the paper. This is the "misplaced value" objection.[26]

Second, I pinpoint that the revision of the list of substantive values needed by the thick atomist fails to meet one of the constraints that, as the atomist acknowledges, it should meet: mutual independence of these values (This is unfortunate because mutually dependent factors yield additivity trivially.). I highlight that, although the factors are logically independent, they are still hyperintensionally dependent. This is the "dependent features" objection.

Third, I take a step back to evaluate the purported benefits of assuming additivity as an heuristic tool to help us identify substantive values. I put to the test the atomist proposal by applying it consistently to the variety of alleged organic complexes one can find in the literature. This exercise reveals that the new substantive values which the atomist suggests solely include normative features that are in fact fit features and not value features. This is the "fit is no good" objection.

After showing that these new allegedly evaluative properties, such as being aesthetic *calibrated, being hedonically appropriate or getting one's just deserts are in fact features*, I explain how the *fit is no good* point in fact underlies the misplaced value objection and the dependence objection above. It also sheds light on the intuitive feel of 'ad-hocness' of the thick atomist response to the theory of organic unities and its failure to fulfill its promise that additivity, as an heuristic assumption, would help us separate the wheat from the chaff in our list of substantive values. Finally, I end on an original framework to distinguish more systematically between normative features, in particular deontic, fittingness, and evaluative features. I explain why it matters not only for the debate on organic complexes, but also, as anticipated by Moore and Nozick, for the very nature of value.

Perceptual Seemings, Epistemic Overdetermination, and Defeat

Tommaso Piazza (Università di Pavia)

13.09, 15:40 – 16:15, Room: Sala de Actos

It is a commonplace in epistemology that the justificatory status of one's beliefs depends on the mental states one is in; for instance, on which perceptual experiences, or on which further justified beliefs, one has. It is not often sufficiently emphasized, however, that we talk of the contribution of one's mental states to the justification of one's beliefs in at least two different senses. In one first sense, when we say that a mental state X contributes to the justification of S's belief B what we mean is that B is (or can be) justified on the basis of X. In a second and somewhat looser sense, when we say that X contributes to the justification of B what we mean is that B is (or can be) justified on the basis of, not X, but the introspective belief that X occurs in S's mind. To distinguish between both senses, it is useful to talk of X's *direct* contribution in the first case, and of X's *indirect* contribution, in the second case.

[26] Cf. the analogous point – against organic unities – made by Olson (2004).

When X is a belief, it normally cannot contribute at the same time *directly* and *indirectly* to the justification of the same beliefs. If it contributes *directly*, it cannot also contribute *indirectly* (and vice-versa). Perceptual seemings appear to behave differently. When it perceptually seems to S that P, S can form the justified belief that P simply on the basis of the seeming. But S, if her background knowledge is just normal, can also form the justified belief that P by inference from the introspective belief that she has the seeming (Pollock 1987; Pryor 2000; Cohen 2002).

When S's seeming that P contributes directly and indirectly to the justification of S's belief that P, this justification is *overdetermined*, as it flows from two different sources: S's *seeming* that P and S's introspective *belief* that she has the seeming that P. In general, when it flows from two different mental states M1 and M2, S's justification for believing P is to be expected to be *better* than it would otherwise have been, and in particular to be more *stable*. Since M1 is different than M2, in fact, what undercuts S's justification flowing from M1 need not also undercut S's justification flowing from M2. Much to the contrary, unless the relevant undercutting defeater calls into question S's epistemic reasons *en masse*, it is to be expected to leave the justification flowing from M2 unaffected. When M1 is the *seeming* that P and M2 is the introspective *belief* that S has the seeming that P, this general expectation is however intuitively frustrated. For it is very intuitive that evidence undercutting the inference to P from the introspective belief that S has the seeming – for instance, evidence that S's seeming in not a reliable indication of P's truth – *eo ipso* also annihilates the justification for P from the seeming itself.

According to the standard explanation, possibly stemming from J. Pollock (1987), this happens because an undercutting defeater of S's *indirect* justification from the seeming also undercuts S's *direct* justification from the seeming. On close inspection, this answer is however unsatisfactory. For it is not clear why information bearing on the propriety of the *inference* from the premise that S has the seeming that P to the conclusion that P should harm justification, like the one directly flowing from the seeming, that is not based on such an inference. In this paper I will consider some alternatives to this traditional explanation, and propose my own account of how an undercutting defeater of the *indirect* justification flowing from the seeming can also harm or downgrade the *direct* justification flowing from the seeming in spite of its non-inferential nature.

The talk, in particular, is articulated as follows. Firstly, I introduce the distinction between the direct and the indirect contribution of a mental state to the justification of S's beliefs. Secondly, I rehearse Pollock's distinction between undercutting and rebutting defeaters. Then I observe that an undercutting defeater of S's *indirect* justification from the seeming that P has also the power to downgrade S's *direct* justification from the seeming, and criticize the standard explanation of this phenomenon stemming from Pollock. Finally, I critically examine two alternative explanations, and I present my own account in some detail.

The Science of Ethics and Evolutionary Moral Realism

João Pinheiro (Centre for Philosophy of Sciences of the University of Lisbon)

14.09, 15:05 – 15:40, Room: Mattos Romão

Since Darwin [1871:§4-5], but especially over the last few decades, Ethics has seen its empirical study scientifically developed, becoming itself a cross-disciplinary field of Applied Science, with import from such a myriad of disciplines as Evolutionary Biology, Ethology, Evolutionary, Developmental, Social, and Moral Psychology, Cognitive Neuroscience, Sociocultural Anthropology,

Ethnography, etc.. When brought together, the insights from these disciplines allow us to model Ethics as a facet of human behavioural ecology [e.g. Flanagan 1996, and Flack & Krakauer 2009], where moral cognition is subject to EvoDevo constraints (not unlike any other biological trait), and plays a role in the construction of our social niche, for which it can be causally selected.

Whence the origin of its biological function. The biofunction of moral systems is to control interactions between individuals (and, possibly, groups), so as to maximize individual (and, possibly, group) fitness, which will consist, by and large, in (potential) social conflict management, and therefore in boosting cooperation, which has been shown [e.g. Hardin 1968, Axelrod & Hamilton 1981, & Nowak 2006] to be the behavioural strategy that best benefits all individuals' (and groups') fitness, were all the interactors to cooperate.

However, despite the proliferation of scientific approaches to Ethics stimulating methodological naturalism in Metaethics, there remains serious doubts regarding the tenability of a fully naturalistic and realist programme, especially because moral facts, understood realistically (i.e. as independent from our evaluative attitudes), and values, are still widely believed to be non- analysable in scientific terms.

The pseudo-dilemma posed by "Street-style" evolutionary debunking arguments of moral realism [Street 2006] are exemplary of this scepticism [cf. Copp 2008, & Kahane 2011]. These arguments, founded on evolutionary theory, depend on the truth-value of one fundamental empirical & epistemic premise: moral cognition is truth-tracking in virtue of its natural history. Naturalists sceptical about moral realism (as previously conceived), claim that, while denying the truth of this proposition, the realist is left blind as to the possibility of a naturalistic explanation of moral realism; whereas while accepting it, the realist is left with the difficult task of providing us with a cogent "just-so-story" of truth-tropic cognitive evolution.

Until quite recently, most naturalists would bite the bullet and agree that the empirical hypothesis that there's only positive selection for truth-tropic moral cognition is false, thus recognizing the frailty of the moral realist's position. But it so happens that there's a novel scientific research programme/ metaethical theory which might just turn the tides. According to this theory, which we'll call evolutionary moral realism, moral truths are grounded in facts about cooperation [e.g. Brosnan 2011, and Sterelny & Fraser 2016] or sociability [e.g. Copp 2007 & 2009], or, if you prefer, there are facts against which moral propositions can be judged to be true or false, and these are «objective facts about the practices and norms that would promote stable cooperation» [Sterelny & Fraser 2016:5].

This theory establishes a version of moral realism compatible with the thesis that «the primal scene of morality (...) is (...) one in which we do something together» [Korsgaard 1996:275] (fitting nicely in the social niche construction theory of morality as part of human's behavioural ecology). What is more, by equating payoffs with fitness values and strategies with behaviours according to norms, evolutionary game theoretical approaches formally predict which moral practices will become evolutionary stable strategies in specific environments, for normative strategies that maximize utility will thrive, meaning that they'll enjoy of higher fitness-values [cf. Hammerstein 2012, el Moulden et al. 2012, & Sterelny 2012].

As Street would argue, one should accept the empirical hypothesis that natural selection is blind to truth-tracking as being more parsimonious. But, in fact, epistemic rules of inferences to the best explanation actually recommend we pick the loveliest of alternative explanations [Lipton 1991 & 2000], and, according to evolutionary moral realism, evolutionary explanations of moral cognition in terms of fitness-tracking and in terms of truth-tracking are not alternatives, but, rather, complementary explanations at different levels of analysis (at "the consequence level", and "the ecological

level", respectively) [cf. Goode & Griffiths 1995, and Griffiths & Wilkins 2015], such that one may safely infer both.

However, and notwithstanding the truth-tropic nature of our moral cognition, it remains to be shown whether «in favourable cases there is a reliable causal connection between moral opinion and these facts» [Sterelny & Fraser 2016:17] about which behavioural strategies are evolutionarily stable. For example, if the evolutionary psychologists are right in that our cognition is adapted to the environment of the Pleistocene [Richerson & Boyd 2000], then, given the role played by kin & group selection in human evolution, and because we inhabit substantially different environments from the Pleistocene, we're unlikely to produce as many adaptive moral beliefs qua moral truths when inter-group conflict and cooperation with strangers are in case [Robinson 2013:75]. As such, our moral cognition might not be reliable in our present and global state of affairs, and moral error theorists [e.g. Joyce 2016] may still rejoice in the fact that moral error is rife.

Notoriously, evolutionary moral realism also has something to say regarding the question of the naturalization of values: irrespectively of how we're to understand the semantics of moral good (e.g. as well-being), it must be fitness-enhancing. Contrafactually speaking, if weren't so, moral cognition would be selected against [cf. Enoch 2010]. But then Moore's purportedly destructive "open-question argument" enters the scene: is fitness-enhancing good? Well, being naturalists, as we are, such an equation comes, quite naturally, a posteriori, nonetheless remaining an informative analytic truth. But moral value, while functionally reduced to fitness-value, is extrinsic [cf. Brandon 1990:39], such hat fitness-enhancing is only good for the organism/from the organismic point-of-view, and not for the Cosmos/from an Archimedean perspective, contrary to some naturalists' will [e.g. de Lazari-Radek & Singer 2014].

Fitness's extrinsicality is also the reason why one's to expect some measure of pluralism about norms (though not necessarily about all norms) to be true. This is because the functionality of norms depends on their environment, viz. norms ought to be understood as environment-relative (purported) solutions to problems, where the environmental feature to which they constitute a solution might not be universal, nor repeatable, given that different evolving populations are subject to partially different and changing environments [e.g. Wong 1995, & Dean 2012]. The resulting picture is one where moral truths are objective/"mind-independent" (albeit not necessarily universal), and environment-relative (including socio-cultural sensitivity).

The Problem of the Ontological Value of the Variable in Russell's Theory of Denotation

Ivory Pribam (Université Panthéon Sorbonne Paris 1)

14.09, 15:05 – 15:40 Room: 2.13

My paper discusses the ontological value of the variable in Russell's theory of denotation. The variable is essential in Russell's theory of denotation, which logically supported what he thought objects are and how to talk of them. Yet, the use of the variable poses problems for Russell's theory of denotation. These problems were discussed by Russell, Frege, and Moore, but have gone largely unaddressed within the history of philosophy. It is the intention of this paper to bring these problems to light. The problem can be summarised as follows: the variable is the fundamental denoting-position of a formal theory that is meant to explain the structure of the ontological. If such a formal theory is meant to ground the ontological, then the formal cannot solely be symbolic but

must also represent the actual structure of the ontological. Yet the variable, the fundamental symbol of denotation in a theory that defines objects, is ontologically indefinable per Russell's epistemology and ontology.

I first set out Russell's ontological, epistemological, and theory of denotation where applicable. I then explain the problems the variable poses within Russell's epistemological-ontological system according to the aforementioned three founding figures of analytic philosophy. I address in Russell's ontology how all existents could be reduced to particular relations and two kinds of entities: particulars and universals. Particulars can be subdivided into objects and simples (although the latter were later abandoned), and universals into universals of quality and relation. I explain how Russell brings all objects down to existing in the same way but able to be grouped into sets per their logical structure.

Russell's ontology is intimately connected to his epistemology. As such the ontological value of the variable need be consistent with his epistemology. For Russell, what we know of the ontological is either through knowledge by acquaintance or knowledge by description. This I elaborate in accordance with his ontology.

The final piece of setup is Russell's theory of denotation. For Russell, this provides the logical structure of objects and how they are referred to. The basic structure of the theory of denotation is the formula C(x). C represents the necessary constant in meaning-position, providing information about x. x in entity-position is a variable standing for the variable denotation of particular entities. So when C is given, the variable is replaceable by a constant – the denoted object has been described, and so determined.

But only the fundamental type-instance, which is the type-formula of definite denotation represented as C('x); is of concern. It symbolises any singular instance of denotation without constant predication having been given.

Having explained this background, I discuss the problems of the variable of the type-instance as explained by Frege, Moore and Russell. Firstly, Frege highlighted the ambiguous dual role of the variable in one of Russell's attempts to define it. He explains the ambiguity comes from trying to define the variable through its logical function both as something which is assumed determinate but unknown and as an entity that is itself indeterminate. Frege notes that what is needed instead is an ontological definition of the variable outside of its symbolic use. However, no such definition is forthcoming as Frege concludes. The only plausible definition Frege sees is that the variable is one of a certain set of values, without its being decided which one. However, he notes, being the value of a set is a characteristic of each and every denoted entity in Russell's theory of denotation. In other words, this merely states of the variable that it exists without coming any closer to an ontological definition of it.

Elsewhere Russell opted to talk of the variable ontologically as an indeterminate entity. The variable is itself anything, where anything is necessarily singular and undetermined. However, this definition faces problems in relation to Russell's theory of knowledge of objects.

Moore raised these problems, asking whether the variable is an object, or something other than an object. In response Russell suggested that it might be something of which we have immediate acquaintance, but which is not an entity. Alternatively, he suggested it might be an entity, but an indeterminate one.

As Russell admits; both options are problematic. I argue that the following explains why: If the variable is something of acquaintance, it cannot be an object. An object of acquaintance is necessarily determinate, whereas the variable is necessarily indeterminate. But, if it is not an object,

it is unclear what else the variable might be within Russell's conjoint epistemology and ontology. Further, if not an object it would logically have to be also represented in meaning-position. But anything placed in meaning-position is necessarily a constant. But the variable is, in essence, not a constant. If something of acquaintance the variable creates inconsistency between Russell's logic and his theory of knowledge of objects.

Supposing the variable is instead an object but not of acquaintance. The first problem is that we can only entertain a proposition in Russell's epistemology if we are acquainted with the denoted object or with the constituents that describe it. Yet as the variable is a fundamental constituent in the fundamental formula of his canonical notation that grounds his epistemology, it should be something of acquaintance.

Additionally if an object but not of acquaintance, it should work within the theory of denotation as an object of description. But a problem arises when we try to formally denote the variable as a singular ontological existent. In the theory of denotation, once predication is given, the denoted object is determined. Logically, this means that if we denote the variable with predication, we have determined it. But, in something close to a paradox, this means we have determined a fundamentally indeterminate entity.

My conclusion is that there is no ontological and logical reconciliation of the variable in Russell's philosophy. When Russell stated "I only profess to reduce the problem of denoting to the problem of the variable;" this should have been a red flag for the congruency of certain aspects of his theory of denotation, ontology and epistemology.

The Ontology of Musical Versions: Introducing the Hypothesis of Nested Types

Nemésio Puy (University of Granada)

14.09, 14:30 – 15:05, Room: 2.13

This paper is devoted to explore the ontological status of musical versions. A version is usually defined as a revision of a previous work, involving modifications on a previous sound structure and, sometimes, on the original instrumentation (cf. Davies, 2007; Rohrbaugh, 2003). For instance, the flute solo of the third movement of Tchaikovsky's Symphony No. 2 is completely remodelled in the 1879 version with respect to the 1872 version, and the melody of the A theme proceeds now by arpeggios on the violins, rather than by syncopes.

We currently say that musical versions are of a musical work. We speak, for example, about 'the 1872 version of Tchaikovsky's Symphony No. 2'. The way in which we typically refer to versions suggests a kind of relation between a version and a musical work. Accordingly, a widely shared intuition in our musical practices is that in a performance of a version we are hearing the musical work this version is said to be of. In a performance of the 1872 version of Tchaikovsky's Symphony No. 2, and in a performance of the 1879 version of Tchaikovsky's Symphony No. 2, listeners broadly assume to be hearing the same work: Tchaikovsky's Symphony No. 2. A one-many relationship is assumed to hold between Tchaikovsky's Symphony No. 2 and the performances of its versions indicated in 1872 and 1879. These performances are regarded as occurrences of Tchaikovsky's Symphony No. 2 in which we can encounter, experience, hear and know this very work. Consequently, three things are assumed to be heard in a performance of the 1872 version of Tchaikovsky's Symphony No. 2:

1. That particular performance.

2. The version that this performance is of (the 1872 version of Tchaikovsky's Symphony No. 2).
3. The work that this version is of (Tchaikovsky's Symphony No. 2).

Let us call it the standard view on versions. It challenges a broadly expanded view within the ontology of music, labelled here as structural monism: the idea that a musical work is individuated by one, and only one, sound structure. Since a work's versions always exhibit different sound structures, ontological views assuming structural monism seem not to be able to accommodate the standard view. Versions would count as different musical works from the works versioned and, consequently, the work versioned would not occur in the performances of its versions.

Type/token theories, which identify musical works with the ontological category of types, have been typically considered as the best explanation of the ontological nature of musical works (cf. Wollheim, 1980; Wolterstorff, 1980; Davies, 2003; Dodd, 2007). However, type/token theories are presumed to endorse structural monism and to be committed to the revisionary consequences for the standard view on versions presented above. For Dodd's theory, versions would be different musical works because they exhibit different sound structures and, hence, they determine different sets of conditions for something to be a properly formed token of a type (cf. Dodd, 2007: 90). A similar problem is faced by the action-type theory (cf. Currie, 1989: 70). In addition, for the initiated-type theory, versions constitute different musical works from the work versioned not only in virtue of structural differences, but also for being indicated at different times and, sometimes, by different composers (cf. Levinson, 2011: 79-80). The thesis defended in this paper is that the hypothesis of nested types can preserve the explanatory virtues of a type/token theory without entailing these revisionary consequences, being able to accommodate both structural monism and the standard view on versions.

According to the hypothesis of nested types, a musical work is a higher order type that is instantiated in lower order types. Versions are lower order types that instantiate a higher order type (a work) and that are instantiated in musical performances. It will be shown that types appealed by the hypothesis of nested types are types in a full sense. Firstly, they are not an ad hoc category, having applicability in other domains different from the ontology of musical works. Secondly, they exhibit a feature that distinguishes types from properties or other universals (cf. Dodd, 2007: 17; cf. Trivedi, 2002: 74): the transmission of predicates between higher order types and their tokens (lower order types), and between higher order types and the tokens of lower order types (performances). Thirdly, the explanation of the access to a type by means of its tokens provided by the notion of deferred ostension is also available for the hypothesis of nested types (cf. Quine, 1969: 40; Dodd, 2007: 11): since a token 'stands proxy for the type' that lies behind it, we hear a lower order type (a version) passing through its tokens (performances), but since a lower order type is also a token of a higher order type (the work versioned), we hear the higher order type passing through the lower order type. Perceptibility is transmitted from tokens of a lower order type to the higher order type of which that lower order type is a token.

In addition, the hypothesis of nested types can explain how the versions' different sound structures can be associated to a same musical work without rejecting structural monism. Two compatibles explanations can be provided:

1) The differences between the sound structures of a work's versions may be explained in terms of imperfect instantiation: versions are not properly formed instances of the higher order type. This explains those cases in which a composer is not satisfied with a previous version and wants to improve it.

2) The differences between the sound structures of a work's versions may be regarded as a consequence of the variability of the properly formed instances of a higher order type. Accordingly, the sound structure that individuates the work qua higher order type is more generic than the ones that individuate any of its versions. It has variables to be filled in different ways by the sound structures of its lower order types. This explains those cases in which a new version is understood as a new right mode of presentation of a same work.

Ambivalence is Central to Personhood, Mental Unity, and Practical and Epistemic Rationality

Hili Razinsky (LanCog, Universidade de Lisboa)

14.09, 17:55 – 18:30, Room: 2.13

An ambivalent person may want and yet not want one and the same thing, love and hate the same person, find a job offer fair and yet unfair, or ambivalently believe and disbelieve that her spouse is cheating on her. Yet how should ambivalence be understood? In my work I analyse ambivalence in terms of two opposing attitudes, while, complementarily, ambivalence is revealed as a unitary way to engage with people and things. Defending the oppositional unity between the plural poles of ambivalence, I also give a new account of personhood and attitudes. I argue, with Donald Davidson, that the plurality of the person's engagements (i.e., attitudes, behaviour etc.) presupposes and constitutes a unitary and basically rational person. In contrast to Davidson's view, ambivalence does not undermine personhood thus understood, but rather it forms a basically rational connection between attitudes.

The character of this connection changes however from one instance to another. Thus, a person can have opposing attitudes towards a single object, or have a negative attitude towards another attitude she maintains (e.g., she dislikes her judgement that John is unpleasant). Again, we can act in creative ways on both opposing attitudes together, or live a limited life, or resolutely act on single pole. I focus in particular of four modes, each consisting in having opposing mental attitudes of a particular type, namely, emotions, beliefs, value judgements and desires.

Ambivalence bears on various issues from the phenomenal unity of consciousness, to emotional integration, self-deception, cognitivism and non-cognitivism about value judgement, political agency and artistic expression. In studying ambivalence I turn to reexamine these other issues, and in some areas, including consciousness, emotion, value, belief, action, rationality, defend new accounts.

Ambivalence is predicated primarily on the person and her mental attitudes, while a mental attitude may be understood, as an intentional disposition to behaviour, consciousness and other attitudes. Qua disposition, an attitude goes beyond an action or a thought. Yet this should not suggest that a momentary consciousness or action cannot be ambivalent. I argue to the contrary in both cases. Qua intentional, the linguistic concepts underlying the person's deeds and attitudes do not generally constitute attitudes. It is however essential to concepts that they may become ambivalent attitudes, and *conceptual ambivalence* is also an aspect of central forms of ambivalent judgements.

Ambivalence is an ordinary and central possibility of human life, successful agency, epistemic achievement, moral conduct and wellbeing. Yet a wide variety of accounts of personhood, action, consciousness, belief, emotion, value, basic rationality and language make strict ambivalence impossible. If ambivalence cannot in fact be dispensed with, it requires to be investigated together

with these other elements. In this talk I will make some suggestions about the possibility and character of ambivalence and of its different forms, and ask how ambivalence bears on such issues as mental unity, intentionality, action–desire relations, and the logic of belief.

Three Stages in the Meta-History of Analytic Philosophy

Henrique Ribeiro (Universidade de Coimbra)

13.09, 17:20 – 17:55, Room: Mattos Romão

The main purpose of this presentation is to show how, from its beginnings to the present, analytic philosophy has successively developed different representations of the place it occupies in the history of western philosophy which have identified it and constituted it as such ("analytic philosophy"), particularly as an entity essentially distinct from the so-called "continental philosophy". The expression and the concept of "analytic philosophy" were already episodically and circumstantially used, both by Russell and by the Viennese logical positivists, but they only started to became current, as designating a specific philosophical school, since the 1960s. A significant, international example is the publishing of a book by Jean Wahl *et al.* (eds) (1962), in connection with a Colloquium which had the participation of distinguished analytic philosophers such as Alfred Ayer, Gilbert Ryle, and Willard van Quine. At the time, the official historian of the analytic movement was James Urmson, who also took part in the Colloquium. As becomes clear in his presentation, the movement was far from constituting a philosophically unified movement with convergent aims and approaches. Urmson distinguishes at least three different strands: the "Cambridge school", which included Russell, along with Wittgenstein; logical positivism (in its Viennese version); and the "Oxford school", which the analytic philosopher Michael Dummett later termed "English ordinary language philosophy" (Dummett 1978, 432ff) and which included, notably, Gilbert Ryle, Peter Strawson, and John Austin (Urmson 1962). A representation such as Urmson's may have two or three meta-historical implications: first, there clearly existed at the time no common guidelines in the analytic movement, i.e., the concept of analysis was not the same in all of the abovementioned approaches (in the text mentioned, Dummett goes so far as to say that the worst enemy of the Oxford philosophers was not Martin Heidegger, but rather the positivists); second, this was not yet about foundations for analytic philosophy having to do with this or that specific philosopher, although Russell's and Wittgenstein's role is acknowledged (while Gottlob Frege's, for example, is completely ignored); and lastly, in a representation like Urmson's, analytic philosophy is essentially British, or devoid of a trans-national dimension (there is no place in it for authors such as Quine, the American philosopher who was to have a key role in the development of this philosophy).—Let us call the representation I have just briefly described "the first stage in the meta-history of analytic philosophy". It was only to be changed, or transformed, with the emergence of a new, perhaps more decisive representation, in the mid-1960s, largely triggered by such philosophers as the aforementioned Michael Dummett (Dummett 1973, 1978, 1991). It was crucial for the analytic movement at large to be united by a single program, or under a common banner; and this eventually came to happen via the conception (which became the movement's development model) of a theory of meaning with no metaphysical presuppositions or implications — a theory like the one that had been developed by Quine himself in works such as those published in Quine (1953) and Quine (1969). This conception called for a revision of the very development of the analytic movement from its beginnings: the issue of foundations (historical and philosophical) becomes key; and the philosophy of Frege, who, together with

Russell and Wittgenstein was, after all, one of the founders of this movement, comes into the picture most decisively, while, as was argued by the "Oxford school" in the 1950s, Russell's philosophy was epistemologically and metaphysically contaminated and therefore could in no way serve as a model. According to this new representation, analytic philosophy had its foundations in what may be termed a "proto-history", i.e., in a more or less ideal period which potentially, or in embryonic form, contained the development of this philosophy to date (cf. Dummett 1991a). An entire new historiography (that is, a new way of reading this development) followed, starting with the one on Frege (by M. Dummett), Wittgenstein (by Peter Hacker (1996)), or Russell (by Nicholas Griffin (1991)).—Let us call this latter, more recent representation "the second stage in the meta-history of analytic philosophy". I am not sure whether this stage is now over; however, I daresay that there exists a third one. Starting in the 1980s, analytic philosophy began to think of itself as a broad philosophical tradition in the ambit of the history of western philosophy. As Martinich & Sosa (eds.) (2001) show, it agreed to include some of the philosophers who had previously contested fundamental analytic theses, as was the case of Richard Rorty (as concerns the analytic vs continental distinction) or Karl Popper (as concerns the role of epistemology and the very concept of philosophical analysis). This third stage thus seems to correspond to a period of profound reformulation of its own meta-history. As a conclusion, I question this reformulation and put forward some fundamental suggestions.

Cognitive Disorientation in Self-Deceptive Inquiry

Dion Scott-Kakures (Scripps College, Claremont)

15.09, 10:10 – 10:45, Room: 2.13

Self-deceivers (at least on a deflationary conception of self-deception) and inquirers in general, have something in common: they engage in reflective inquiry with the aim of settling a question. This is denied by some (e.g. intentionalists) who claim that a self-deceiver's behavior is at odds with this characterization. There's no doubt that many characteristics of a self-deceiver's cognitive behavior make plausible the claim that the self-deceiver and the typical inquirer have different aims: her persistent efforts to confirm favored hypotheses; her apparent asymmetric acceptance and rejection thresholds for doxastic embrace and rejection of various propositions; her alternately strikingly credulous and strikingly critical stances – credulous with respect to p- friendly evidence but resistant with respect to p-undermining evidence.

Here I aim to explain these distinctive features of self-deceptive inquiry and, in particular, the apparent control self-deceivers exercise over their belief forming-processes by holding that self-deception is a form of "cognitive disorientation." In cognitive disorientation, data or experiences upon which we typically rely to carry out inquiry come to mislead us systematically. In short, when we engage in self-deceptive inquiry, we systemically undermine or interfere with – by the very reflective efforts we engage in – our success in achieving our reflective aim.

Drawing on an analogy with spatial disorientation, I characterize two elements of the cognitive disorientation that I aim to locate in self-deceptive inquiry. With Confusion of Aim, an agent's intended, reflective aim has been confused with or unwittingly supplanted by another end; the agent takes herself to be pursuing one goal, when, in fact, she is intentionally acting in ways that unwittingly contribute to the securing of some contrary end, thereby, thwarting the realization of her reflective aim. Those who are (spatially or cognitively) disoriented are generally subject to Confusion of Aim. While there are no doubt, many ways of entering a state of Confusion of Aim, there's a particularly

pernicious route in which the agent receives misleading feedback during her activity, feedback she uses to adjust her behavior in light of and in pursuit of her reflectively held aim and that systematically misleads her. Call this Misleading Feedback. I argue that only when Confusion of Aim is the result of Misleading Feedback do we have the makings of cognitive disorientation.

I turn briefly to a consideration of the task of settling a question with an eye towards displaying the role and place of Confusion of Aim and Misleading Feedback in self-deceptive inquiry. When we try to settle a question of the form "p or not-p?" we presume that when and if inquiry is successfully completed, we will have:

a. Come to believe that p, if by our then current lights we have come to possess good and sufficient reason to believe that p; or
b. Come to believe that not-p, if by our then current lights we have come to possess good or sufficient reason to believe that not-p.

In this way, our aim or target in settling a question (of this form) is disjunctive; that is, *one's aim is to come to have either of one of two beliefs depending upon what, via inquiry, one discovers there is good or sufficient reason to believe.*

This, I argue, may suggest spare account of the self-deceiver's cognitive disorientation: She asks and persists in asking directional questions. One aim (the disjunctive aim of settling a question), has been replaced by another – coming to believe that p. As with a confirmation or myside bias, the inquirer begins with the question "Is it the case that p?" She fixes her attention on the search for evidence in favor of p; she spends little of no effort in the task of seeking evidence supportive of not-p. Having discovered a modicum of evidence in favor of p, she may come to believe that p. Might this be a model for the role Confusion of Aim plays in the cognitive disorientation of the self-deceiver?

Such an account will not succeed. While the account explains how it is that an inquirer comes to possess a biased body of evidence, it fails to explain how an inquirer's use of and response to the considerations so generated during her inquiry is, itself, typically disorientate d. A self-deceiver finds certain considerations compelling or probative – considerations a non-self-deceptive inquirer would not so regard. A self-deceiver's alternate credulousness and resistance is not explained. Finally, a self-deceiver's being moved by a limited amount of evidence to embrace a favored belief while requiring great amounts of evidence to embrace a disfavored belief is not explained. In sum, a self-deceiver doesn't merely discover any reasons for p and against not-p, she accepts these, she is moved by them: she regards these as having subjective normative force.

For this, we require an account of the Misleading Feedback the self-deceptive inquirer receives during inquiry. Here I devote attention to the development of such an account by appeal to the theory of cognitive dissonance. Crucial to the account is that a self-deceiver is misled by a signal that mimics one by which the typical inquirer steers aptly. In brief, the story is this: in self- deceptive inquiry, attitudes that are, from an evidential perspective, irrelevant to the question the subject is trying to settle are aroused and made accessible during inquiry. These attitudes are many, important, and resistant to change. These attitudes are dissonant both with the original worrisome datum and with the inquirer's uncertainty. In reducing the dissonance provoked by the presence of these, now salient, evidentially irrelevant attitudes, the self-deceiver's inquiry is subject to motivational bias in the form of misleading feedback.

I end by considering a number of objections to my account and by offering (the beginnings of) a rejoinder to a challenge to deflationary accounts of self-deception posed by an increasingly influential and noteworthy competing family of accounts of the phenomenon of self-deception.

Two Ways to Think about (Implicit) Structure

Georg Schiemer (University of Vienna)

15.09, 9:00 – 9:35, Room: 5.2

According to a dominant view in modern philosophy of mathematics, mathematics can be understood as the study of abstract structures (e.g. [2], [8]). Put differently, structuralism holds that theories of pure mathematics (such as Peano arithmetic, lattice theory, topology, or graph theory) study only the structure or the structural properties of their respective subject fields (namely number systems, lattices, topological spaces, and graphs). But what precisely is the relevant structure of such mathematical entities? How can we think about their structural content? The present talk will address these questions by drawing a distinction between *primitive* and *implicit* structure, that is, between mathematical systems defined by mathematical theories and their implicit structural content.

It is a well-known fact that mathematical objects such as rings, and topological spaces are usually introduced axiomatically today, that is, in terms of formal axiomatic definitions that specify the constitutive properties of the objects in question. For instance, a topological space can be defined in terms so-called neighborhood axioms, first introduced by Felix Hausdorff in 1914, which specify the properties of a neighborhood relation between points and sets of point sets. Similarly, a ring forms an algebraic structure consisting of a set and two binary operations that is defined in terms of the ring axioms. While it is clear that mathematical systems and their primitive properties are specified axiomatically in this sense, less has been said about how mathematicians investigate the implicit structure of these systems. Which methods are employed in modern mathematics to yield information about the implicit structure of a given system? In this talk, we will compare two general ways to think about this implicit structural content of theories of pure mathematics. According to the first approach, the implicit structure or the structural properties of mathematical objects are specified with reference to formal languages, usually based on some notion of definability. Thus, properties of systems such as rings or graphs will count as structural if they are logically definable in a formal language of the correct mathematical signature. According to the second approach, structures are determined in terms of invariance criteria. For instance, the structural properties of a given mathematical system are often said to be those properties invariant under certain transformations of the system or under mappings between similar systems (see [1], [6]). In the talk, we will investigate these two approaches to *implicit* structure by drawing to a particular mathematical case study, namely the study of simple incidence structures in *finite* geometry (compare [4]).

Given this geometrical example, we give a more philosophical analysis of the conceptual differences between the two methods. The talk will focus on three issues. The first concerns the conceptual motivation for treating mathematical structures in terms of the notions of definability and invariance. Why are these criteria adequate means for the specification of structural properties? As we will argue, both methods capture some form of “topic neutrality” underlying the structuralist account of mathematics. In the case of invariance, this is due to the fact that mathematics is indifferent to the intrinsic nature of mathematical objects and thus also indifferent to arbitrary switchings of such objects in a given system. In the case of definability-based approaches, the relevant topic

neutrality is related to the fact that adequate logical definitions should be reducible to statements about the primitive mathematical structure. The structural topic neutrality in mathematics can thus be explained in terms of the formality of logic and the fact that logic has not subject matter of its own (see [3]).

The second point addressed in the talk concerns the logical relation between two approaches to implicit structure. As has been argued in recent work on logic and model theory, the exist a general *symmetry* or *duality* between the method of specifying invariants relative to transformations or mappings and the notion of definability (see, in particular, [5]). We will present a formal account of this duality in terms of a so-called Galois connection between automorphism classes and Galois-closed sets of relations. More specifically, given this framework, it will be shown that the class of definable properties (of the objects) of a given primitive structure always forms a subclass of those properties invariant under the automorphisms of the structure.

Finally, we discuss the relevance of the two ways to think about implicit structure for our present understanding of mathematical structuralism. Here, in particular, our focus will be on the notion of the *equivalence* of mathematical structures (see [7], [8]). Building on the existing literature on the topic, we will propose two notions of structural equivalence that take into account not only the (axiomatically defined) primitive structure, but also its tacit structural content. The first notion is motivated by the idea of definable structure and based on the notion of *interpretability*. According to it, two mathematical structures are equivalent if they are bi-interpretable. The second notion, in turn, is motivated by the invariant approach and based on the notion of "transfers" between structures. According to it, two structures are equivalent if there exists a mapping between their domains that induces an isomorphism of the respective automorphism groups (compare [4]).

Davidson on Metaphors

João Esteves da Silva (Universidade de Lisboa)

13.09, 16:45 – 17:20, Room: Mattos Romão

In "What Metaphors Mean" (1978), Donald Davidson argued that metaphors *mean* what their words, in their most literal interpretation, mean and nothing more. From this follows that notions such as metaphorical meaning or metaphorical truth can do little or nothing in what concerns the explanation of how this kind of utterances work. According to Davidson, what distinguishes metaphors from non-metaphorical utterances is rather *use*, and the purpose of the seemingly special use bestowed on language in such cases – usually false (or, better said, with Stanley Cavell, "wildly false") assertions that nevertheless play a particular role – cannot be to *say* something special, which entails that the very idea that metaphors carry messages is also to be discarded.

In light of this, most theories of metaphor, which are commonly grounded on presuppositions contrary to these, while believing to provide a method for deciphering an encoded content, are in fact focused on the *effects* that metaphors have on us, and thus fall into the trap of taking the thoughts springing from these effects as content of the metaphor itself. Nonetheless, Davidson recognizes that metaphor is ultimately an indispensable device and one of genuine cognitive potential. Metaphors can help to notice details that could otherwise be ignored, bring certain analogies and similarities into focus, offer a sort of lens that filters relevant aspects. Little is said, however, regarding *how* a metaphor is related to that which it makes us notice.

With this essay, Davidson undertakes a true job of cleaning a particularly rough ground which

often lends itself to various sorts of confusion. But has this been enough to free us from illusion while looking at such linguistic phenomenon that, as Nelson Goodman would say, seems to be as important as it is odd? Have the reasons for this importance been made clear and this oddness dissolved? And, moreover, what could be the limits of possible constructive projects concerning metaphors built on this davidsonian foundation? We will attempt to answer such questions by attending to the following points.

§1. The main point will be to understand "What Metaphors Mean". Its main argument will be succinctly presented and contrasted with the positions of Max Black ("Metaphor") and William Empson (*The Structure of Complex Words*). Ultimately, Davidson's position can be seen as an inversion of the empsonian idea of the pregnancy of metaphors, i.e. it is rather our imagination that is pregnant, especially when stimulated by an effective metaphor. We will discuss whether this perspective constitutes a radical case of non-cognitivism.

§2. We will proceed by considering some objections that were directed against it by Goodman in "Metaphor as Moonlighting". It will be argued, with Davidson, that the idea of metaphorical truth or falsity defended by Goodman cannot hold and that it would be more reasonable to simply speak of useful or effective metaphors or of useless or ineffective ones. It will also be argued that the two main inconsistencies pointed out by Goodman in Davidson's essay are but apparent: on the one hand, when the role of an utterance changes definitively (as is the case of a metaphor that, so to speak, dies), such utterance effectively becomes a *different* one (despite the fact that its signs remain the same); on the other, on no occasion whatsoever Davidson claims that any role is exclusive to metaphors.

§3. The main point would be to evaluate some of the consequences of "A Nice Derangement of Epitaphs" (1986) on the thesis of "What Metaphors Mean". It will be argued that, despite some necessary adjustments (that have above all to do with the rejection of the notion of literal meaning in favour of that of *first meaning*), the two essays are not, as might be supposed, incompatible. The coining of the notion of first meaning is in fact a result of the previous essay's attack on the distinction between literal and figurative meaning. Moreover, the conclusion of "A Nice Derangement" – the collapse of a conception of language as "a clearly defined shared structure that users acquire and then apply to cases", which in its turn entails a collapse of any supposed boundary between the philosophy of language and the philosophy of action – contributes decisively to the understanding that a merely semantical explanation of metaphors is inadequate.

§4. Finally, it will be argued that the often regretted fact that Davidson says little or nothing about the relation between metaphors and its respective targets (although he still leads us to suppose that it proceeds through a more or less spontaneous recognition – in a sense similar to the recognition of e.g. the sadness manifested by a person's facial expression – and not by way of inference) is justifiable: its explanation always depends on the attention given to the *particular circumstances* in which a certain metaphor is applied, due to the fact that in practice the variety of such applications and all its possible variations is immense and constantly updatable. By taking this into consideration, instead of a theory that seeks to account for such diversity (a project that would seem to us unfair), Davidson develops what would could be called a minimal theory, which can already be seen in itself as a constructive project, offering a sufficiently solid ground from which each metaphor can be properly recognized and interpreted. As a complement, two additional notions that can help to bring light into this matter will be suggested: *showing* (or elucidation) and *transition* (metaphors work as transitional devices – a sort of linguistic ladders – in order to show certain things or aspects), that we will attempt to clarify.

Aggregates and Their Role in the Problem of Coincident Objects

Jeroen Smid (University of Manchester / Lund University)

13.09, 16:45 – 17:20, Room: 5.2

The problem of coincident objects concerns the relation between distinct composite objects that (seemingly) coincide. Take, for example, Theseus's ship. This composite object can stay the same while changing some of its planks. Besides the ship, the planks seem to compose another object, viz. the aggregate of the ship's planks. Contrary to the ship, the aggregate of the ship's planks cannot survive the replacement of any of the planks but has all of the planks essentially. The question then arises: if the ship completely coincides with its aggregate of planks while they are numerically distinct, then how are they related?

One solution often proposed to the problem is that the two objects stand in a relation of constitution. This view is called 'The Standard View' due to its popularity. (Defenders include Lynne Rudder Baker (1997), Frederick Doepke (1997); Mark Johnston (1992); Peter Simons (1987, 237ff); Judith Jarvis Thomson (1983); and David Wiggins (1968).) But it is not clear how the problem is really solved by introducing the novel relation of constitution (Olson 2001; Smid 2017). A more revisionary approach holds that objects are composed of both spatial and temporal parts. These perdurantistic theories solve the problem of coincident objects by holding that the ship, say, and its aggregate of planks may temporally overlap by sharing a temporal part. (Defenders include Katherine Hawley (2001); Mark Heller (1990); David Lewis (1983); Ted Sider (2001); and Achille Varzi (2003).) Although I have nothing to say against such theories, it would be better to have a solution for the problem of coincident objects that is neutral with respect to the question how objects persist through time. My talk shows that such a solution is available.

The problem of coincident objects crucially depends on the distinction between a real whole or unity and the mere aggregate of its parts. Aggregates go by many names, besides 'aggregate' it is called 'a totality', 'mere sum', 'collection', or 'collective'. The many names may be due to its long heritage because the notion already appears in *Metaphysica Zeta* 17 when Aristotle distinguishes a syllable from the aggregate of its letters. Their existence is not often disputed for they seem to be given almost "automatically": if there are some objects then it seems that one also has an aggregate consisting of exactly those objects.

However, I think we do not have good reasons to postulate the existence of aggregates. To show that eliminativism with respect to aggregates is preferable, I first explicate, on the basis of the debate on the problem of coincident objects, three existence and identity criteria of aggregates. These criteria are as follows:

[Extensionality] *Aggregates are the same if they have the same constituents.*
[Essentiality] *Aggregates cannot gain or lose constituents.*
[Comprehension] *If something is ϕ, then there is an aggregate that consists of all and only everything that is ϕ.*

I compare these with the existence and identity criteria for sets and mereological fusions, showing that aggregates are neither sets nor fusions. Aggregates are not sets because set-formation is not idempotent whereas aggregate-formation is. (Roughly, an operation is idempotent if applying it n+1 times gives the same result as applying it n times.) Neither can aggregates be identified with mereological fusions – *pace* Thomson (1983) – due to what I call the 'Too Many Parts problem'.

Since a fusion overlaps with more things (and thus has more parts) than the constituent of the aggregate with which it is identified, there are cases where we intuitively have distinct aggregates but no distinct fusions. So, if there are aggregates, then they are distinct from both sets and fusions.

We should thus either accept aggregates as *bona fide* entities irreducible to anything else, or instead look for suitable paraphrases so that we can eliminate them from our theory of the world. I explain how plural quantification provides a satisfactory paraphrase because plural quantification tells us that plural *terms* behave very similar to how aggregates are thought to behave. Moreover, by paraphrasing all alleged reference to aggregates into plural quantification we also dissolve the problem of coincident objects. Because instead of having to look for a one-one relation holding between two individual objects (a whole and an aggregate) that coincide, we now have to find a one-many relation holding between an individual object (a whole) and some objects (denoted by a plural term). The most obvious candidate is composition: the whole is *composed of* some objects (i.e., it is composed of its parts). I end my talk by investigating how this solution may (fail to) work in the case of the relation between an object and its matter or stuff – such as the relation between a statue and its marble.

Truth-Conditions and Conceptualization:
Where Formal Semantics Meets Human Cognition

Francesco Spada (University of Modena and Reggio Emilia)

14.09, 14:30 – 15:05, Room: 5.2

At least since the publication of Grice's *William James Lectures*, the debate about the role played by context in the determination of the proposition expressed by an utterance has been characterized by the emergence of a number of distinct positions on the subject-matter: semantic minimalism, indexicalism, moderate contextualism, radical contextualism, semantic relativism, etc.. In this paper I do not set out to defend any of the positions before mentioned. Rather, I would like to enhance the debate by introducing a new position that I call *conceptualism*. Conceptualism is the view that the truth-conditions of a sentence, as well as whether a sentence has truth-conditions, in some cases may depend on how we conceptualize the object(s) the sentence is about. I will show how conceptualism applies to two cases that have been frequently discussed in the semantics/pragmatics literature: colour predications (e.g. 'S is red') and numerically quantified sentences (e.g. 'two bottles are on the table). My claim will be that there is a common disposition underlying the understanding of both kinds of sentences, namely our capacity of representing the very same objects as simple and unstructured entities or as complex and structured unities. This may concern the extension as well as the quantity of such objects. With respect to its extension, a surface may be conceptualized as one simple surface or it may be mentally divided into two or more surfaces and eventually conceptualized as a whole made out of those simpler parts. With respect to their quantity, B_1, B_2 and B_3 may be conceptualized as three bottles or as two bottles plus one (among other alternatives): In the latter case B1, B_2 and B_3 are represented as the composition of two and one bottle; in the former case, such an internal composition is lost. I will argue that this underlying disposition affects the determination of the truth-conditional content of colour predications and numerically quantified sentences so as to justify the introduction of the view I call conceptualism. Then, I will discuss the impact of conceptualism on the endurance of traditional formal semantics. My goal is to argue that the specification of the truth- conditions of a sentence is an eminently cognitive process that a

subject performs, a process that may be affected by the cognitive mechanisms whereby the subject conceptualizes the objects in the outer world, hence a process that must be set in its own cognitive domain for it to be properly understood. If I am right in these observations, this implies that we should revise the role and ambitions of the project of formal semantics.

'What is the case – a fact – is the existence of states of affairs' (*Tractatus*, 2)

Shunichi Takagi (University College London)

13.09, 17:55 – 18:30, Room: Mattos Romão

The aim of this paper is to correct a reading of the *Tractatus* which in my view mistakenly attributes to it an erroneous conception of fact. For this purpose, I shall focus on the contemporary debates between correspondence theories and identity theories of truth. The Tractarian conception of the world as the totality of facts has been resurrected in this field of contemporary philosophy, mainly through *Mind and World* of John McDowell, who is a self-acknowledged identity theorist. According to McDowell, as the world is the totality of facts, there simply is no ontological gap between what one can think and what can be the case, and that accordingly a fact is simply *identical* with a true thought. This is understood to go against correspondence theories of truth because according to identity theorists there is no slightest gap to be bridged by correspondence between what is the case and the content of the thought.

Despite a prominent role it plays in these debates, however, opinions diverge over such a basic issue whether the *Tractatus* itself endorsed a version of identity theory or a kind of correspondence theory. Julian Dodd, whose *An Identity Theory of Truth* is a major work in this field of studies, seems to be among those who believe that Wittgenstein held a correspondence theory of truth and thereby presented a wrong conception of fact.

Was the author of the *Tractatus* committed to a mistaken view on the nature of fact? As a fundamental concern of the identity theorists is the nature of fact, the question we need to examine is what kind of conception of fact Wittgenstein presented in the *Tractatus*.

What is the mistaken conception of fact that correspondence theorists allegedly endorse? And in any event what is a correspondence theory of truth? Dodd specifies it to be a view which endorses the truth maker principle:

> For a proposition *p* to be true, there must exist at least one entity, distinct from *p*, whose existence entails that *p* is true.

Based on this, Dodd regards the correspondence theorists as taking facts to be states of affairs, by which he means:

> i. A fact is a complex of particulars and universals
> ii. Since a fact is a complex, it has its constituents as its parts
> iii. As the constituents of a fact are in the realm of reference, a fact which is a complex of these items is also in the realm of reference.

On the face of it, there seems to be more than a sufficient amount of textual evidence to show that Wittgenstein was a correspondence theorist. For example, does the *Tractatus* not say that 'If an

elementary proposition is true, the state of affairs exists' (TLP, 4.25), where an elementary proposition and the state of affairs it depicts should be distinct? If so, this is nothing but the truth-maker principle we have just seen, and henceforth it is indisputable that the early Wittgenstein was a correspondence theorist of truth.

However, the issue seems to be not so simple, and I shall argue that the *Tractatus* is not committed to any of the three errors Dodd condemns.

In order to dispel the first point, it suffices to quickly review Ramsey's 'Universals', in which he presents the author of the *Tractatus* as holding a different view from one which says that a fact is a unity of a particular and a universal. By way of this, I shall suggest that the mistaken ascription of the first error to the *Tractatus* is rooted in not distinguishing between the *Tractatus* itself and the position Wittgenstein held in *Notes on Logic*, which is essentially the view Russell presented in his lectures on Logical Atomism. At the same time, I shall also point out that the Tractarian view is alien to the trope theory.

Regarding the second and third point, I shall counter these worries simultaneously since they are interconnected. For this objective, I shall address one of the major exegetical points of the *Tractatus*, namely whether a *Sachverhalt* is a kind of *fact* as the Ogden-Ramsey translation 'atomic fact' might imply, or it itself is *not* a fact at all and we should follow the Pears-McGuinness translation 'state of affairs'. To answer this question, I shall suggest that an important question is whether 'existence' in Wittgenstein's explanation of the nature of fact, i.e. 'existence of a state of affairs', makes any contribution to characterising it, and if it does, what it is. For this purpose, I shall carefully examine some historical texts, including Frege's letters to Wittgenstein, Ramsey's *Nachlass* and his 'Facts and Propositions', and thereby show that it *does* make a significant contribution to what a fact is and part of what it does is to distinguish fact and complex.

Based on this, I shall refute the second and the third points. That is, regarding the second worry, I shall argue that the *Tractatus* does not present an erroneous conception of fact which entails that the fact that the sun is rising has as its part the actual object which is emitting rays. And, with respect to the third point, I shall suggest that if we properly understand the nature of fact according to the *Tractatus*, we shall clearly see that the world as the totality of facts in it is as finely grained as the intensional contents of the Fregean realm of sense.

In this way I shall argue that the Tractatus is innocent of any of the three wrong commitments regarding which Dodd criticises correspondence theorists. To conclude this paper, I shall quickly suggest that to decide fully whether Wittgenstein in the Tractatus was a correspondence theorist or not, we need to attend to some of his *Nachlss* in which Wittgenstein left some remarks on the nature of pictoriality.

Epistemological Dogmatism and the Problem of the Criterion

Guido Tana (University of Edinburgh / Universität Leipzig)

14.09, 15:05 – 15:40, Room: Sala de Actos

Dogmatism is an internalist stance on warrant for perceptual knowledge most recently defended by James Pryor (2000, 2004), Micheal Huemer (2001, 2007, with the name of Phenomenal Conservatism), Brit Brogaard (2013), and Elijah Chudnoff (2011). One of its main features lies in offering a refutation of external world scepticism: Dogmatism argues that experience has a distinctive presentational phenomenology which is able to provide immediate defeasible justification for beliefs about

the external world (Pryor 2004:357, Chudnoff 2011:314). The goal of this essay is to argue that Dogmatism's anti-sceptical credentials are fundamentally undermined by its having no resources to solve the Problem of the Criterion, exemplified by the Easy Knowledge objection. This problem is a fundamental one for Dogmatism, because its anti-sceptical stance consists in delivering an explanation of how defeasible warrant is possible in the first place, something which is not possible if this justificatory scepticism is left unscathed.

Dogmatism argues that the immediate justification experience provides does not depend on any background beliefs (Pryor 2005:204). If it seems to a subject that p, then, in absence of relevant defeaters, the subject has some degree of justification for believing that p (Huemer 2007:30). Experiences are not in need of further justification like beliefs are, and we only need to heed our ordinary doxastic practice (2007:54, Pryor 2005:210). Appearances are presumed true until proven false, because it is prima facie reasonable to believe that they are (Huemer 2001:100-3). The conception of scepticism that this stance tries to avoid is an essentially scenario-based one, in which it is considered epistemically inappropriate to bestow reasonableness on the sceptic's distrust of the senses, provided that no empirical reasons underlie this kind of doubt. Furthermore, Dogmatism, while being an internalist proposal, substantially rejects the KK thesis applied to justified beliefs: "What makes you justified in believing p is one thing; what makes you justified in believing that you have justification for believing that p is something else" (Pryor 2000:535).

This perspective is a fundamentally *neo-Moorean* proposal. It suffers however of the same lack of dialectical bite against the external-world sceptic as G.E. Moore's own answer. Its particularist and experientialist methodology rules out scepticism from the beginning (Fumerton 2008:42), labelling it "a[n] [epistemological] disease" (Pryor 2004:368). The two opponents appear to speak one past the other on what is required for justification and knowledge, because Dogmatism rejects entirely the dialectic on which scepticism grounds its perceived strength. However, even granting that external world scepticism cannot be based on any adequate piece of evidence does not grant Dogmatism the means to steer us in clear waters. This is because some elements of its reasoning can be legitimately disagreed upon: scepticism needs not to be based only on deceptive scenarios, and the belief that perception has a determinate content or phenomenological presentation that is able on its own to warrant prima facie justification to our beliefs is not out of contention (Neta 2004:204-5).Ultimately, if Dogmatism engages with scepticism on an epistemic level that does not concern scepticism, its contribution risks to be something marginal or trivial (Wright 2007:47).

However, notwithstanding its irrelevance to traditional Cartesian worries, Dogmatism can be reassessed if conceived as a claim concerning the possibility of having any justification in general, which Dogmatism localizes in experience itself (Coliva 2008:241-42). Dogmatism would therefore have its role in answering the normativist justificatory doubt which undermines knowledge assertions. Its proposal would be one of redeeming our claims to justification by offering an intuitive and natural variety of justificatory warrant.

This will be shown unfortunately to be a false hope. In delivering an empirical answer to scepticism concerning warrant about empirical justification in general, Dogmatism shows itself fatally liable to the *Problem of the Criterion* and the sceptical outcome lying at its core: *Agrippa's Trilemma* (Williams 2001:63).

The objection that sets this upshot is the *Easy Knowledge* argument popularized by Stewart Cohen and Jonathan Vogel. It states that theories such as Dogmatism reject the KR principle: "a potential source K can yield knowledge for S only if S knows K to be reliable" (Cohen 2002:309). By doing this, Dogmatism allows for *Basic Knowledge* in its justificatory structure. However, this step

appears to licence a way of obtaining knowledge that is intuitively too easy to be epistemically satisfactory, independently from any particular sceptical scenario (2002:310). *Easy Knowledge* casts a shadow on the entailment from "the table appears red" to "the table is red", making us wonder whether we had basic knowledge in the first place. The Dogmatist might try to answer this sceptical possibility by avoiding the dialectic again, arguing that evidence is required to doubt the entailment. However, on this epistemic level regarding justification, the issue cannot be escaped; the aim of settling whether a particular piece of justification is true (Davies 2004:240) appears to be a reasonable one if we are interested in redeeming our claims to epistemic warrant (Cohen 2005:427).

This is because, by adopting a particularist methodology, Dogmatism is providing an answer to the Problem of the Criterion, which establishes the mutual epistemic dependence of the two questions "What do we know?" and "how are we to decide whether we know?" (Chisholm 1989:6). Dogmatism chooses a specific criterion, fully partaking in the sceptical dialectic.

However, this answer to the Problem of the Criterion leads directly to Pyrrhonian scepticism under the guise of Agrippa's Trilemma due to the legitimacy of rational disagreement on the choice of the criterion. The Pyrrhonian dialectic concerns "having an epistemically significant reason for a belief" (Cling 2009:334). About this matter a rational disagreement on which criterion to choose is something the Dogmatist cannot evade, as "doubts might be raised because doubts have been raised" (Barnes 1990:26-7, 116). Dogmatism commits itself to the Assumption trope of the Trilemma, by subtracting the epistemic role of perception from rational evaluation, disqualifying it from achieving the goal of redeeming our claims to justificatory warrant, as a bare assertion has no rational standing when a plausible rational disagreement is raised.

Finally, even if more leeway is granted to Dogmatism, its methodology can be identified as circular, falling within the second trope of Agrippa's Trilemma, Reciprocity. The arbitrary decision for a criterion which permits Easy Knowledge allows for what has been called *Cognitive Penetration* (Siegel 2012, Markie 2005, 2006). Through the arbitrary recourse to Basic Knowledge, Dogmatism is guilty of elevating epistemically inappropriate beliefs to the status of justified, epistemically appropriate beliefs, with the consequence that appearances in general cannot warrant any justification anymore (Moretti 2015:306). Dogmatism permits to the criterion of prima facie defeasible justification based on phenomenological presentation of appearances to be grounded on itself, elevating to appropriateness whatever belief we have on the basis of our having it. This unfavourable outcome could not be supported by any rational justificatory theory, exhibiting Dogmatism's inadequacy in both answering scepticism as well as in providing a warranted epistemic criterion for justification.

Inheritance Reasoning From an Ecological Perspective

Paul D. Thorn (HHU Düsseldorf)
Joint work with Gerhard Schurz (HHU Düsseldorf)
15.09, 9:35 – 10:10, Room: 5.2

In executing an inheritance inference, one reasons from a premise stating that a given property is 'typical' among a class of individuals, and concludes that the property is typical among a subclass of the class:

Property ϕ is typical among the members of class C.
SC is a subclass of C.

ϕ is typical among SC.

In our talk, we present results from a simulation-based study of inheritance inference. The study endeavours to ascertain which kinds of inheritances inferences are reliable, with attention to variations in reliability that are contingent upon the type of environment in which inferences are made. Within the study, reliability is measured by the tendency of inheritance inferences to deliver true conclusions (given true premises), where a property is said to be "typical" of a class or subclass just in case the relative frequency of the property among the class/subclass exceeds some bound *r*. We evaluate the reliability of inheritance inference for varied *r*.

One goal of the study is to address whether inheritance inference is reliable in the case of 'exceptional subclasses'. In cases where the following propositions are true of a tuple $\langle SC, C, \psi \rangle$, we say that SC is an exceptional subclass of C, with respect to ψ:

ψ is typical among C.
SC is a subclass of C.
ψ is not typical among SC.

Another goal of the study is to illustrate that the reliability of inheritance inference depends upon the criteria that are used in selecting the classes that serve as the basis for inheritance inference. One approach to inheritance inference proceeds by treating any atomic property as determining an admissible class. A second approach identifies classes with the cells of a partition (of size k) of the set of objects that minimizes the dissimilarity between objects that are assigned to the same class. More precisely, we select some partition $\Pi = \{c_1\}$ of the relevant domain of objects, whose value is minimal for f:[27]

$f(\Pi) = \Sigma c_i,\ldots,c_k \in \Pi\ \Sigma\ x \in c_i \,||\, \phi(x) - \mu_i \,||^2$, where where μ is the mean value of the elements of c_i.

In addition to permitting more inferences, our study shows that the second approach results in inheritance inferences that are far more reliable (for all $k > 1$). The two approaches also produce different results in the case of exceptional subclasses: When paired with the first approach to selecting classes, inheritance inference in the case of exceptional subclasses is horrendously unreliable. With the second approach, exceptionality generally results in no more than a small decrease in reliability (for all $k > 1$).

The following figures illustrate some results of our simulations. Each bar represents mean values for 10,000 randomly generated environments. For all figures, *error rates* are equated with the relative frequency of cases *not* satisfying the 'conclusion condition' for an inheritance inference among the cases that satisfy the 'premise conditions'.

[27] We thus apply so called "k-means clustering", due to its simplicity, popularity, and satisfactory performance in trial simulations.

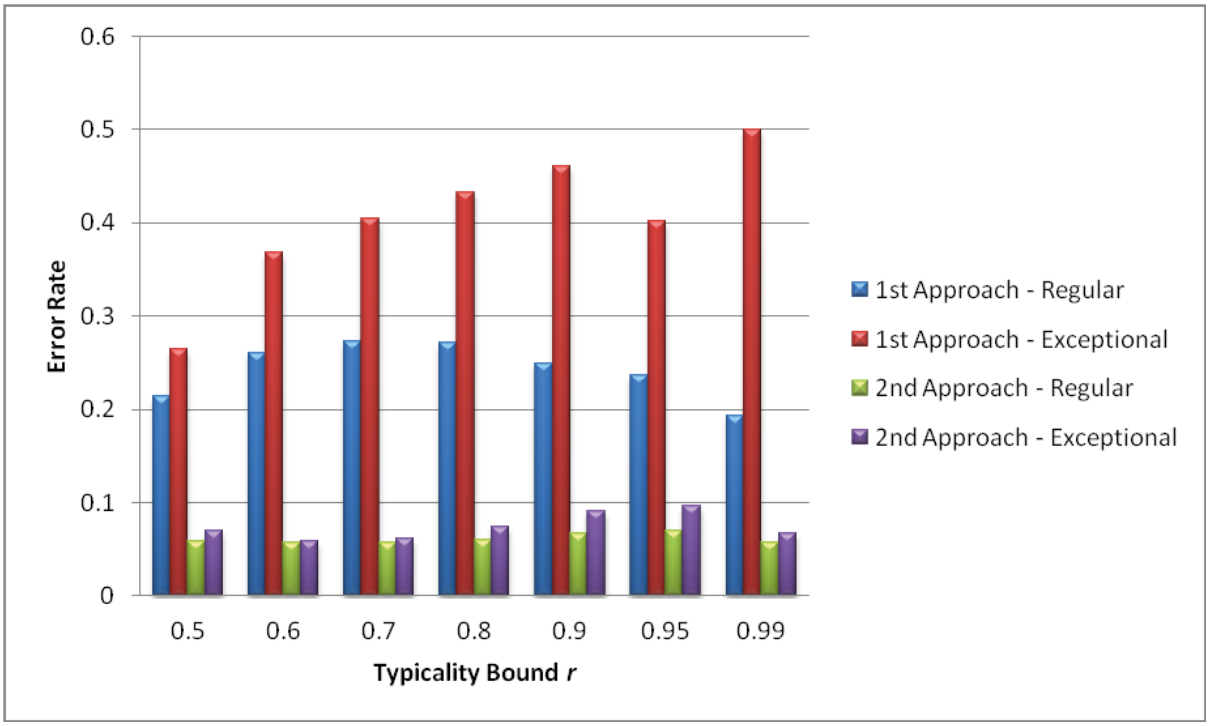


Figure 1: Error Rate as a Function of Typicality Bound r (with k = 8)

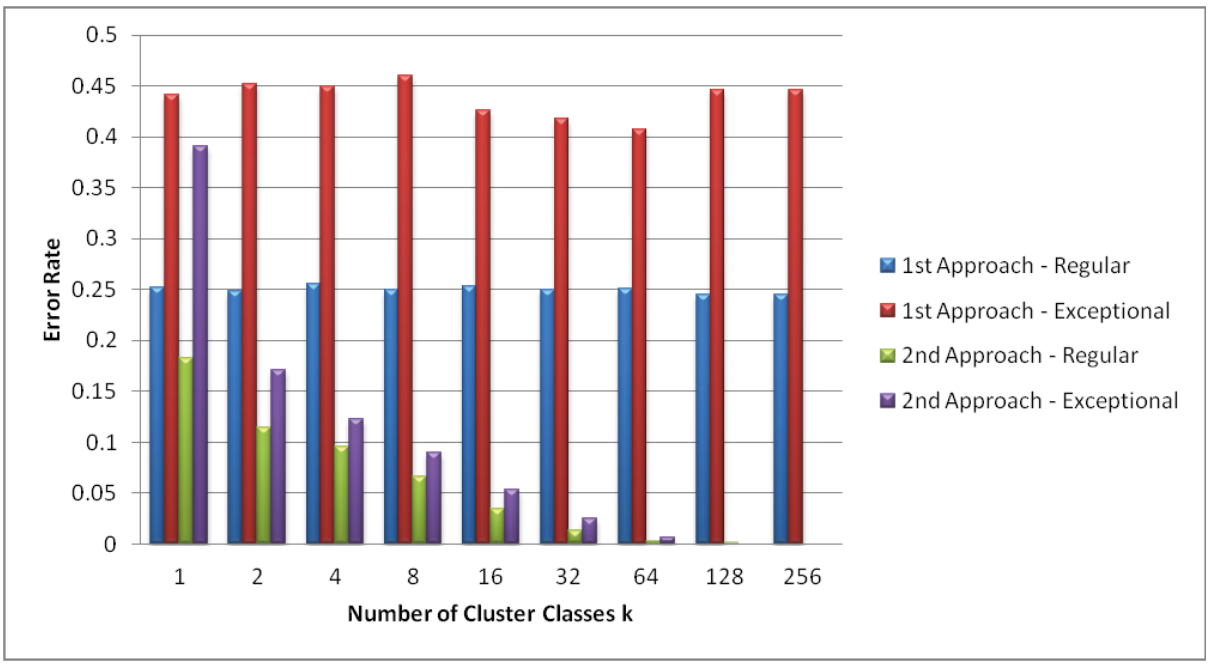


Figure 2: Error Rate as a Function of Number of Cluster Classes k (with r = 0.9)

Inbetween Good and Bad Luck: Responsibility-Sensitive Justice and Disadvantage

Manuel Valente (UCLouvain / Leiden University)

15.09, 9:00 – 9:35, Room: Mattos Romão

Over the past four decades, responsibility-sensitive justice has become central to contemporary political philosophy (Axelsen et al., 2018, p.4). Such conception of justice accepts inequalities that result from people´s choices and rejects (or compensates for) disadvantages individuals are not responsible for. As a result, it focuses mainly on disadvantage and aims at combatting the influence of bad luck in a distribution of resources. Yet, whereas many political philosophers agreed that the goal of distributive justice is to eliminate (involuntary) disadvantage, the question of involuntary advantage has not been addressed as a question in itself. Perhaps rightly, political philosophers have situated the acceptance of good luck in relation to bad luck. But does the commitment to extinguishing the consequences of bad luck imply rejecting advantages for which individuals are not responsible? How should this conception of justice situate both types of luck in relation to each other?

In this presentation I analyse the extent to which the goal of eliminating involuntary disadvantage also requires combatting involuntary advantage. I offer two interpretations of disadvantage and explain how they might, or not, entail the acceptance of good luck. My main argument is that if we take disadvantage comparatively, proceeds from good luck will almost certainly be regarded as unfair. If, on the other hand, disadvantage is understood non-comparatively, gains from good luck do not necessarily need to be corrected towards equality.

In order to make my argument more engaging, but also improve its clarity, I use a thought experiment that I will call `the helicopter drop´. This experiment starts by imagining a just society. In this society no individual is disadvantaged through no fault of their own and all inequalities reflect the responsibility of individuals for the choices they have made. One day, a helicopter flies over the land and drops large sums of money upon the citizens of that land. As a result of the natural fall of money, citizens will be unequally advantaged by the "drop". Assuming all individuals had an equal chance of moving to locations where the helicopter dropped the money, this event produces two consequences. First no one will be left worse off in absolute terms, which means that disadvantage can only emerge by comparison. Second it is presumed that no one will be responsible for the advantages they now enjoy. If the situation before the helicopter drop was just and everyone is now better off than before, should the new situation be rendered unjust?

I use this example as means of testing our intuitions about the new distribution that emerges after the helicopter drop. When looking at comparative disadvantage, i.e., the situation where one is worse off than another, the helicopter drop left some disadvantaged when it made some people better off than others. I argue that comparative disadvantage does not enable that advantage leaves nobody disadvantaged as the relative advantage of one will always entail the relative disadvantage of another. If the purpose of justice is to eliminate involuntary disadvantage, one´s advantage can only not be unjust when inequality is the responsibility of the disadvantaged. This, I will argue, entails a commitment to (almost certainly) rejecting gains from good luck, as much in the helicopter example as in real life.

When looking at non-comparative disadvantage, i.e., the situation where one is worse off than he would otherwise be, the helicopter drop has left nobody disadvantaged as no one is worse off than they would be if the drop had not occurred. Non-comparative disadvantage enables situations where the advantage of some disadvantages nobody (in absolute terms). This interpretation opens

the possibility of deciding how to deal with advantages from good luck since, according to this reading, the helicopter drop left no one disadvantaged. Does this make advantage from good luck no longer unjust? I suggest that the answer here depends on how we weight the values of efficiency and responsibility against each other.

Efficiency refers to the idea that it is good if some persons (at least one) is better off without other persons (at least one) being worse off in the process (Cohen, 2008, p.87-90). Giving more weight to efficiency, as opposed to responsibility, means accepting involuntary advantages when nobody is worse off than they would otherwise be. It supports the principle that gains from good luck are only acceptable when they leave nobody disadvantaged (Knight, 2009). In the helicopter drop, this would mean that the money dropped would be `up for grabs´ and not subject to further taxation.

Defenders of responsibility-sensitive conceptions of justice might, however, believe that, in such conflicting cases, responsibility should prevail over efficiency. This would require that advantages not only leave nobody disadvantaged, but also are used to improve the situation of the disadvantaged. Emphasising responsibility in this case would imply that individuals should be equal in respect to what they are not responsible for. I argue that opting for this alternative would support financing an equal basic income. That is, the helicopter drop would serve as means of boosting every one´s inheritance in this society. That is, in a world where involuntary disadvantage has been eliminated, such as the one where the helicopter had dropped the money, involuntary advantage should be to the equal advantage of everyone.

This presentation does not aim to defend one position above all others. Instead, it identifies two ways of conceiving disadvantage and their implications for assessing whether advantage is just. This point is important for three main reasons. First because it questions the sensitivity of inequality to responsibility, which helps clarifying possible ambiguities existing in the literature. Secondly, it sheds a new light over the famous distinction by Joel Feinberg (1974) between comparative and non- comparative justice. Thirdly, and perhaps its utopian goal, it asks if the justice sought by responsibility-sensitive justice has been achieved once involuntary disadvantages have been eliminated.

Works, Texts, and the Anti-Intentionalist Fallacy

Simon Walgenbach (University of Manchester)

14.09, 16:45 – 17:20, Room: Mattos Romão

One of the major points of contention in the analytic philosophy of literature is the relevance of authors' intentions to the interpretation of their works. As interpretation is widely construed as the attribution of meanings to works of literature, this comes down to the question to what extent such intentions determine these meanings. While anti-intentionalists take literary meaning to be independent of authorial intentions, intentionalists hold that they should play at least some role in literary interpretation. One of the most influential formulations of anti-intentionalism has been put forward by Monroe Beardsley (1970) who argues that *textual meaning* – i.e. what a text means – is independent from *authorial meaning* – i.e. what its author meant with it – and that literary interpretation should be concerned solely with the former. Consequently, he regards authorial intentions to be irrelevant for literary interpretation. I argue that Beardsley's considerations are unsound because he fails to recognise the difference between works of literature and the texts that make them up to the effect that even if his arguments about the meaning of texts were sound, this would not tell us anything about the interpretation of literary works.

To understand why Beardsley's theory of interpretation fails, it is necessary to reflect on the relation between two kinds of entities philosophers of literature usually postulate. On the one hand, there are *literary works* which are the primary objects of literary interpretation. On the other hand, these works are somehow constituted by syntactic-semantically defined and individuated *texts*. Bearing this distinction in mind, I discuss Beardsley's three arguments against the identity of textual and authorial meaning:

1. "There are textual meanings without authorial meanings. Therefore textual meaning is not identical to authorial meaning."[28]
2. "The meaning of a text can change after its author has died. But the author cannot change his meaning after he has died. Therefore, the textual meaning is not identical to the authorial meaning."[29]
3. "A text can have meanings that its author is not aware of. Therefore, it can have meanings that its author did not intend. Therefore, textual meaning is not identical to authorial meaning."[30]

I point out that Beardsley is inaccurate here: to establish his strong anti- intentionalism, he needs textual and authorial meaning to be independent of each other (and not just non-identical). I concede that his first argument seems to support even this stronger claim. However, I argue that all three arguments suffer from the same defect: while they may or may not be sound in establishing that texts mean something independently of authors' intentions, they fail to prove that the same holds for works. For the transition from one to the other, one must assume an additional principle that identifies the two meanings. Admittedly, Beardsley need not argue for such a principle as he explicitly commits himself to textualism, according to which works of literature are identical to the texts that constitute them, and he plausibly suggests that this implies the identity of text and work meaning. Apart from its being implied by textualism, however, Beardsley does not offer us an independent reason to accept that works draw their meaning from their texts alone. Hence, if textualism is false, Beardsley's arguments tell us nothing about the meaning of literary works.

I then argue that textualism is indeed false. I concede that a prominent argument put forward by Arthur Danto (1981) and Gregory Currie (1991), albeit compelling, cannot be used to refute textualism in the context of the debate about intentionalism. Since both demonstrate the non-identity of textually identical works with appeal to their different admissible interpretations, this argument cannot be used within an ontology that is supposed to settle debates about interpretation without begging the question. Instead, I present an argument that can be sketched as follows:

(B1.1) If textualism is true, then it is not possible that some but not other copies of one text are copies of a work;
(B1.2) however, it is possible that some but not other copies of one text are copies of a work;
(B1.3) therefore, textualism is false.

As the textualist is likely to deny (B1.2), I present a thought experiment to support my case. Imagine that Smith and Jones are pen pals. To celebrate the 10th anniversary of their friendship, Smith collects, typewrites, and binds their entire correspondence. Let us call the resulting book C*1* and say that it is produced at a time t*1*. Imagine further that at t*2* and oblivious of these events, Brown writes an epistolary novel N. Finally, imagine that N is constituted by a text T and that accidentally,

[28] Beardsley 1970: 19.
[29] *Ibid*.
[30] *Ibid.*: 20.

T consists of the exact same word sequence that appears in C*1*. The situation, I argue, is as follows:

(B2.1) If textualism is true, then every copy of N is a copy of T;
(B2.2) if every copy of N is a copy of T and N is a work, then every copy of T is a copy of a work;
(B2.3) N is a work;
(B2.4) C*1* is not a copy of a work;
(B2.5) C*1* is a copy of T;
(B2.6) therefore, textualism is false.

Hence, what we end up with is one text that constitutes a work of literature in some but not other cases. I consider possible textualist objections against (B2.3), (B2.4), and (B2.5), arguing that they all fail. The most plausible conclusion, I claim, is that textualism is false.

As Beardsley needs to assume textualism and textualism is false, Beardsley's arguments against intentionalism fail. Furthermore, it turns out that the interpretation of texts and the interpretation of works are two issues that need to be treated separately. The interesting question for the philosopher of literature will then be how a work's meaning is determined and how it relates to its text's meaning. It remains open whether or not authorial intentions are another determinant.

Truth-Tracking and Knowledge in Virtual Reality

Billy Wheeler (Sun Yat-Sen University)

13.09, 15:05 – 15:40, Room: Sala de Actos

Virtual Reality (VR) has been used successfully to train individuals in a diverse range of occupations, including pilots, surgeons and engineers, in the skills needed to carry out complex tasks. In the language of epistemology, much of what these individuals have learned could be characterized as 'knowledge-how' or skills-based knowledge. But with advances in VR technology and a lowering of its cost, VR is now making its way into the classroom, university, and other academic settings. But this raises the question: can 'knowledge-that' or factual knowledge about the real world also be gained through experiences in VR?

One reason for thinking it cannot comes from the history of philosophical skepticism. Hilary Putnam's (1981) famous 'brain in a vat' scenario shows that if we are in a simulated universe then we cannot take ourselves to know much of what we ordinarily believe. Nozick (1981) too agrees with Putnam: if we are indeed a brain in a vat, we would not know much because our beliefs inside a simulation cannot 'track the truth' of the real world. The programmer could design the program in any way they see fit, and it need not match the way the actual world is.

James McBain (2017) has recently argued that this shows truth-tracking is an inappropriate theory of knowledge for explaining how knowledge through VR is possible. In its place, he advocates an 'information-theoretic' account based around Fred Dretske's (1981) theory of knowledge as the 'flow of information'.

In this talk, I will show that contra McBain, a truth-tracking account of knowledge from VR is possible. I start by raising a problem for his account which shows that knowledge from VR is more than just the reliable flow of true information. What this criticism reveals, is that there are in fact two types of knowledge one can associate with VR. Firstly, there is *knowledge-in* VR, or virtual knowledge, which is a feature of correct belief about the content of the VR world. Secondly, there

is *knowledge-from* VR, which is a feature of correct belief about the actual world. I show how these two aspects can be connected by a 'double truth-tracking account'. Put simply, true belief in the virtual world amounts to knowledge if it tracks the content of the virtual world, and true belief about the actual world amounts to knowledge when the content of the virtual world tracks the truths of the actual world.

How the content of the virtual world tracks the actual world is explored through a number of options, but the most important is when there exists a method for reliably connecting the content of the virtual world to the actual world. This mechanism is usually played by the role of the computer programmer who satisfies the 'safety' and 'adherence' conditions Nozick places on his tracking conditions.

The argument relies upon transitivity of the tracking relation: (i) if beliefs track the virtual world and (ii) the content of the virtual world tracks the actual world, then (iii) beliefs in VR track the actual world. I finish by defending the argument from the charge that as counterfactual conditionals are not generally transitive in nature, the double truth-tracking account cannot guarantee that beliefs formed in VR will track the actual world.

Infinite Utility and Risk

Hayden Wilkinson (University of Oxford / Australian National University)

14.09, 9:35 – 10:10, Room: Mattos Romão

How should we evaluate worlds containing infinite populations? This is a crucial question for ethical consequentialists – their judgements of acts are determined by evaluations of worlds, and some of our best physical theories now predict that our world will inevitably contain an infinite population (see Knobe et al., 2006; Gott, 2008; Carroll, 2017). It is also not an easy question to answer – after all, standard analysis does not allow us to say that any infinite total value is greater than any other (of the same order of infinity).

Various answers have been proposed. Vallentyne & Kagan (1997) propose an expansionist approach, Bostrom (2011) explores the use of hyperreal numbers and various other methods, and Jonsson & Voorneveld (forthcoming) recommend the use of limit-discounting. Each of these approaches has its share of problems, but here I wish to focus on just one: decisions under risk.

We live in a risky and uncertain world, so ethical theories which judge only in cases of certainty are of no use to us (Jackson, 1991). How then should we evaluate options involving infinite worlds, when we are not certain of which world will be produced? This is a source of frustration for both Bostrom (2011) and Arntzenius (2014), who point out that none of our previous methods work in such cases – none assign cardinal values to worlds, so none allow us to produce expected values.

Only one solution has so far been proposed – Arntzenius (2014) describes a principle which first takes expected utilities over finite sets of locations (which may be people or physical locations) and only then performs an expansionist aggregation. This sidesteps the requirement of each world having a cardinal value to input into an expected utility calculation.

Specifically, Artnzenius (ibid.: 53) proposes:

> [For any acts *A* and *B*,] say that $EU(A) > EU(B)$ iff for all allowable expansions R_1, R_2, R_3, of the decision region R_1 there exists an integer n such that for all $k>n$, $EU(A \text{ in } R_k) > EU(B \text{ in } R_k)$. (And the same goes for $EU(A) < EU(B)$ and $EU(A) = EU(B)$, mutatis mutandis.)

Regions here are all sets of individual locations (perhaps people or perhaps physical coordinates) and expansions of those regions are sets with additional members added (with some restrictions).

In this paper, I demonstrate a substantive problem for Arntzenius approach. There exists a decision scenario in which Arntzenius method diverges sharply from the intuitively correct judgement.

In this scenario, an agent must choose between acts *A* and *B*. Action *A* will yield, with probability 1, world W0 which contains the following local values (with locations Li labelled, although they need not be ordered).

	L_1	L_2	L_3	L_4	L_5	L_6	L_7	L_8	…
W_0:	0	0	0	0	0	0	0	0	…

Action B will yield any of the following worlds W_n, with probability $1/2^n$. Note that no matter which world results, the total value for that with world is clearly probability negative – the sum for each approaches negative infinity.

	L_1	L_2	L_3	L_4	L_5	L_6	L_7	L_8	…
W_1	2	-2	-2	-2	-2	-2	-2	-2	…
W_2	0	8	-8	-8	-8	-8	-8	-8	…
W_3	0	0	32	-32	-32	-32	-32	-32	…
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	

This continues for all positive integers. In general, for all W_n, the value located at L_i is given by:

$$L_i = \begin{cases} 0 \text{ for } i<n \\ 2^{2n-1} \text{ for } i = n \\ -2^{2n-1} \text{ for } i>n \end{cases}$$

Crucially, if we take the expected value for each location under act B, as Arntzenius would recommend, we obtain the following:

	L_1	L_2	L_3	L_4	L_5	L_6	L_7	L_8	…
$EV(L_i)$:	1	1	1	1	1	1	1	1	…

We can now take any set of locations we like any region *R* and that region will contain greater expected value for *B* than for *A*. We can expand the set to include any other locations we want, no matter how many, and *EV*(*A* in *R*) will remain greater. Thus, we have satisfied Arntzenius requirement, and we should say that act *B* has greater expected value than *A*.

Of course, this cannot be right. Every possible state of the world, and every W_n that could result, gives a world which is clearly worse than W_0. And yet Arntzenius rule recommends that we take the worse outcome.

To avoid this conclusion, we must reject expected utilities taken locally. I show that, given some basic assumptions, we need a cardinal value for the world as a whole after all or else we necessarily face this problem.

In this paper, I demonstrate that we can avoid this problem through a formal apparatus similar to Arntzenius and to Vallentyne & Kagan (1997). Just as Vallentyne & Kagan extend finite judgements to infinite cases in cases of certainty, I show that we can do much the same in probabilistic cases.

I present a weak and a strong form of my view. The weak form is necessary to preserve our basic intuitions in finite cases, and resembles Vallentyne & Kagan's Strengthened Basic Idea 1 for cases of certainty.

The strong form matches their General Metaprinciple, and has a controversial implication: that evaluations are relative to the physical location of the agent. I argue that this is a worthwhile cost; it enables us to provide a far more complete evaluation of worlds.

Both forms satisfy transitivity, strong Pareto dominance over states and over locations (with the latter only for value rather than expected value), and finite anonymity over lo- cations. Like their counterpart principles which operate only under certainty, neither view satisfies completeness.

Disputatio about Scientific Realism

Lisa Zorzato and Antonella Foligno (Università di Urbino Carlo Bo)

15.09, 9:35 – 10:10, Room: Sala de Actos

Today the debate between realist and antirealist is of central importance with respect to the truth value attributed to our best scientific theories. Scientific realism is a realism regarding whatever is described by our best scientific theories and it aims at dealing with the following questions: i) can we have compelling reasons to believe in the truth of our scientific beliefs?; ii) which are the criteria used to attribute truth value to scientific theories?

On the one hand it seems fair to admit that scientific realism is the best option to embrace in order to give a definitive answer to these questions; but on the other hand, scientific realism seems hard to defend because, unlike antirealism, it has to bear the burden of proof.

In our presentation we aim at presenting Stanford antirealism position put forward in his *Exceeding our grasp: Science, history, and the Problem of Unconceived Alternatives* (2006), and then at giving some realist's replies inspire us by Chakravartty (2008), Magnus (2010) Forber (2008), Godfrey-Smith (2008), Ruhmkorff (2011), and Devitt (2011).

Besides the two main antirealist arguments (i.e. the "empirical underdetermination of theories by data" (EU); and the "pessimistic induction" (PI) also labelled "pessimistic meta-induction" (PMI)) there is a new challenges offered by Stanford in 2006 which has been labelled "the problem of unconceived alternatives" (PUA) or the "new induction" over the history of science (NI) which combines (EU) with (PMI) suggesting that, because over the history of science there were scientifically plausible alternatives to past accepted theories that past scientists failed to conceive of (some of which have been accepted at a later time, while the theories actually accepted at the time have been later shown to be false), probably there are equally good alternatives to currently accepted theories that modern scientists fail to conceive of (and current theories are equally false). Thus, there are no compelling reasons to believe that our current theories are true.

Towards this latest antirealist argument there are many realist replies which aim at showing that Stanford's argument is inappropriate and it diverts the attention from the main realist claim, namely

the induction over scientific theories in Stanford becomes an induction over scientists and their cognitive abilities in exhausting the set of all plausible alternatives.

Assuming that PUA is similar to PI, we are going to prove that a possible reply to the classic Pessimistic Induction can be also used as a reply to PUA because they are both based on a historical induction. First, a principle of continuity can be established between the different formulations of a theory in order to see which elements of the theory are retained over its historical development. This allow to save at least a partial version of scientific realism.

Secondly, Stanford's argument does not work as a real argument against scientific realism because relies on a distinction between community level proprieties and individual level properties; in fact Stanford appeals to the cognitive limits of scientist's without focusing the attention on the real realist's claim: the comparison between the content of our best scientific theories and the physical reality they aim at describing, explaining and predicting.

Third, the cognitive limits of past scientists would not necessary be the limits of future theorizers because history teaches us that science has been getting better and better. Let us just take for example the last century physics which has undergone to an exponential growth if compared with the centuries before.

Finally – as Devitt suggests – to undermine the PUA challenge we can appeal to the methodological and technological improvements shown in our scientific In fact, this version (with respect to the classic realist reply, namely that there is a difference in the breadth, precision, novelty, or other important features of the predictive and explanatory accomplishment of past and present theories) explains why present theories are more successful and hence removes the whiff of ad-hoc-ery.

LIST OF PARTICIPANTS

Eno Agolli	University of Connecticut, Storrs
Luís Duarte d'Almeida	University of Edinburgh
Lorenzo Azzano	University of Padua
Augustin Baas	University of Geneva / University Paris-Sorbonne
Riccardo Baratella	University of Salzburg
Filippo Batisti	Ca' Foscari University of Venice
Tim Bayne	Monash University
Saba Bazargan-Forward	UC San Diego
Berit Brogaard	University of Miami
Massimiliano Carrara	FISSPA, University of Padua
Eduardo Castro	LanCog, Universidade da Beira Interior
Daniele Cassaghi	University of Milan
Olivier Chassaing	Shopiapol, Paris Nanterre University
Demetra Christopoulou	University of Athens
Ludovica Conti	Università di Pavia
Bogdan Dicher	LanCog, Universidade de Lisboa
Pedro Dinis	LanCog, Universidade de Lisboa
Julian Dodd	University of Manchester
Casey Doyle	University of Oxford
Ainhoa Fernández	University of the Basque Country
Filippo Ferrari	University of Bonn
Maria Paola Sforza Fogliani	University School for Advanced Studies, IUSS Pavia
Antonella Foligno	Università di Urbino Carlo Bo
Martin Fricke	Universidad Nacional Autónoma de México
Aldo Frigerio	Università Cattolica del S. Cuore di Milano
Manuel García-Carpintero	LOGOS-BIAP, University of Barcelona
Joaquim Giannotti	University of Glasgow
Seth Goldwasser	University of Pittsburgh
Jorge Gonçalves	IFILNOVA
Steven S. Gouveia	Universidade do Minho
Zhiwei Gu	Central European University
Tanner Burke Hammond	Boston University
Jaakko Hirvelä	University of Helsinki
Madeleine Hyde	Stockholm University

Silvia Ivani	Tilburg University
Tvrtko Jolić	Institute of Philosophy Zagreb
Bartosz Kaluziński	Adam Mickiewicz University in Poznań
Thomas Kivatinos	The City University of New York
Baptiste Le Bihan	University of Geneva
Jesús López-Campillo	University of Murcia
Giuliana Manca	King's College London
Maria Matuszkiewicz	University of Warsaw
Tom McClelland	University of Warwick
Luís Fernández Moreno	Complutense University of Madrid
Diana Neiva	University of Porto
Marina Ortega-Andrés	University of the Basque Country
Niall Paterson	University of Bristol
Alain Pe-Curto	Yale University
Tommaso Piazza	Università di Pavia
João Pinheiro	Centre for Philosophy of Sciences of the University of Lisbon
Ivory Pribam	Université Panthéon Sorbonne Paris 1
Nemésio Puy	University of Granada
Hili Razinsky	LanCog, Universidade de Lisboa
Henrique Ribeiro	Universidade de Coimbra
Dion Scott-Kakures	Scripps College, Claremont
Georg Schiemer	University of Vienna
João Esteves da Silva	Universidade de Lisboa
Jeroen Smid	University of Manchester / Lund University
Francesco Spada	University of Modena and Reggio Emilia
Shunichi Takagi	University College London
Guido Tana	University of Edinburgh / Universität Leipzig
Maria Paola Tenchini	Università Cattolica del S. Cuore di Milano
Paul D. Thorn	HHU Düsseldorf
Manuel Valente	UCLouvain / Leiden University
Simon Walgenbach	University of Manchester
Billy Wheeler	Sun Yat-Sen University
Hayden Wilkinson	University of Oxford / Australian National University
Lisa Zorzato	Università di Urbino Carlo Bo

U LISBOA | UNIVERSIDADE DE LISBOA
FLUL LETRAS LISBOA
VNIVERSITAS OLISIPONENSIS
SOCIEDADE PORTUGUESA DE FILOSOFIA ANALÍTICA
Instituto de Filosofia UNIVERSIDADE DO PORTO
mlag MIND LANGUAGE & ACTION GROUP
FCT Fundação para a Ciência e a Tecnologia