Holistic Inferential Criteria of Adequate Formalization

FRIEDRICH REINMUTH

Holistic Inferential Criteria of Adequate Formalization

Friedrich Reinmuth

Peregrin and Svoboda propose an inferential and holistic approach to formalization, and a similar approach (to correctness) is considered by Brun. However, while the inferential criteria of adequacy explicitly endorsed by these authors may be holistic "in spirit," they are formulated for single formulas. More importantly, they allow the trivialization of equivalence and face problems when materially correct arguments come into play. Against this background, this paper tries to motivate holistic inferential criteria that compel us to distinguish carefully between non-trivially equivalent formalizations as well as between materially and logically correct arguments on an inferential basis.

The first section of the paper (section 1) discusses some problems faced by the inferential (and semantic) criteria of adequacy proposed by Brun (2004, 2012, 2014) and Peregrin and Svoboda (2013, 2017). According to these authors, inferential criteria are to be applied holistically. Yet, their criteria are formulated for single formulas, which leads to some application problems. More importantly, the criteria face problems that are due to their lack of syntactic sensitivity, e.g. the problem of trivialized equivalence. It is argued that postulating additional subsidiary criteria is not a satisfying option if one wants to defend an inferentialist approach to formalization and holds that there is a systematic connection between syntactic features and inferential roles. In contrast, Brun's postulate of hierarchical structure should be accepted as an important systematic constraint on our judgments of adequacy, albeit one that appears weaker than hoped for in some cases.

In section 2, I will propose holistic inferential criteria in the spirit of Peregrin and Svoboda and provide a more detailed discussion of some of the problems raised in section 1. While the criteria can be used to assess the adequacy of formalizations relative to sets of "sample arguments," they are too weak to distinguish properly between non-trivially equivalent formal-

izations, and face difficulties when materially correct arguments are taken into consideration. Section 3 then turns to the role of sentences in inferential contexts that are not reduced to premise-conclusion arguments, namely, to informal derivations. It is argued that if we see the development of calculi as an attempt to account for the logical correctness of arguments in a systematic way and take the distinctions in inferential roles they offer seriously, we have good reasons to strengthen our inferential criteria so that they compel us to choose between non-trivially equivalent formalizations and to distinguish carefully between materially and logically correct arguments. The last section (section 4) indicates some directions for future research.

1 Adequate Formalization, Inferential Criteria, and Trivialized Equivalence

Brun, who has provided a detailed and thorough investigation of the problems of adequate formalization (2004), and most other authors assume that a basic requirement of adequacy is that formalizations do not render intuitively incorrect arguments formally correct (correctness). Some authors, notably Baumgartner and Lampert (2008; 2010), also advocate views of different strength to the effect that adequate formalizations should not render intuitively correct arguments formally incorrect (completeness).

Peregrin and Svoboda have recently put forward an account of logic in terms of reflective equilibrium in which they promote two such criteria as "inferential" criteria of adequate formalization which they contrast with and prefer to so called "semantic" criteria which rely on comparisons of truth conditions (see 2017, esp. ch. 5 and 6). For them, adequate formalizations (logical forms) "are products of the logicians' efforts to account for the inferential structure of a language, especially to envisage the roles of individual statements within the structure" (2017, 4). Since the formalization of a sentence S aims at "making explicit the place of [...] S within the inferential structure of its natural language by means of associating S with a formula of a logical language" (Peregrin and Svoboda 2017, 69), inferential criteria provide the measure of success.

Before discussing the criteria, I want to introduce the main example used in the following, (the conclusion of) "an inference traditionally attributed to De Morgan" (Brun 2012, 325):

DE MORGAN'S ARGUMENT (DMA).

> Every horse is an animal.
> ∴ Every head of a horse is a head of an animal.

For this example, Brun (2012) discusses the formalization

(P1) $\forall x(Hx \rightarrow Jx)$

of the premise

(PDM) Every horse is an animal

and the formalizations

(C1) $\forall x(Fx \rightarrow Gx)$
(C2) $\forall x(\exists y(Hy \land Ixy) \rightarrow \exists y(Jy \land Ixy))$
(C3) $\forall x \forall y(Hy \land Ixy \rightarrow Jy \land Ixy)$
(C4) $\forall x(Hx \land \exists y Iyx \rightarrow Jx \land \exists y Iyx)$

of the conclusion

(CDM) Every head of a horse is a head of an animal

with a correspondence scheme which agrees with the following, in which entries for "a" and "b" are added:

> CORRESPONDENCE SCHEME: HEADS OF HORSES.
> Fx : x is a head of a horse
> Gx : x is a head of an animal
> Hx : x is a horse
> Ixy : x is a head of y
> Jx : x is an animal
> a : Fury
> b : Batu[1]

Note that Peregrin and Svoboda do not consider the correspondence scheme, which assigns natural language expressions to the non-logical symbols in the formalizing formula, to be part of the formalization. I will follow Peregrin

---

[1] The argument and the formalizations (C1), (C2), and (C3) are also extensively discussed in Brun (2004), while (C4) was introduced by Lampert and Baumgartner (2010).

and Svoboda in this, because correspondence schemes provide a kind of formalization at the atomic level, while I want to pursue an account of the adequacy of formalizations that does not take for granted the adequacy of other formalizations.

The first inferential criterion proposed by Peregrin and Svoboda is labelled "*principle of reliability*" and provides a criterion for the correctness of formalizations:

> REL. Φ counts as an adequate formalization of the sentence *S* in the logical system **L** only if the following holds: If an argument form in which Φ occurs as a premise or as the conclusion is valid in **L**, then all its perspicuous natural language instances in which *S* appears as a natural language instance of Φ are intuitively correct arguments. (2017, 70) [2]

If we assume, for example, that De Morgan's argument is an instance of the classically valid

$$\forall x(Hx \to Jx)$$
$$\therefore \ \forall x(\exists y(Hy \land Ixy) \to \exists y(Jy \land Ixy))$$

then it has to be intuitively correct for (C2) to be an adequate formalization of (CDM) if the logical system is classical logic (which will be the general framework in the following).

As Peregrin and Svoboda point out, (REL) is quite similar to an inferential criterion of correctness proposed by Brun (2014, 104). Peregrin and Svoboda also propose a (comparative) completeness criterion with their "*principle of ambitiousness*":

> AMB. Φ is the more adequate formalization of the sentence *S* in the logical system **L** the more natural language arguments in which *S* occurs as a premise or as the conclusion, which fall into the intended scope of **L** and which are intuitively perspicuous and correct, are instances of valid argument forms of **L** in which Φ appears as the formalization of *S*. (2017, 71) [3]

---

2 To simplify the following discussion, I will largely ignore the restriction to perspicuous arguments.
3 The intended scope of a logical system consists "of the arguments whose correctness is to be demonstrable by means of the [logical] language" (Peregrin and Svoboda 2017, 64–65). Peregrin

It seems clear that (AMB) is intended as a means of comparing formalizations where at least the one to be judged to be more adequate meets (REL). It also seems clear that "more natural language arguments" is to be understood in the sense of "the larger and more varied" (Peregrin and Svoboda 2017, 72). Inferential criteria such as (REL) and (AMB) that are not restricted to manageable sets of arguments can hardly be used to judge formalizations to be (more) adequate as their application obviously faces a, as Baumgartner and Lampert put it, "*termination problem*" (2008, 97).

According to Peregrin and Svoboda, the following holds:

> We can, and [...] do, base our (provisional) selection of the formalization on considering a limited number of sample arguments. Thus, a humanly manageable version of (REL) would not simply require that *all* perspicuous natural language instances of a valid argument form in which Φ occurs in place of S are intuitively correct, but only that this holds for those which are among the actual set of sample arguments. Similarly, we could easily reformulate (AMB) so that it (tentatively) prefers the formalization which merely reveals more intuitively correct *sample* arguments as logically correct. In such case, of course, the procedure of selecting the preferable (tentatively adequate) formalization would yield more reliable results the larger and more varied the set of sample arguments is. (2017, 72)

Moreover, they as well as Brun stress that the (intended) application of their respective criteria presupposes that "the formalizations of all sentences, save the one on which we focus our attention, is unproblematic" (Peregrin and Svoboda 2017, 70; see Brun 2014, 104).

All three authors agree that this, as Brun puts it,

> motivates a holistic approach to formalizing which proceeds by bootstrapping: as a starting point, some formalizations are presumed to be correct and used to test others, but such tests may also lead to revising some of the starting-point formalizations [...]. (2014, 104–5)

---

and Svoboda (2017, 71) relate (AMB) to the definition of the completeness of formalizations in (Baumgartner and Lampert 2008, 103).

However, while it may be the case that "we always test a kind of holistic structure, though we perceive it as testing the single formula" (Peregrin and Svoboda 2017, 70), the criteria are formulated for single formulas. This leads to another application problem: even if we restrict our attention to manageable sets of arguments and even if we assume certain formalizations to be adequate, we still cannot apply (REL) and (AMB) in a "humanly manageable" way. Assume, for example, that our sample set only consists of

(1)   Every head of a horse is a head of an animal.
        ∴ Batu is a head of an animal.

and that we consider (1) not to be intuitively correct. Assume that we want to use (REL) to assess the correctness of (C2) as a formalization of (CDM). Then, we still would have to go through all valid argument forms in which (C2) appears as the only premise and check if one of the conclusions is an adequate formalization of the conclusion of (1). Only if no such argument form exists can we judge (C2) to fulfill the criterion of correctness for the sample set. This holds even if we assume that the conclusion of (1) is adequately formalized by

$$\exists y(Jy \wedge Iby)$$

That the latter formula does not follow from (C2) does not entail that there are no adequate formalizations of the conclusion of (1) which follow from (C2). So, in order to apply (REL) (or (AMB)), we do not only have to assume that other formalizations are "unproblematic," but that they are "fixed" (Peregrin and Svoboda 2017, 75).

However, this makes it difficult to assess the respective merits of alternative formalizations of a sentence since we might want to rely on different formalizations of other sentences. For example, if we want to test (C1), we might want to use another formalization of the conclusion of (1), namely, "*Gb*."

Apart from facing application problems, (REL) and (AMB) are highly insensitive to the syntactic features of formalized sentences and their formalizations. Consequently, the "two principles alone [...] do not seem to be sufficient. The main problem is that they do not distinguish between very dissimilar equivalent formulas" (Peregrin and Svoboda 2017, 72). The reason is that (REL) and (AMB) only consider the validity of argument forms, which for many logical systems, e.g. classical logic, is not affected by the substitution of equivalent formulas. This failure to distinguish between equivalent formulas opens

the way to "unacceptably trivial proofs for inferences involving equivalent sentences" (Brun 2014, 105). As an example, consider (C3) and (C4) and the following two sentences:

(CDM-a) Every horse that has a head is an animal that has *that* head.
(CDM-b) Every horse that has a head is an animal that has *a* head.

Since (C3) and (C4) are equivalent, they can be substituted for each other in classically valid argument forms. Now assume that (C3) is an adequate formalization of (CDM-a) and (C4) is an adequate formalization of (CDM-b). Then, (C4), being equivalent to (C3), should also be considered an adequate formalization of (CDM-a), as substituting (C4) for (C3) does not change the validity of the argument forms used to establish the adequacy of (C3). Similarly, (C3) should also be considered an adequate formalization of (CDM-b). Consequently, one could use just one of the two formulas as an adequate formalization for both sentences and "capture" the intuitive equivalence of the sentences by a trivial argument form in which the one premise is identical to the conclusion. This seems worrisome if one holds that "equivalence is subject to logical proof and should not be trivialized by simply choosing the same formalization for any two equivalent sentences" (Brun 2014, 101).

The trivialization of equivalence is a symptom of the lack of "syntactic sensitivity" of (REL) and (AMB)—and similar criteria that are formulated for premise-conclusion arguments. As Brun rightly remarks: "If there are sentences which are in a non-trivial way equivalent [...], this is a matter not only of their truth-conditions but also of their syntactical features" (2014, 107). Brun, Svoboda and Peregrin also point to a desire for a compositional account of logical analysis, which seems to require some systematic sensitivity to syntactic features of the formalized sentences (Brun 2012, 328; 2014, 108; Peregrin and Svoboda 2017, 73).

In order to achieve "some kind of anchoring of the 'logical form' in the grammatical form of the statement of which it is a logical form" (Peregrin and Svoboda 2017, 73), they propose additional criteria such as the following:

> PT. Other things being equal, Φ is the more adequate formalization of the statement *S* in the logical system **L** the more the grammatical structure of Φ is similar to that of *S*. (2017, 72) [4]

---

4 Brun gives the following examples: "the logical symbols in a formalization Φ must have a counterpart in *S*; Φ's correspondence scheme must not include ordinary language expressions not occurring in *S*" (2012, 326–27).

However, Peregrin and Svoboda consider these criteria to be "more-or-less auxiliary" (2013, p. 2919). Brun comments:

> Rules operating on the syntactical surface implicitly guide the common practice of formalization, but if they are not to classify a great deal of standard formalizations as inadequate, they cannot be taken as strict requirements but must be interpreted very liberally or qualified by a virtually endless list of exceptions. (2014, 107)

Brun suggests that a more sophisticated grammar (and maybe also a more sophisticated logical system) is needed for precise and working syntactic criteria (2012, 328; 2014, 109). Peregrin and Svoboda seem to suggest that the very project of formalization and the development of logical systems go hand in hand with developing a (logical) syntax for the sentences in the intended scope of the logic which is projected into the syntax of the developed logical system(s) (see 2017, esp. chap. 7.3). They seem to presuppose that the non-logical symbols of logical languages are parameters that can be used to replace natural language expressions in order to arrive at (logical) forms of sentences and arguments which can then again be instantiated by natural language sentences and arguments (see 2017, esp. chap. 2.3). In this vein, they speak of "the theory of natural language syntax that has been projected into the language of predicate logic" (2017, 52).

However, if the grammatical theory we use applying (PT) is essentially a logico-syntactic theory that finds expression in the syntax of the logical system in question, applications of (PT) to formalizations of a natural language sentence *S* would presuppose that we have already settled on a formalization of *S* in order to test whether the grammatical structure of formalizations is (more) similar to the grammatical structure of *S*.

As indicated, all three authors seem to assume some connection between syntactic features and inferential roles. Given this presumed connection, one might ask why one does not try to approach syntactic features via inferential roles instead of postulating additional "rules of thumb" or hoping for a more sophisticated grammar, an approach Peregrin and Svoboda seem to advocate and which Brun seems to consider as an option (Brun 2014, 115).

If one tries to develop such an approach, one is well advised to impose systematic constraints on the choice of formalizations. Brun, who advocates systematic formalization, distinguishes two aspects, namely "formalizing

analogous sentences analogously," and "formalizing step by step" (2012, 327). However, Brun is as skeptical about the strict application of these precepts as he is regarding surface rules:

> The common theme behind surface rules and the principles of analogous and step-by-step formalization is that they all become more convincing the more we can spell out in a precise and general manner how sentences are to be formalized based on some syntactic description. (2014, 109)

Again, one might ask why one should not rather use inferential criteria to determine which logico-syntactic structure one should impose on natural language sentences. Why not use inferential criteria to specify "the classes of sentences which can be formalized as instances of the same scheme" (Brun 2014, 108) and base syntactic descriptions on how sentences are to be formalized w.r.t. inferential criteria?

While the syntactic criteria and the precepts of formalizing step-by-step and analogously are, according to their authors, not strictly applicable, Brun also offers a powerful postulate (or criterion) for adequate formalizations that enforces systematic syntactic relations between non-equivalent adequate formalizations of the same sentence, the "postulate of hierarchical structure":

> PHS. If $\Phi = \langle \varphi, \kappa \rangle$ and $\Psi = \langle \psi, \kappa \rangle$ are two adequate formalizations of a sentence S in **L** then either (i) Φ and Ψ are equivalent, or (ii) Φ is more specific than Ψ, or (iii) Ψ is more specific than Φ, or (iv) there is an adequate formalization of S that is more specific than both Φ and Ψ. (2014, 109) [5]

One purpose of (PHS) is that it lets us "argue about the adequacy of formalizations by pointing out that they could (not) plausibly be the product of a systematic procedure" (Brun 2014, 109). The deeper motivation is that it ensures "that the various adequate formalizations of an inference constitute

---

5 Note that for Brun formalizations also contain a correspondence scheme. For this formulation of (PHS) with a fixed correspondence scheme κ, $\Phi = \langle \varphi, \kappa \rangle$ is (**L**-)equivalent to $\Psi = \langle \psi, \kappa \rangle$ iff φ and ψ are (**L**-)equivalent; and Φ is more specific than Ψ "iff φ can be generated from ψ by substitutions $[\alpha/\beta]$ such that either (i) α is a sentence-letter occurring in ψ and β is a formula containing at least one sentential connective or a predicate-letter, or (ii) α is an n-place predicate-letter occurring in ψ and β is an open formula with n free variables containing at least one sentential connective, quantifier or predicate-letter with more than n places" (Brun 2014, 109).

a certain unity" (Brun 2014, 110). Postulates like (PHS) are needed if we want adequate formalization to play a part in a systematic account of the (in)correctness of inferences, e.g. by reaching a state of reflective equilibrium, as envisaged by Brun and Peregrin and Svoboda.

Still, (PHS) explicitly allows equivalent formalizations of the same sentence. Moreover, as noted by Lampert and Baumgartner (2010, 95), (C4) and (C2) are both more specific than (C1). Thus, while one can rule out (C3) as an adequate formalization of (CDM) if (C1) is an adequate formalization of this sentence, the same does not hold for (C4). So, all (PHS) (or the equivalent criterion (HCS), which Brun uses in his 2012 paper)[6] does is that "it rules out that (C2) and (C4) are both adequate without telling us which one is inadequate" (Brun 2012, 329). Brun also holds that "(C4) and (C2) fare equally well with respect to (TC) and surface rules" (2012, 329) where (TC), Brun's semantic criterion, (with an added explanation) reads:

> TC. A formalization $\langle \varphi, \kappa \rangle$ of a sentence S in a logical system **L** is correct iff for every condition c, for every **L**-interpretation $\langle \mathcal{D}, \mathcal{I} \rangle$ corresponding to c and κ, $\mathcal{I}(\varphi)$ matches the truth value of S in c. An **L**-interpretation corresponding to a condition c and a correspondence scheme $\{\langle \alpha_1, a_1 \rangle, ..., \langle \alpha_n, a_n \rangle\}$ is an **L**-structure $\langle \mathcal{D}, \mathcal{I} \rangle$ with domain $\mathcal{D}$ and an interpretation function $\mathcal{I}$, such that $\mathcal{I}(\alpha_i)$ matches the semantic value of a_i in c (for all $1 \leq i \leq n$). (Brun 2014, 105; see 2014, 105–6)

This is due to the fact that "(TC) is not distinctive enough if materially i−valid inferences are involved" (Brun 2012, 327), i.e. informally materially correct inferences. Without going into the details of Brun's argument against the adequacy of (C4), we can note that it relies on the "strategy of analogous formalizations" (Brun 2012, 330) and is thus, according to Brun's own standards, not decisive. As we will see in the next section, inferential criteria for premise-conclusion arguments are "not distinctive enough" either if materially correct arguments are involved.

Up to now, the following picture has emerged: the inferential criteria, promoted in particular by Peregrin and Svoboda (as well as Brun's "semantic" criterion (TC) and, to some extent, (PHS)) do not incorporate the presumed

---

6 (HCS) reads informally: "at least one of two non-equivalent formalizations of the same sentence must be inadequate if neither is more specific than the other and there is not a third adequate formalization more specific than both" (Brun 2012, 329).

systematic relation between syntactic structure and inferential role. While this is especially obvious in the case of non-trivially equivalent formalizations, it also leads to problems when materially correct arguments are involved in the assessment of formalizations. To make up for this, the authors propose auxiliary criteria referring to syntactic features, formalizing step-by-step and the analogous formalization of analogous sentences.

This seems rather strange: if one assumes a systematic connection between syntactic features of sentences and the role they can play in inferences, then inferential criteria of adequacy should not rely on additional side-criteria of dubious applicability to ensure a systematic connection between the syntactic features of sentences and their formalizations. Rather, such a connection should result from the application of inferential criteria.

In the section after the next, I will try to outline such an inferentially oriented approach to the adequacy of formalizations. In the next section, some of the problems raised in this section will be discussed in more detail with respect to holistic inferential criteria in the spirit of (REL) and (AMB).

2 Adequacy and Premise-Conclusion Arguments

As already noted above, Peregrin and Svoboda hold that at least considerations of completeness relative to a logical system have to take into account the "*intended scope* of a logical language, consisting of the arguments whose correctness is to be demonstrable by means of the language" (2017, 64–65). They specify:

> Let us call the set of all the perspicuous arguments which char-acterize the behavior of *S* within the intended scope of a logical system **L** the **L**-*reference arguments for S* and any of its non-empty subsets which consists of arguments considered during a particu-lar procedure of assessing alternative formalizations the **L**-*sample arguments for S*. (2017, 65)

Note that the intended scope of a logical system is not something given. Which arguments we consider to be (more) important reference arguments is part of the "bootstrapping" that Peregrin and Svoboda describe (2017, 74–76). The need for choosing sample arguments (and, importantly, other inferential con-texts) will become clearer once the holistic inferential criteria are formulated. To do this, we need some preparatory definitions. These definitions will be

given for formalizations of English sentences but can easily be generalized. First, we define:

> FORMALIZATION-FUNCTION. Φ *is an* **L**-*formalization function for* **S** if and only if
>
> i) **L** is a logical system; and
> ii) **S** is a non-empty set of English sentences; and
> iii) Φ is a function from **S** to a set of **L**-formulas.

The following table provides examples of first-order formalization functions. The sentences in the domain are noted to the left, while the respective values are noted to the right:

Table 1: Formalization functions (Φ1), (Φ2), (Φ3), and (Φ4)

Sentences in the domain	*Values for*			
	(Φ1)	(Φ2)	(Φ3)	(Φ4)
(CDM): Every head of a horse is a head of an animal	(C1)	(C2)	(C3)	(C4)
(PDM): Every horse is an animal		(P1)		
Batu is a head of a horse	Fb	$\exists y(Hy \wedge Iby)$		
Batu is a head of an animal	Gb	$\exists y(Jy \wedge Iby)$		

The value of a formalization function Φ for a natural language sentence S will be called the *formalization of S w.r.t.* Φ. Thus, the four formalization functions differ in their formalizations of (CDM). They agree in their formalization of (PDM), and (Φ1) also differs from the other three formalization functions in its formalizations of the remaining two sentences.

Now we can define:

> INSTANCE OF AN ARGUMENT FORM. *A is an instance of AF w.r.t. the formalization function* Φ iff there are **S** and **L** such that Φ is an **L**-formalization function for **S** and there are sentences $S_1, ..., S_n$ ($n \geq 1$) in **S** such that $A = \langle S_1, ..., S_n \rangle$ and $AF = \langle \Phi(S_1), ..., \Phi(S_n) \rangle$.

If A is an instance of AF w.r.t. Φ, we will call AF a *formalization of A w.r.t.* Φ. Let us say that A is *an argument over* **S** iff **S** is a set of English sentences and

A is a non-empty finite sequence such that every member of *A* is an element of **S**. So, for example, (DMA) and

(2)   Every head of a horse is a head of an animal.
        Batu is a head of a horse.
    ∴ Batu is a head of an animal.

are arguments over the domain of the formalization functions above.

Note that if Φ is a formalization function for a set **S** of sentences and *A* is an argument over **S**, then there is exactly one formalization of *A* w.r.t. Φ. So, for example,

(3) $\forall x(Hx \rightarrow Jx)$
    ∴ $\forall x(Fx \rightarrow Gx)$

is the formalization of (DMA) w.r.t. (Φ1), while

$\forall x(Hx \rightarrow Jx)$
∴ $\forall x(\exists y(Hy \wedge Ixy) \rightarrow \exists y(Jy \wedge Ixy))$

is its formalization w.r.t. (Φ2).

If *A* is an argument over the domain **S** of an **L**-formalization function Φ, then we will say that *A is* **L**-*correct w.r.t.* Φ iff the formalization of *A* w.r.t. Φ is an **L**-valid argument form. So, for example, (DMA) is classically correct w.r.t. (Φ2), but not w.r.t. (Φ1).

Now, we will formulate relativized criteria in the spirit of (REL) and (AMB) for formalization functions. A relativization to sample classes is not only in order because it may be difficult to survey all arguments over the domain of a formalization function. It also holds—as pointed out above—that we have to decide which arguments to admit to the sample classes and which not. The criteria have the form of definitions, but they refer to the intuitive correctness of arguments and should therefore not be treated as definitions of predicates in terms of other, well-established predicates. For the correctness of formalization functions, we postulate:

COR. Φ *is a correct* **L**-*formalization function for* **S** *w.r.t.* **A** iff

  i) Φ is an **L**-formalization function for **S**; and
  ii) **A** is a non-empty set of arguments over **S**; and
  iii) for every argument *A* in **A** it holds: if *A* is **L**-correct w.r.t. Φ, then *A* is an intuitively correct argument

So, for example, if we consider just the unit set of (2) and take classical first-order logic as the logical system, (Φ1), (Φ2), (Φ3) and (Φ4) are correct formalization functions for their common domain w.r.t. this set if we take (2) to be an intuitively correct argument (which I will assume for the following). Note that we will only consider classical first-order logic for the formal side and therefore largely omit mentioning of the logical system in the remaining part of this section.

We set for complete formalization functions:

> COMP. Φ *is an* **L**-*complete formalization function for* **S** *w.r.t.* **A** iff

   i) Φ is an **L**-formalization function for **S**; and
  ii) **A** is a non-empty set of arguments over **S**; and
 iii) for every argument *A* in **A** it holds: if *A* is an intuitively correct argument, then *A* is **L**-correct w.r.t. Φ.

So, for example, if we consider again just the unit set of (2) , (Φ1), (Φ2), (Φ3) and (Φ4) are all complete formalization functions w.r.t. this set. However, if we extend the set of arguments to include (DMA) (and consider it to be intuitively correct), only (Φ2), (Φ3) and (Φ4) are complete formalization functions w.r.t. the extended set.

Adequacy w.r.t. a set of arguments over the domain of a formalization function is postulated to consist in correctness and completeness w.r.t. that set:

> AD. Φ *is an* **L**-*adequate formalization function for* **S** *w.r.t.* **A** iff

  i) Φ is an **L**-correct formalization function for **S** w.r.t. **A**; and
 ii) Φ is an **L**-complete formalization function for **S** w.r.t. **A**.

So, for example, if we consider again the set {(DMA), (2)}, (Φ2), (Φ3) and (Φ4) are all adequate formalization functions w.r.t. this set, while (Φ1) is not. Note that if a formalization function is correct, complete, or adequate w.r.t. some set of arguments, it is so w.r.t. every non-empty subset of this set.

To make comparative judgments of correctness, completeness, and adequacy, it seems natural to extend the formalization functions in question by adding new pairs of sentences and formulas and to consider different sets of arguments over the (extended) domain. Surely, it seems advisable to assume that "the procedure of selecting the preferable (tentatively adequate)

formalization would yield more reliable results the larger and more varied the set of sample arguments is" (Peregrin and Svoboda 2017, 72). However, we also have to decide which "sample arguments we use to demarcate the scope of the [...] logical system" (Peregrin and Svoboda 2017, 70). The scope of a logical system is not something beyond dispute. So, for example, Lampert and Baumgartner want to use classical first-order logic to cover all kinds of intuitively correct arguments (see 2008; 2010), while Peregrin and Svoboda only want to include "*as many logically correct arguments as possible*" (2017, 71). However, they themselves hold "that no clear boundary between logically correct arguments and those that are correct but not logically correct exists in natural language" (2017, 37). Such a boundary can be drawn w.r.t. a logical system and adequate formalizations but this strategy is not straightforwardly applicable if one still has to determine which formalizations one wants to accept as adequate.

To base our discussion on richer examples, let us consider the following extensions of (Φ2), (Φ3) and (Φ4):

Table 2: Extension of (Φ2), (Φ3) and (Φ4) to (Φ2.1), (Φ3.1) and (Φ4.1)

Sentences in the domain	*Values for*		
	(Φ2.1)	(Φ3.1)	(Φ4.1)
(CDM): Every head of a horse is a head of an animal	(C2)	(C3)	(C4)
(PDM): Every horse is an animal		(P1)	
Batu is a head of a horse		$\exists y(Hy \land Iby)$	
Batu is a head of an animal		$\exists y(Jy \land Iby)$	
(CDM-a): Every horse that has a head is an animal that has that head	(C3)	(C3)	(C4)
(CDM-b): Every horse that has a head is an animal that has a head	(C4)	(C3)	(C4)
Batu is a head of Fury		*Iba*	
Fury is a horse		*Ha*	
Fury has a head		$\exists y Iya$	
Fury is a horse that has a head		$Ha \land \exists y Iya$	
Fury is a horse and Batu is a head of Fury		$Ha \land Iba$	

Fury is an animal	Ja
Fury is an animal that has a head	$Ja \land \exists y Iya$
If Fury is a horse that has a head, then Fury is an animal that has a head	$Ha \land \exists y Iya \rightarrow Ja \land \exists y Iya$
Fury is an animal and Batu is a head of Fury	$Ja \land Iba$
If Fury is a horse and Batu is a head of Fury, then Fury is an animal and Batu is a head of Fury	$Ha \land Iba \rightarrow Ja \land Iba$
If Batu is a head of Fury, then Batu is a head of an animal	$Iba \rightarrow \exists y (Jy \land Iby)$
It holds for everything: if it is a horse and Batu is a head of it, then it is an animal and Batu is a head of it	$\forall y (Hy \land Iby \rightarrow Jy \land Iby)$
Everything is a head of an animal	$\forall x \exists y (Jy \land Ixy)$

The extended formalization functions have a common domain and differ only in their formalizations of (CDM), and (CDM-a) and (CDM-b), respectively.

Now consider the following arguments over the common domain of (Φ2.1), (Φ3.1), (Φ4.1):

(4)   Every horse that has a head is an animal that has that head.
   ∴ Every horse that has a head is an animal that has a head.

and

(5)   Every horse that has a head is an animal that has a head.
   ∴ Every horse that has a head is an animal that has that head.

If we assume that both arguments are intuitively correct, (Φ2.1), (Φ3.1), (Φ4.1) are adequate w.r.t. {(DMA), (2) , (4), (5)}. The difference is that (Φ3.1) and (Φ4.1) trivialize the equivalence between (CDM-a) and (CDM-b).

We can (for our purposes) define two **L**-formalization functions Φ, Φ* to be **L**-*equivalent formalization functions* iff they share the same domain **S** and it holds for every S in **S** that $\Phi(S)$ is **L**-equivalent to $\Phi^*(S)$. According to this definition, (Φ3.1) and (Φ4.1) are equivalent formalization functions w.r.t. classical logic. Thus, they render the same arguments over their common domain classically correct and are not distinguishable regarding their correctness, completeness, or adequacy by applying (COR), (COMP), and (AD). This holds

in general: If **L** is a logical system that allows the substitution of **L**-equivalent formulas, e.g. classical logic, then **L**-equivalent formalization functions cannot be distinguished w.r.t. their correctness, completeness or adequacy by (COR), (COMP), and (AD).

Moreover, these criteria face difficulties when materially correct arguments come into play. To see this, let us turn to the relation between (Φ2.1) on the one hand and (Φ3.1) and (Φ4.1) on the other.

Consider the following arguments over the common domain:

(6)   Every head of a horse is a head of an animal.
∴ If Batu is a head of a horse, then Batu is a head of an animal.

(7)   Every head of a horse is a head of an animal.
∴ If Fury is a horse and Batu is a head of Fury, then Fury is an animal and Batu is a head of Fury.

(8)   Every head of a horse is a head of an animal.
∴ If Fury is a horse that has a head, then Fury is an animal that has a head.

If we assume that all three arguments are intuitively correct, (Φ3.1) and (Φ4.1) are adequate w.r.t. {(DMA), (2) , (4), (5), (6), (7), (8)}, while (Φ2.1) is only adequate w.r.t. {(DMA), (2) , (4), (5), (6)}. How could one argue that one should anyhow prefer (Φ2.1)?

First, we can note that the criterion of correctness put forward by Peregrin and Svoboda, namely

*CorArg**: An argument is correct if the step from its premises to its conclusion is a generally acceptable move in an argumentation, or if it can be reconstructed as composed from such generally acceptable moves (2017, 46)

does not clearly rule out any of the arguments as intuitively incorrect if we do not put further constraints on what moves are "generally acceptable." Given their further explanation of their notion of correctness, namely

> that an argument is correct iff it is safe to move from its premises to its conclusion in the sense that whoever accepts the premises cannot reject the conclusion or, more precisely, whoever *does* reject them will be taken to be either unreasonable, or not un-

derstanding the language in which they are formulated (2017, 46),

the arguments presumably have to be counted as correct since it seems hard to imagine that many competent speakers of English will hold that someone who has accepted the respective premises can reject the respective conclusion.[7]

That Peregrin and Svoboda want to include materially correct arguments amongst the correct arguments[8] is not the only reason why we cannot simply restrict "generally acceptable" to "logically acceptable" to exclude (7) and (8). Another reason is that to apply a notion of logical correctness we would need an account of logical form. If we follow Peregrin and Svoboda's explanation of formal and then logical correctness, we would have to come up with something like logical forms of these arguments and then show that these logical forms have incorrect instances.[9] However, we are just trying to determine a logical form for the arguments in question. Therefore, it seems not an admissible move to just claim that, for example, the logical form of (7) is actually

$$\forall x(\exists y(Hy \wedge Ixy) \rightarrow \exists y(Jy \wedge Ixy))$$
$$\therefore\ Ha \wedge Iba \rightarrow Ja \wedge Iba$$

and that the logical form of (8) is actually

$$\forall x(\exists y(Hy \wedge Ixy) \rightarrow \exists y(Jy \wedge Ixy))$$
$$\therefore\ Ha \wedge \exists yIya \rightarrow Ja \wedge \exists yIya$$

and that (apart from not being classically valid) these have clearly incorrect instances such as

Every child of a mother is a child of a father.

---

7 As mentioned in the preceding section, (TC) faces similar problems, as pointed out by Brun (2012, 327; 2014, 106).

8 More precisely, for Peregrin and Svoboda, correct arguments encompass logically correct, analytically correct and status quo correct arguments, where the latter are "correct due to some fixed and stable (though perhaps not eternal and unalterable) state of the world" (2017, 27).

9 That is at least what one would have to do according to the account offered in chap. 2.3 of (Peregrin and Svoboda 2017). Later, they hold that in the process of reflective equilibrium those arguments come out as logically correct whose "logical form is authorized as valid by logic" (2017, 113). In the present scenario, this would not change much since we would still face the question which logical form we are to assign to the arguments in question.

∴ If Martha is a mother and Rachel is a child of Martha, then Martha is a father and Rachel is a child of Martha.

and

Every child of a mother is a child of a father.
∴ If Martha is a mother that has a child, then Martha is a father that has a child.

respectively. A defender of (Φ3.1) or (Φ4.1) could rightly point out that that would just beg the question since we would simply choose the formalizations of (7) and (8) w.r.t. (Φ2.1) as the appropriate logical forms.

Moreover, a defender of (Φ3.1) or (Φ4.1) could even concede that we should formalize "analogous sentences analogously" (Brun 2012, 327) and that the incorrect instances we produced are to be formalized in line with the formalizations of (7) and (8) w.r.t. (Φ2.1): in the absence of a clear concept of "analogous sentence," a defender of (Φ3.1) or (Φ4.1) can simply hold that the sentences in question are not analogous (see Lampert and Baumgartner 2010, 100–102).

Let us consider another argument, which is used by Lampert and Baumgartner (2010, 97–98) in their argument against Brun's account of formalization:

(9)   Everything is a head of an animal.
   ∴ Every head of a horse is a head of an animal.

If we assume that this argument is intuitively correct, we have an argument for whose unit set it holds that (Φ2.1) is adequate w.r.t. it, while (Φ3.1) and (Φ4.1) are not. However, Peregrin and Svoboda would probably not assume that speakers of English take this argument to be correct. They hold that "the paradoxes of material implication" lead to argument forms that "have instances that hardly any speaker of English would consider to be correct" (2017, 76). Given that the argument in question is basically a quantified version of one of the "paradoxes" (at least regarding its formalization w.r.t. Φ2.1), they would probably assume that not many speakers of English would judge it to be correct. Moreover, speakers (not already indoctrinated logically) might shy away from considering it to be correct because the premise seems not simply false, but absurd.

We could of course consider more arguments, but presumably the problems already encountered would persist. In particular, (Φ2.1) cannot beat (Φ3.1) or

(Φ4.1) on the completeness side if the premise position of the formalization of (CDM) is concerned. On the other hand, we have arguments such as (9) which concern the conclusion position of the formalization of (CDM) and for which (Φ2.1) beats (Φ3.1) and (Φ4.1) on the completeness side. However, such arguments will have premises that seem quite absurd. Concerning correctness, the trouble is that if the premises and the conclusions of arguments seem quite reasonable, it is unclear why competent speakers of English would hold that one can reject the conclusion if one accepts the premises.

In the next section, I will argue that we should not restrict our attention to premise-conclusion arguments but also consider how inferential relations between premises and conclusions can be accounted for inferentially by deriving conclusions from premises.

3   Adequacy and Inferential Sequences

Up to now we have only considered premise-conclusion arguments, such as

    (10)   Every head of a horse is a head of an animal.
                Fury is a horse.
   ∴ If Batu is a head of Fury, then Batu is a head of an animal.

The following is not simply a premise-conclusion argument:

    (11)   Assume every head of a horse is a head of an animal. Then it holds that if Batu is a head of a horse, then Batu is a head of an animal. Now assume Fury is a horse. Assume further that Batu is a head of Fury. Then Fury is a horse and Batu is a head of Fury. Thus, Batu is a head of a horse. Then Batu is a head of an animal. Thus, if Batu is a head of Fury, then Batu is a head of an animal.

Rather, (11) may be called an informal derivation. In it, the premises of (10) and an additional sentence are assumed. That last assumption is discharged in the last step, in which the conclusion of (10) is inferred, so that an informal derivation of the conclusion of (10) from the premises of (10) results.

At least from an inferential perspective, the derivation could be taken to show why the argument is logically correct by deriving its conclusion from its premises only using immediate inference steps that rely only on logico-syntactic features of the sentences involved. Such derivations can be

formalized (more or less) "naturally" in natural deduction calculi such as Lemmon's (1998):[10]

(12)

1	(1)	$\forall x(\exists y(Hy \land Ixy) \to \exists y(Jy \land Ixy))$	Assumption (A)
1	(2)	$\exists y(Hy \land Iby) \to \exists y(Jy \land Iby)$	1 Universal quantifier elimination (UE)
3	(3)	Ha	A
4	(4)	Iba	A
3,4	(5)	$Ha \land Iba$	3, 4 $\land$-introduction ($\land$ I)
3,4	(6)	$\exists y(Hy \land Iby)$	5 Existential quantifier introduction (EI)
1,3,4	(7)	$\exists y(Jy \land Iby)$	2, 6 Modus ponendo ponens (MPP)
1,3	(8)	$Iba \to \exists y(Jy \land Iby)$	4, 7 Conditional proof (CP)

Of course, one can view calculi as technical devices that can be used to prove that a certain formula follows from certain formulas, provided the calculi in question are correct w.r.t. the semantic consequence relation one chooses. However, one can also view logical calculi as an attempt to provide a systematic account of logical inferential relations in terms of syntactic features of formulas, and, via the "bridge" of formalization, of sentences in the scope of the logical system in question.[11] Such a view should appeal to Peregrin and Svoboda, who hold

> that language does not exist in the form of its set of sentences and a relation of inferability, but rather in the form of their generators: words and grammatical rules and basic ('axiomatic') instances of inference, plus rules of their composition. [...] the inferential competence, *viz.* the ability to tell correct inferences from incorrect ones, rests at the bottom on the knowledge of the elementary cases and in the knowledge of the ways of composition of simpler inferences into more complex ones. (2017, 159)

10 The leftmost column records the assumptions on which the formulas depend.
11 Such a view seems (at least partly) attributable to Jaśkowski and Gentzen, the founders of natural deduction, as regards their natural deduction calculi (see Jaśkowski 1934; Gentzen 1969a).

If taken as part of the systematic side in a reflective-equilibrium scenario, one can of course adjust the calculus, but a chosen calculus can radically constrain our commitments to the adequacy of formalizations if we also try to take the generation of logical inferability relations into account. So, for example, (Φ2.1), (Φ3.1), and (Φ4.1) all provide formalizations of the premises and the conclusion of (10) which render this argument classically correct. However, only with (Φ2.1) can we directly formalize (11) by (12).

My aim in the following is to make this idea more precise for natural deduction calculi with linear (and not tree) derivations.[12] Now we cannot speak anymore just of a logical system if a logical system is simply identified by a certain syntax and a consequence relation. Instead, we have to use something more fine-grained, namely a logical calculus w.r.t. which a given sequence of formulas is a derivation or not. I will call a sequence of formulas a *determined sequence of formulas w.r.t.* a calculus iff for every member in the sequence it is determined (for example by some form of commentary or by graphical means) if it is an assumption or an inference (in accordance with some rule). An example is the above derivation in Lemmon's system.

For the natural language side, I will speak of inferential sequences. An example is the introductory example of an informal derivation. Yet, to keep things simple, I will assume that *inferential sequences over* a set **S** of English sentences are finite non-empty sequences of expressions of the form

Assume S

and

Thus S'

where S, S' are in **S**, with "Assume" indicating assumptions and "Thus" inferences.

I will assume that logical calculi are logical systems where an argument form is valid w.r.t. a calculus iff its conclusion can be derived from its premises

---

12 Note that this choice is motivated by the relative ease with which informal derivations are formalized and formal derivations instantiated while the view on calculi sketched above can be applied to other types of calculi as well. Cordes and Reinmuth (2017) discuss the formalization of informal derivations in different types of linear calculi of natural deduction.

in that calculus.[13] Under these assumptions, the criteria of the preceding section can be applied to logical calculi. For ease of exposition, I will restrict the following discussion to formalizations in Lemmon's system. However, they can easily be generalized or applied to other natural deduction calculi with linear derivations:

ADAPTATION OF SOME TERMINOLOGY FOR LEMMON'S CALCULUS.

- Φ *is a formalization function* for* **S** iff Φ is a formalization function for **S** w.r.t. Lemmon's calculus.
- *I is an instance* of H w.r.t. the formalization function* Φ iff there is **S** such that Φ is a formalization function* for **S** and there are sentences $S_1, ..., S_n$ ($n \geq 1$) in **S** such that $I = \langle \ulcorner P_1 S_1 \urcorner, ..., \ulcorner P_n S_n \urcorner \rangle$ and H is a determined sequence of formulas of length n w.r.t. Lemmon's calculus such that for all $i \leq n$ it holds: $H_i = \Phi(S_i)$ and $[[P_i = $ "Assume" and $\Phi(S_i)$ is assumed in line i of H] or [$P_i = $ "Thus" and $\Phi(S_i)$ is inferred in line i of H]].
- *H is a formalization* of I w.r.t. the formalization function* Φ iff I is an instance* of H w.r.t. the formalization function Φ.

Note that I will continue to speak simply of instances, formalizations and formalization functions if I assume there is no danger of confusion. Note also that all formalization functions from the preceding section are also formalization functions w.r.t Lemmon's calculus.

First, let us consider the following inferential sequence over the domain of the formalization functions (Φ2.1), (Φ3.1), and (Φ4.1) from the previous section:

(13)

1. Assume every head of a horse is a head of an animal.
2. Assume Fury is a horse.
3. Assume Batu is a head of Fury.
4. Thus Batu is a head of an animal.
5. Thus if Batu is a head of Fury, then Batu is a head of an animal.

Obviously, this inferential sequence is a shortened version of (11). I take it that many of us would accept it as an informal derivation. However, w.r.t. the basic rules of most natural deduction calculi, its formalization would not be

---

13 Of course, for the usual calculi for classical first-order logic it holds that an argument form is valid in this sense iff it is valid according to the model-theoretic definition of validity for classical first-order logic.

a derivation. While a calculus aims at covering some notion of derivability, it is intended to do so in a way which operates on the syntactic structure of formulas in a systematic way. Of course, we do not have to accept the way in which a given calculus does this. On the other hand, w.r.t. a given calculus, we have to make decisions as to which steps to count as immediate, as "the most distinctive patterns of the inferential landscape" (Peregrin and Svoboda 2017, 161). If we choose a certain formalization function, we also choose which inferential steps from natural language sentences to natural language sentences are instances* of derivations w.r.t. this formalization function and the given calculus and thus immediate in the sense that no intermediate steps are "missing."[14]

So, for example, if we choose one of the formalization functions (Φ2.1), (Φ3.1) and (Φ4.1), we also choose which of the following inferential sequences is an instance* of a derivation in Lemmon's system:

(14) An instance* of a derivation w.r.t. (Φ2.1)

     1. Assume every head of a horse is a head of an animal.
     2. Thus if Batu is a head of a horse, then Batu is a head of an animal.

(15) An instance* of a derivation w.r.t. (Φ3.1)

     1. Assume every head of a horse is a head of an animal.
     2. Thus it holds for everything: if it is a horse and Batu is a head of it, then it is an animal and Batu is a head of it.
     3. Thus if Fury is a horse and Batu is a head of Fury, then Fury is an animal and Batu is a head of Fury.

(16) An instance* of a derivation w.r.t. (Φ4.1)

     1. Assume every head of a horse is a head of an animal.
     2. Thus if Fury is a horse that has a head, then Fury is an animal that has a head.

Each of the three formalization functions renders only one of the three inferential sequences as an instance* of a derivation in Lemmon's system: (Φ2.1) does this for (14), (Φ3.1) for (15), and (Φ4.1) for (16). So, even if someone is

---

14 The task of determining which inferences to count as immediate should be seen as integral to the project of formalization if we hold that one aim of formalization is to make "explicit the inferential properties of expressions of natural language" (Peregrin and Svoboda 2017, 109).

inclined to judge each of the arguments (6), (7), and (8) from the preceding section to be intuitively correct, they ipso facto single out some inferential steps as (not) immediate if they choose one of the formalization functions.

One consideration that takes up the discussion from the preceding section is that if we want to proceed systematically and if we are interested in logical correctness, we may want to endorse inferences as immediate which seem acceptable in prima facie analogous cases. So, if we want to treat

> Every child of a mother is a child of a father.

in the same way as

> Every head of a horse is a head of an animal.

then only (14) seems an option w.r.t. the choice between (14), (15) and (16). Of course, as in the case of premise-conclusion arguments, such considerations are not decisive. However, they have another relevance in the new scenario. Choosing a formalization can be seen as choosing an account of how a sentence functions as a premise or conclusion. If we choose one account, we exclude others. Assume, for example, that we take (6), (7), (8) and (10) to be correct, while we have doubts about (9) and therefore choose, for example, (Φ4.1). Then we can be content in the setting of the preceding section because it renders the first four, but not the last argument correct. However, if we consider (11) an account of how and why (10) is logically correct, we cannot formalize this account if we choose (Φ4.1): if we choose (Φ4.1), we have to view (11) as an elliptical informal derivation in which certain steps are left out. Thus, if we take

(17)
1. Assume every child of a mother is a child of a father.
2. Thus if Rachel is a child of a mother, then Rachel is a child of a father.
3. Assume Martha is a mother.
4. Assume Rachel is a child of Martha.
5. Thus Martha is a mother and Rachel is a child of Martha.
6. Thus Rachel is a child of a mother.
7. Thus Rachel is a child of a father.
8. Thus if Rachel is a child of Martha, then Rachel is a child of a father.

to provide an account of the correctness of

> Every child of a mother is a child of a father.
> Martha is a mother.
> ∴ If Rachel is a child of Martha, then Rachel is a child of a father.

we would also have to explain why we cannot replace "child" by "head," "mother" by "horse," "father" by "animal," "Martha" by "Fury" and "Rachel" by "Batu" to get an account of the correctness of (10).

To illustrate the need to make choices, we can consider another example. Suppose we take the inferential sequence

(18)
  1. Assume every horse that has a head is an animal that has that head.
  2. Thus it holds for everything: if it is a horse and Batu is a head of it, then it is an animal and Batu is a head of it.
  3. Thus if Fury is a horse and Batu is a head of Fury, then Fury is an animal and Batu is a head of Fury.

to provide an account of the logical correctness of

(19)   Every horse that has a head is an animal that has that head.
     ∴ If Fury is a horse and Batu is a head of Fury, then Fury is an animal and Batu is a head of Fury.

and the inferential sequence

(20)
  1. Assume every horse that has a head is an animal that has a head.
  2. Thus if Fury is a horse that has a head, then Fury is an animal that has a head.

to account for the logical correctness of

(21)   Every horse that has a head is an animal that has a head.
     ∴ If Fury is a horse that has a head, then Fury is an animal that has a head.

Then, we can choose (Φ2.1) but neither (Φ3.1) nor (Φ4.1) as each of the latter formalization functions only offers a formalization* of one of the two informal derivations, namely (Φ3.1) of (18) and (Φ4.1) of (20).

One obvious option to make the costs in choosing one or another formalization function explicit is to reformulate the criteria of correctness, com-

pleteness, and adequacy from the preceding section directly for inferential sequences and determined sequences of formulas. Then, for example, (Φ2.1), but neither (Φ3.1) nor (Φ4.1) would be adequate w.r.t. the set {(18), (20)}. Moreover, this would also allow us to treat non-trivially equivalent formalization functions differently since they would be adequate for different sets of inferential sequences. For example, (Φ3.1) but not (Φ4.1) would be adequate w.r.t. {(18)} and (Φ4.1) but not (Φ3.1) would be adequate for {(20)} if we judge (18) and (20) to be informal derivations.

For reasons of space, I will not make this explicit, but propose an inferential version of (PHS) that takes into account the discussion so far and puts systematic inferential constraints on adequacy judgements concerning formalization functions. Still, some preparatory work is required. We can set (remember that the whole discussion is carried out for Lemmon's system):

> DERARG. *H is a derivation for A w.r.t. the formalization function* Φ*
> if and only if
>
> i) there is **S** such that Φ is a formalization function* for **S**, and *A* is an argument over **S**; and
> ii) *H* is a derivation in Lemmon's calculus such that
>
>> a. {φ | φ is an undischarged assumption in *H*} = {φ | φ is a premise in the formalization of *A* w.r.t. Φ}; and
>> b. the conclusion of *H* = the conclusion of the formalization of *A* w.r.t. Φ; and
>> c. every non-logical symbol that occurs in *H* also occurs in the formalization of *A* w.r.t. Φ.

According to this definition, (12) is a derivation for (10) w.r.t. (Φ2.1). Of course, there are also derivations for (10) w.r.t. (Φ3.1) and (Φ4.1). However, these will differ from (12) and will not be formalizations of (a standardized version of) (11). Similarly, while there are derivations for (2) w.r.t. (Φ3) and (Φ4), these will differ considerably from

(22)

1	(1)	$\forall x(Fx \rightarrow Gx)$	A
1	(2)	$Fb \rightarrow Gb$	1 UE
3	(3)	Fb	A
1,3	(4)	Gb	2, 3 MPP

which is a derivation for (2) w.r.t. (Φ1). In contrast to this,

(23)

1	(1)	$\forall x(\exists y(Hy \wedge Ixy) \rightarrow \exists y(Jy \wedge Ixy))$	A
1	(2)	$\exists y(Hy \wedge Iby) \rightarrow \exists y(Jy \wedge Iby)$	1 UE
3	(3)	$\exists y(Hy \wedge Iby)$	A
1,3	(4)	$\exists y(Jy \wedge Iby)$	2, 3 MPP

which is a derivation for (2) w.r.t (Φ2), corresponds closely to (22) and to the informal

1. Assume every head of a horse is a head of an animal
2. Thus if Batu is a head of a horse, then Batu is a head of an animal
3. Assume Batu is a head of a horse
4. Thus Batu is a head of an animal

To make the notion of correspondence more precise, we set:

> CORDER. *H is a derivation that corresponds to H* w.r.t. Φ*, Φ and* **S⁰**
> if and only if

   i) there is **S** such that Φ is a formalization function* for **S**, and **S⁰** ⊆ **S**; and
   ii) there is **S*** such that Φ* is a formalization function* for **S***, and **S⁰** ⊆ **S***; and
   iii) *H* and *H** are derivations in Lemmon's calculus and there is an *n* such that

       a. the length of $H = n =$ the length of *H**, and
       b. for every $i \leq n$ it holds:

  i. if H^*_i is an assumption in line i of H^*, then H_i is an assumption in line i of H, and

  ii. if H^*_i is inferred in line i of H^*, then H_i is inferred in line i of H, and

  iii. for every R: if R is an inference rule of Lemmon's calculus and H^*_i can be inferred in line i of H^* in accordance with R, then H_i can be inferred in line i of H in accordance with R, and

  iv. for every S in $\mathbf{S^o}$: if $H^*_i = \Phi^*(S)$, then $H_i = \Phi(S)$.

According to this definition, (23) is a derivation that corresponds to (22) w.r.t (Φ1) and (Φ2) and their common domain. Apart from the obvious correspondence on the formal side, it holds that those formulas in a certain line that are values for a sentence from the common domain of the two formalization functions are the respective values of the same sentence. Note that (23) also corresponds to (22) w.r.t. (Φ1) and (Φ2.1) and the domain of (Φ1).

 On the other hand, there can be no derivation H^o in Lemmon's system such that H^o corresponds to (22) w.r.t. (Φ1) and (Φ3) or (Φ4). For example, if we tried to find a corresponding derivation for (Φ3), we would come to:

(24)

1	(1)	$\forall x \forall y (Hy \land Ixy \rightarrow Jy \land Ixy)$	A
1	(2)	?	1 UE
3	(3)	$\exists y (Hy \land Iby)$	A
1,3	(4)	$\exists y (Jy \land Iby)$	2, 3 MPP

Obviously, whatever formula we choose to infer by UE in line (2) will itself have a universal quantifier as main operator and thus be an unfit premise for the MPP in the last line. Similarly, for (Φ4), we would arrive at:

(25)

1	(1)	$\forall x (Hx \land \exists y Iyx \rightarrow Jx \land \exists y Iyx)$	A
1	(2)	?	1 UE
3	(3)	$\exists y (Hy \land Iby)$	A
1,3	(4)	$\exists y (Jy \land Iby)$	2, 3 MPP

In this case, whatever formula we choose to infer by UE in line (2) will have an antecedent that differs from the formula in line (3) and a consequent that differs from the formula in line (4) and thus again be an unfit premise for the MPP in the last line. Of course, these results for (Φ3) and (Φ4) carry over to their extensions (Φ3.1) and (Φ4.1).

Considerations concerning corresponding derivations w.r.t. different formalization functions and a subset of their domains could be used to take inferential sequences into account whose formalizations are not derivations, and which may be viewed as elliptical informal derivations. However, I will leave this for another occasion and just put forward an inferential version of (PHS) for formalization functions:[15]

> PHS-INF. If Φ is a formalization function* for **S**, and Φ* is a formalization function* for **S***, and **A** is a non-empty set of arguments over **S** ∩ **S***, then:

> i) Φ is not an adequate formalization function* for **S** w.r.t. **A**; or
> ii) Φ* is not an adequate formalization function* for **S*** w.r.t. **A**; or
> iii) for every A, every H^*: if A is in **A** and H^* is a derivation for A w.r.t. Φ*, then there is an H such that H is a derivation that corresponds to H^* w.r.t. Φ*, Φ, **S** ∩ **S***; or
> iv) for every A, every H: if A is in **A** and H is a derivation for A w.r.t. Φ, then there is an H^* such that H^* is a derivation that corresponds to H w.r.t. Φ, Φ*, **S** ∩ **S***.

This criterion extends the "unity of logical form" (Baumgartner and Lampert 2008, 95) which (PHS) is meant to ensure to the role of formalizations in derivations and thereby strongly constrains judgements of adequacy. So, for example, if we judge (Φ1) to be adequate w.r.t. {(2)}, we cannot judge (Φ3), (Φ4) or any extension of either to be adequate w.r.t. any set **A** of arguments over their respective domains if {(2)} ⊆ **A**. On the one hand, (22) is a derivation for (2) w.r.t. (Φ1) and, as shown above, there are no derivations that correspond to (22) w.r.t. (Φ1) and (Φ3) or (Φ4) and their common domain, a result which carries over to extensions of (Φ3) and (Φ4). On the other hand, we have derivations for (2) w.r.t. (Φ3) and its extensions for which there are no corresponding derivations w.r.t. (Φ3), (Φ1) and the domain of (Φ1), and the same holds for (Φ4) and its extensions. Thus, if (Φ1) is an adequate formalization

---

function* for its domain w.r.t. {(2)}, then (Φ3), (Φ4) as well as their extensions are not.

Also, this inferential version of Brun's (PHS) "punishes" the trivialization of equivalence, because non-trivially equivalent formulas behave differently in the context of derivations. So, for example, we can show that either (Φ3.1) or (Φ4.1) is not an adequate formalization function* for their common domain w.r.t. {(19), (21)} if we accept (PHS-Inf): on the one hand, the formalization of (18) w.r.t. (Φ3.1) is a derivation for (19) w.r.t. (Φ3.1) to which no derivation corresponds w.r.t. (Φ3.1), (Φ4.1) and their common domain. On the other hand, the formalization of (20) w.r.t. (Φ4.1) is a derivation for (21) w.r.t. (Φ4.1) to which no derivation corresponds w.r.t. (Φ4.1), (Φ3.1) and their common domain. Thus, according to (PHS-Inf), at least one of the two formalization functions cannot be adequate. These results also show that (PHS-Inf) cannot be used consistently with the criteria from the preceding section. Rather, these criteria have to be adapted, e.g. by taking into account the derivations for the arguments in the respective sample sets.

If we put the discussion so far in the context of providing a systematic inferential account of logical correctness (e.g. in the context of reaching some form of reflective equilibrium), we should (or, at least, can) treat a derivation for an argument w.r.t. a formalization function as a way of accounting for the logical correctness of that argument. Choosing among formalization functions against the background of a calculus is thus a way of choosing between different ways of accounting for the supposed logical correctness of natural language arguments. Doing this, we have to make decisions and (probably) revise initial judgements. If we want a systematic account of logical correctness in terms of inferential role, where the inferential role of a sentence is tightly connected to the logical form we assign to it, then we have to make some choices.

These choices will also determine which arguments are to be counted as logically correct w.r.t. a certain logical system. If we choose a certain formalization function to be adequate w.r.t. a set **A** of arguments, we also choose which arguments in **A** come out as logically correct. Even if one does not hold that "[w]hatever is informally valid must be shown to be valid on formal grounds by means of a logical formalization involving conceptual analysis" (Baumgartner and Lampert 2008, 105), but tries to formalize intuitively logically correct arguments as logically correct w.r.t. the chosen logical system, one encounters the problem that without a notion of logical correctness, and, in turn, of logical form, one faces just a plentitude of intuitively (in)correct

arguments, as described in the preceding section. But if we see the choice between formalization functions also as a choice as how to account for the logical correctness of arguments, the scenario changes, as choosing a formalization function that covers more intuitively correct arguments than another may well mean choosing a formalization function that does not cover inferential steps we take as immediate, and thus accounts of logical correctness that we want to accept.

So, for example, compared with the scenario at the end of the preceding section, (Φ3.1) and (Φ4.1) do not seem to fare better than (Φ2.1): they also offer an account of the logical correctness of arguments in which (CDM) appears as a premise. However, they cannot offer the account that (Φ2.1) provides. Moreover, if we accept (PHS-Inf), we can show that either (Φ3.1) or (Φ4.1) is not adequate w.r.t. relevant sets of arguments over their domain. Thus, choosing them over (Φ2.1) does not just leave out some intuitively maybe rather dubious arguments. Concerning such arguments, we can see their logical correctness as a by-product of the process of systematization. As Brun stresses, what we want is "a *system*, not merely a list of our commitments" (2014, 113), which forces us, as Peregrin and Svoboda put it, to "impose more order on our language and our reasoning than we are able to *find* there, even at the cost of some Procrustean trimming and stretching" (2017, 102).

4 Directions for Future Research

The approach suggested in the preceding section does not aim at a merely "technical" solution to the problems encountered by inferential criteria in a premise-conclusion setting. Rather, it rests on taking seriously a notion of logical correctness in terms of inferability of the conclusion from the premises in accordance with a finite set of rules for certain expressions and ways of combining them. It would be interesting to investigate to what extent this approach allows us to use the syntax of a logical system to structure and classify natural language sentences relative to that logical system. This should include an investigation into which predicates from the metalanguage for the (syntax of) the logical system can fruitfully be adapted to describe syntactical features of the sentences in the intended scope of the logical system. For example, one could try to account for the sub-sentences of a sentence by recourse to the subformula relation for formulas.

While "an approach to logic [that] is closely allied to inferentialism in the philosophy of language and to theories underlying the so-called proof-

theoretic semantics in logic" (Peregrin and Svoboda 2017, 4) should be more than compatible with taking not only premise-conclusion arguments but also inferential sequences into account when trying to determine the inferential roles of natural language sentences, it seems unclear to which extent the inferential criteria developed here fit into other approaches. This applies in particular to Baumgartner and Lampert's "*new picture* of adequate formalization" (2008, 95), according to which "the difference between informal formal and informal material validity must be dropped" (2008, 105). As the strengthened inferential criteria provide incentives to draw the line between materially and logically correct arguments more sharply, it would be interesting to assess them in the context of the debate surrounding Baumgartner and Lampert's "new picture" (see Baumgartner and Lampert 2008; Lampert and Baumgartner 2010; Brun 2012; Peregrin and Svoboda 2013).

The strengthened inferential criteria force us to make fine-grained choices and, in particular, to choose between equivalent formalizations of the same sentence. However, with this might also come the worry that the choices forced on us are too fine-grained. For example, w.r.t. the basic rules of most natural deduction calculi we have to choose between

1. Assume Fury is a horse and Fury is an animal and Batu is a head of Fury
2. Thus Fury is a horse

and

1. Assume Fury is a horse and Fury is an animal and Batu is a head of Fury
2. Thus Batu is a head of Fury

and, if we keep the order of the conjuncts, cannot choose

1. Assume Fury is a horse and Fury is an animal and Batu is a head of Fury
2. Thus Fury is an animal

Intuitively, all three inferences seem immediate and at least the being forced to choose between the first two seems rather strange. One option is to take this simply as the prize of a systematic account of such inferences in terms of the introduction and elimination rules for conjunction. The decision we have to make is arbitrary and comes at the price of excluding some intuitively

immediate inferences, but it is, according to this option, a price we have to pay.

An alternative option is to liberalize the rules of the calculus to allow, for example, a direct formalization of the three inferential sequences as formal derivations. One direction of future research is a weighing of these options. A related line of inquiry is how one could use the notion of corresponding derivations to take also elliptical informal derivations into account. Last, but not least, one should investigate how the strengthened inferential criteria can be put to work when we deal with argumentative texts which are not readily formalizable but have first to be subjected to some form of argument analysis which may involve hermeneutical considerations (see e.g. Brun 2014; Brun and Betz 2016; Reinmuth 2014).*

Friedrich Reinmuth
University of Greifswald
reinmuth.fr@gmail.com

References

Baumgartner, Michael, and Timm Lampert. 2008. "Adequate Formalization." *Synthese* 164 (1): 93–115. doi:10.1007/s11229-007-9218-1.

Brun, Georg. 2004. *Die richtige Formel. Philosophische Probleme der logischen Formalisierung.* 2nd ed. Heusenstamm b. Frankfurt: Ontos Verlag.

———. 2012. "Adequate Formalization and de Morgan's Argument." *Grazer Philosophische Studien* 85: 325–35. doi:10.1163/9789401208338_017.

———. 2014. "Reconstructing Arguments – Formalization and Reflective Equilibrium." In *Theory and Practice of Logical Reconstruction. Anselm as a Model Case*, edited by Friedrich Reinmuth, Geo Siegward, and Christian Tapp, 94–129. Logical Analysis and History of Philosophy 17. Münster: Mentis Verlag.

Brun, Georg, and Gregor Betz. 2016. "Analysing Practical Argumentation." In *The Argumentative Turn in Policy Analysis. Reasoning about Uncertainty*, edited by Sven Ove Hansson and Gertrude Hirsch Hadorn, 39–78. Logic, Argumentation & Reasoning 10. Cham: Springer International Publishing.

Cordes, Moritz, and Friedrich Reinmuth. 2017. "Commentary and Illocutionary Expressions in Linear Calculi of Natural Deduction." *Logic and Logical Philosophy* 26 (2): 163–96. doi:10.12775/llp.2017.002.

GENTZEN, Gerhard. 1935. "Untersuchungen über das logische Schliessen." *Mathematische Zeitschrift* 39 (2 and 3): 176–210, 405–31. Republished asGentzen (1969b).

———. 1969a. "Investigations into Logical Deduction." In *The Collected Papers of Gerhard Gentzen*, 68–131. Studies in Logic and the Foundations of Mathematics 55. Amsterdam: North-Holland Publishing Co. Translation of Gentzen (1935) by M.E. Szabo.

———. 1969b. *Untersuchungen über das logische Schliessen*. Darmstadt: Wissenschaftliche Buchgesellschaft.

JAŚKOWSKI, Stanisław. 1934. "On the Rules of Suppositions in Formal Logic." *Studia Logica (Poznan Panstwowe Wydawnictowo Naukowe)* 1: 5–32.

LAMPERT, Timm, and Michael BAUMGARTNER. 2010. "The Problem of Validity Proofs." *Grazer Philosophische Studien* 80: 79–109. doi:10.1163/18756735-90000872.

LEMMON, Edward John. 1978. *Beginning Logic*. 2nd ed. Indianapolis, Indiana: Hackett Publishing Co.

———. 1998. *Beginning Logic*. Boca Raton, Florida: CRC Press. Republication ofLemmon (1978).

PEREGRIN, Jaroslav, and Vladimír SVOBODA. 2013. "Criteria for Logical Formalization." *Synthese* 190 (14): 2897–2924. doi:10.1007/s11229-012-0104-0.

———. 2017. *Reflective Equilibrium and the Principles of Logical Analysis. Understanding the Laws of Logic*. London: Routledge.

REINMUTH, Friedrich. 2014. "Hermeneutics, Logic and Reconstruction." In *Theory and Practice of Logical Reconstruction. Anselm as a Model Case*, edited by Friedrich Reinmuth, Geo Siegward, and Christian Tapp, 152–90. Logical Analysis and History of Philosophy 17. Münster: Mentis Verlag.