Ruhr-Universität Bochum
Fakultät für Philosophie und Erziehungswissenschaft
Institut für Philosophie I

# Meaning and identity of proofs in (bilateralist) proof-theoretic semantics

Inaugural-Dissertation zur Erlangung des akademischen Grades eines Doktors der Philosophie der Fakultät für Philosophie und Erziehungswissenschaft der Ruhr-Universität Bochum

vorgelegt von:

Sara Ayhan
geboren am 11.07.1990 in Datteln

Dekanin: Prof. Dr. Grit im Brahm
Erstgutachter: Prof. Dr. Heinrich Wansing
Zweitgutachter: Prof. Dr. Greg Restall
Disputation: 20.07.2023

Tag der Einreichung: 17.03.2023

# Acknowledgements

There are many people who, in some way or another, supported me during my PhD studies and thus, contributed to completing this thesis.

First and foremost, I want to thank **Heinrich Wansing**, my supervisor but more than that a real 'Doktorvater' and mentor to me. I am deeply grateful for the time he spent reading and discussing my papers with me, for his thoughtful and thorough comments, his constant encouragement, and simply for making the PhD a real opportunity for me in the first place. I would also like to thank **Greg Restall** for so kindly agreeing to be my second supervisor, for showing genuine interest in my work and for encouraging me to continue this path. Furthermore, I want to thank the people who read drafts of my papers, spent generously time discussing them with me, and provided most helpful comments, among them **Dave Ripley**, **Luca Tranchini**, **Lloyd Humberstone**, **Hitoshi Omori**, **Ansten Klev** and **Will Stafford**. To all the members and guests of our "Logic in Bochum" group over the past years I want to say thank you for listening to my presentations on parts of this thesis, for providing helpful and constructive feedback, and in general for creating such a lively and friendly atmosphere for research. Among the Bochum group a special thanks goes to **Christian Straßer** and **Nils Kürbis** for agreeing to be part of my PhD committee and to **Dustin Gooßens** and **Daniel Skurt** for proofreading parts of my thesis.

Last but not least, I want to thank my parents, **Roswitha Hansen-Ayhan** and **Mehmet Ayhan**, for unconditionally supporting me in every possible way even though my studies remain a mystery to them, my brother, **Ozan Ayhan**, for being what feels like the biggest believer in my career choice and my abilities, and **Daniel Suer**, my partner and most frequent proofreader, for the constant support, for never doubting my goals and for always finding the right words when I doubted myself.

# List of Articles

The chapters 2 to 5 of this thesis are already published, accepted, and/or submitted to international peer-reviewed philosophical logic journals and essentially unchanged. Chapter 4 has been uploaded in this form at an open-access archive but is also my contribution to an extended, co-authored paper with Heinrich Wansing, to appear in a journal, so I list both references below. For Chapter 5, Section 5.3.4 is not included in the publication indicated below due to publishing restrictions concerning the length of the paper. I added this section for this thesis.

- **Chapter 2:** Ayhan, S. (2021). What is the meaning of proofs? A Fregean distinction in proof-theoretic semantics. *Journal of Philosophical Logic*, 50, 571-591.

- **Chapter 3:** Ayhan, S. (submitted). What are acceptable reductions? Perspectives from proof-theoretic semantics and type theory. *Australasian Journal of Logic*.

- **Chapter 4:**

  - Ayhan, S. (2020). A cut-free sequent calculus for the bi-intuitionistic logic `2Int`. (Unpublished Note) https://arxiv.org/abs/2009.14787.
  - Ayhan, S. & Wansing, H. (forthcoming): On synonymy in proof-theoretic semantics. The case of `2Int`, *Bulletin of the Section of Logic*.

- **Chapter 5:** Ayhan, S. (2021). Uniqueness of logical connectives in a bilateralist setting. In M. Blicha & I. Sedlár (Eds.), *The Logica Yearbook 2020* (pp. 1–16). London: College Publications.

**Chapter 6** is planned to be submitted for publication soon. The most significant changes involve a unified citation style and format as well as consecutive chapter numbering. Furthermore, the abstracts and keywords of the respective papers are omitted.

# Contents

# 1 Introduction

## 1.1 Proof-theoretic semantics - The general idea, development and current state of research

The questions with which the chapters constituting this thesis deal are all located in the field of *proof-theoretic semantics* (PTS), an approach to the meaning of (logical) expressions which is based on the concept of *proof*. As such, PTS is opposed to the standard semantical approach, namely model-theoretic semantics, or truth-conditional semantics, which is based on the notion of truth and characterizes the meaning of connectives in terms of model-theoretic notions.[1] The idea of PTS, on the other hand, is to give the meaning of logical constants in terms of the rules of inference governing them. PTS is, thus, a semantics *in terms of proofs* but also a semantics *of proofs* (Schroeder-Heister, 2022).

Since in PTS meaning is based on how we use expressions in proofs and how we draw inferences - so on fundamentally human activities - PTS is an anti-realist approach: meaning is not considered to be given independently of us but, since proofs are taken to be mental constructions, the semantic values that PTS supposes must be conceived of as (at least in principle) always recognizable.[2] These conceptions - taking proofs over truth, proofs as mental constructions and, thus, only allowing *constructive* proofs - are all also fundamental for intuitionistic logic, which is why this is the logic to which PTS in its usual form is most closely related.

Furthermore, PTS is counted among *inferentialism* or *inferential role semantics* (Brandom, 2000), which can be seen as the broader view that the meaning of linguistic expressions is determined by how the expressions are used in inferences. Thus, PTS very broadly also belongs to the tradition based on the slogan "meaning is use" going back to Wittgenstein's famous remarks in his "Philosophische Untersuchungen".[3] The advantage of such a semantic position and thus, also the advantage of PTS over, e.g., model-theoretic semantics is usually seen in the fact that it manages with less assumptions about metaphysically controversial concepts such as truth, possible worlds, etc. In the spirit of 'Ockham's Razor' (Spade & Panaccio, 2019), having an ontology that is as parsimonious as possible is usually taken as preferable

[1] Of course, there are also other non-standard approaches to semantics, like game-theoretic semantics, algebraic semantics, etc., see (Wansing, 2000) for an overview. However, model-theoretic semantics is usually seen as the main opponent of PTS (Francez, 2015, p. 3).

[2] Following a rather standard practice in the literature, throughout this thesis I will use the terms "proof" and "derivation" interchangeably whenever it is clear from the context whether proofs as epistemic entities or their representation in the form of formal proofs are meant. Where this distinction is important, I will use "proof" as the epistemic entity that is denoted by its "derivation", the linguistic representation.

[3] "Man kann für eine *große* Klasse von Fällen der Benützung des Wortes 'Bedeutung' - wenn auch nicht für *alle* Fälle seiner Benützung - dieses Wort so erklären: Die Bedeutung eines Wortes ist sein Gebrauch in der Sprache" (Wittgenstein, 2006[1953], p. 43).

in philosophy.

Usually, the roots of the idea of PTS are ascribed to Gentzen due to his frequently cited remarks concerning his calculus of Natural Deduction:

> The introductions represent, as it were, the 'definitions' of the symbols concerned, and the eliminations are no more, in the final analysis, than the consequences of these definitions. (Gentzen, 1964[1935], p. 295)

This idea has been picked up and further developed in the second half of the 20th century by - among others - von Kutschera (1968), Prawitz (1973; 1974; 2006) and Dummett (1975; 1991). The term "proof-theoretic semantics", though, was only coined at a conference in 1991 by Schroeder-Heister in order to express a particular focus in the overarching field of *general proof theory* (Schroeder-Heister, 2022), a program which was strongly advocated in the early 1970s by representatives such as Prawitz (1971; 1973), Kreisel (1971), and Martin-Löf (1975). Their motivation was - in opposition to Hilbert's understanding of proof theory - to generate a new appreciation of proofs, namely not as mere tools, which are only studied to achieve specific aims, like establishing the consistency of mathematics, but as objects which are worth to be studied in their own right. On such a view it is, e.g., not only an interesting question *what* can be proved, but rather *how* something can be proved. In this spirit it is also much more interesting, then, to focus on proof systems containing rules for the logical connectives instead of axiomatic proof systems.

Since these beginnings the field of PTS has grown and developed different directions. Some of the questions that are widely discussed concern the exact nature of justification of rules as giving the meaning of logical constants, the relation of PTS to specific systems of logic, or which proof-theoretic format of representation to use. I will give an overview of these discussions in the next subsections. Other approaches, which are worth mentioning, although they are not further relevant for this thesis, are works centering around *base-extension semantics*, i.e., a semantics building upon sets of inference rules for atomic sentences,[4] or attempts, e.g., in (Francez, 2015) (the first and, until now, only monograph on PTS), to transfer the ideas of PTS onto natural languages.

## 1.2 Criteria for meaning-conferring rules

A core problem that PTS has to deal with is which rules of inference are actually admissible in order to be considered as giving meaning to logical connectives. Paradigmatically, this is usually illustrated by the connective `tonk`, which was introduced in Prior's (1960) highly influential paper "The Runabout Inference-Ticket" with the aim of showing such a proof-theoretic approach to semantics to be absurd.

[4] See, e.g., (Makinson, 2014; Sandqvist, 2009) for classical logic, (Sandqvist, 2015) for intuitionistic logic, or (Gheorghiu & Pym, Submitted) on 'traditional' PTS vs. base-extension semantics.

Prior argues that if proof rules were indeed all we needed for a semantics of connectives, nothing would prevent us from considering a connective as meaningful, such as `tonk`, which may be used with the following rules: for any formulas $A$, $B$, from $A$ the formula $A$ `tonk` $B$ may be inferred, and from $A$ `tonk` $B$ the formula $B$ may be inferred. Thus, the rules for `tonk` are a mixture of one of the (usual) introduction rules for disjunction and one of the (usual) elimination rules for conjunction and are problematic insofar as they would allow a derivation of any $B$ from any $A$. One could, in other words, derive everything from anything and thus `tonk` would trivialize the consequence relation. With this objection of Prior, PTS has been given the non-trivial task of coming up with certain criteria for rules so that those, roughly speaking, include the connectives we intuitively assume to be acceptable (such as conjunction, implication, etc.), while excluding the unacceptable connectives, such as `tonk`. Without claiming the following overview to be exhaustive or going into too much detail, I will sketch some of the most prominent approaches to this task here. A certain familiarity with the most basic proof-theoretic concepts, like the types of rules, reductions, normalization, etc. will be assumed for this purpose.

Belnap (1962) was one of the first to give an answer and an attempted solution to the problem posed by `tonk`. His approach is to restrict the set of permissible rules by means of a *conservativity* constraint (see also Dummett, 1991, pp. 217-220). In his opinion, what must not be disregarded is that we do not define rules of inference out of nothing but that we have certain background assumptions about deducibility guiding us. As Belnap (1962, p. 131) observes, Prior also readily uses transitivity as a feature of deduction when he wants to show how `tonk` trivializes inferences. Thus, we must bear in mind, while defining the rules of inference for various constants, that these rules must be consistent with our background assumptions concerning deducibility. So we can distinguish between seemingly meaningful rules, like the rules for conjunction, and meaningless rules, like the ones for `tonk`, "on the grounds of consistency - *i.e.*, consistency with antecedent assumptions" (ibid.). Belnap, then, assumes a system adhering to the usual structural rules of weakening, permutation, contraction and cut to determine the context of deducibility completely, so, according to him, these structural rules say everything universally valid about inferences that can be said. The system can then be extended by definitions of logical connectives in form of giving the rules of inference that govern them. The extension is called *conservative* iff there are no *new* deducibility statements within the extended system of the form $\Gamma \vdash A$ that do *not* contain the new connective in $A$ or any element of $\Gamma$. Thus, there can be new derivations in the extended system, but only ones which contain the new connective in them; anything else must be already derivable before the extension. The constraint on new vocabulary in order to preserve consistency, then, is that an extension of the system must be conservative. This is justified, according to Belnap (1962, p. 132), because with our 'bare' system of structural rules we already have all the universally valid deducibility statements that hold without any reference to specific vocabulary. This way we have a foundation to distinguish between, e.g., conjunction and `tonk`: `tonk` is a problematic case - as opposed to conjunction - in that we get a new derivation, namely $A \vdash B$ for arbitrary $A$ and $B$, not containing `tonk`. Hence, an extension of the logic with a `tonk`-like operator is not conservative because it allows us to make new derivations that do not contain `tonk` in the premises or the conclusion and thus, such a definition is inconsistent.[5]

Another notion which occurs frequently in this context is the notion of *harmony* which should govern the relation between the introduction and elimination rules (henceforth: 'I-rules' and 'E-rules').[6] What exactly constitutes this harmony, however, is not uncontroversial in the literature. Informally, harmony is used as a concept of some kind of balance between I- and E-rules, which somehow ensures an exclusion of `tonk`-like rules. For the formal sketch of the concept various proposals exist (Francez & Dyckhoff, 2012, p. 614).[7]

Figuring prominently here is the so-called *inversion principle* introduced by Prawitz (1965, p. 33), pondering over Gentzen's remarks cited above, and more recently defended by Read (2010) under the name "general-elimination harmony". The inversion principle is supposed to give an answer to the question of how the E-rules could be justified given the I-rules. If a set of I- and E-rules adheres to the inversion principle, this means that "a proof of the conclusion of an elimination is already 'contained' in the proofs of the premisses when the major premiss is inferred by introduction" (Prawitz, 1971, p. 246 f.). In this sense, according to Prawitz, the E-rules are justified by the meaning of the logical constants, which are stated by the I-rules, because the conclusion of the E-rule says not more than what is already given by the meaning of its major premise. `Tonk` fails adhering to this principle, of course, as it cannot be said that in the following derivation using the `tonk` I- and E-rule subsequently, a proof of $B$ is already contained in a proof of $A$:

$$\dfrac{\dfrac{A}{A \ \texttt{tonk} \ B}\ \texttt{tonk}I}{B}\ \texttt{tonk}E$$

Another notion of harmony I want to consider is connected to reduction steps and normalization processes.[8] I- and E-rules are supposed to be in harmony iff they

[5] There is also a different account on how to go about the fact that conservativity depends on our background assumptions about the nature of derivability by Ripley (2015): If we do not assume transitivity, an extension with `tonk` would not be non-conservative and thus, there would be no inconsistency. This is also the core of Cook's (2005) analysis. See also (Wansing, 2006), however, showing that the problems of `tonk` avoided in a non-transitive system can be recreated by other `tonk`-like connectives.

[6] Note, that we are usually working with natural deduction systems in PTS; hence the focus on I- and E-rules. This focus on natural deduction will be the subject of Section 1.3.

[7] Conservativity, for example, is also treated as an approach to harmony (Dummett, 1975, p. 103).

[8] This account is not at all unrelated to Prawitz' account, though, since he argues (1971, p. 251)

have appropriate reduction steps or, to put it differently, iff the proofs containing the rules can be normalized (Dummett, 1991, p. 248). To be more precise, this is what Dummett calls "intrinsic" or "local harmony" as opposed to "total harmony", which he uses for conservativity and which he prefers over intrinsic harmony because with the latter non-conservative extensions of the language might still be possible. The existence of reduction procedures stands for a kind of local harmony because it is a "property solely of the rules governing the logical constant in question" (ibid., 250). However, it is possible to have intrinsic harmony of the rules in a system without having total harmony because the addition of one of the constants to the system can lead to a non-conservative extension of the system (ibid., 290). Read (2010, p. 572), understanding harmony as the inversion principle, does not find it problematic that harmony does not guarantee conservative extension and in his opinion neither does it guarantee normalization. Thus, those concepts should not be mixed up, he claims. He considers cases of inconsistent connectives, for example, self-contradictory ones, in order to show that these can still be said to be in harmony because they follow the inversion principle but they allow for non-conservative extensions of the language and normalization may not be possible with them (ibid., pp. 570-575). According to him, harmony does not need to ensure consistency but rather *coherency*, the lack of which is supposedly the actual problem with the `tonk` rules.

Building up on Read, Tranchini (2015) goes for yet another approach, namely of combing normalizability and conservativity (harmony as *conservativity over normal deducibility*) to achieve a requirement for harmony equivalent to the inversion principle. With this at hand, he is also able to distinguish between paradoxical connectives and `tonk`-like connectives (ibid., p. 412) since paradoxical connectives (satisfying the inversion principle) can be shown to yield conservativity over normal deducibility, whereas `tonk` (not satisfying inversion) is not conservative over normal deducibility.[9]

## 1.3 The proof-theoretic format and its connection to logical systems

The two classes of calculi usually considered in the context of PTS are Natural Deduction (ND) and Sequent Calculus (SC), which were both developed by Gentzen

that the reductions "simply [...] make the inversion principle explicit for the different cases that can arise".

[9]Only on a special understanding of "normal derivations", though, namely taking a derivation to be normal iff no reduction steps can be applied to it as opposed to the more common definition of normality as "containing no maximal formula"; two features, which often coincide, though not necessarily. For definitions along the latter lines, see (Prawitz, 1965, p. 34), (Negri & von Plato, 2001, p. 9), (Francez, 2015, p. 84), (Read, 2010, p. 560), (Tennant, 1982, p. 270) implicitly and explicitly in (Tennant, 1978, p. 76). Van Dalen (2004, p. 192) does define a normal derivation with reference to the inapplicability of further reduction steps, but he simultaneously equates this feature with there being no maximal formula (which he calls "cut" instead).

(1964 [1935]) (not exclusively, though).[10] The presentation of rules according to the ND calculus has been the more popular one in the tradition of PTS. Reasons for this are probably that this seems to adhere best to Gentzen's view expressed in his remarks above or that ND displays certain PTS-desirable properties the best or that it is considered indeed especially *natural* in presenting our way of reasoning. There are some works more recently, though, which try to establish SC as a solid basis for PTS and want to show the possibilities, advantages or even the superiority of this calculus over ND (Ripley, 2015; Schroeder-Heister, 2012a, 2012b; Wansing, 2000). Additionally, there can also be some mixed forms, e.g., having an ND variant of the SC (Schroeder-Heister, 2009).

As mentioned above, the project of PTS is closely related to intuitionistic logic. According to (Schroeder-Heister, 2022) this close link is due to the preference in PTS for ND frameworks and certain features of these which make them especially suitable for intuitionistic logic. One of these features is the *subformula property*, which holds due to the rules we suppose (in this case an intuitionistic Gentzen-style ND calculus) plus the fact that we have a normalization result for this calculus. The subformula property says that for every formula $A$ being deducible from a (possibly empty) set of formulas $\Gamma$ there is a deduction such that every formula occurring within the deduction is a subformula of $A$ or of one of the formulas belonging to $\Gamma$. It means intuitively that there is no need in such a calculus to go outside the given syntactic realm (Sandqvist, 2012, p. 710). From the viewpoint of a PTS theorist classical ND is inferior, then, because it does not have the subformula property due to the *reductio ad absurdum* rule, by which it extends the intuitionistic ND calculus:

$$\begin{array}{c}[A \to \bot] \\ \vdots \\ \dfrac{\bot}{A}\end{array}$$

With this rule it is possible to deduce formulas using syntactically irrelevant formulas. From a PTS view this is inconvenient because it would mean in consequence to say that the validity of some formulas is dependent on the meaning of expressions neither occurring in the formula itself nor in the formulas from which it is deduced. Additionally, Schroeder-Heister (2022) mentions further reasons why the *reductio* rule, as displayed above, is out of order, such as that it does not exhibit the property of *separation* of logical constants because two constants occur, namely $\bot$ and $\to$.[11] To deal with the problem classical logic has with the subformula property, there are, on the one hand, approaches, like (Sandqvist, 2012), trying to create

[10] These terms refer to *classes of calculi* because both of them comprise several specific calculi. If I just use "ND" and "SC" in the thesis, then no further specification is needed in the respective context, either because I refer to certain specified rules or it is rather the *style* of rule representation that is important than the specific calculus.

[11] Zucker & Tragesser (1978, p. 503) formulate this separation condition as follows: "The rules for each constant $c$ [...] are *separated*, i.e., they do not refer to any constant other than $c$". This

classical ND calculi which do have the subformula property.[12] On the other hand, as is mentioned in (Schroeder-Heister, 2022), part of the problem is connected to the ND format, so that a solution can also be to drop this and adopt the SC format instead. Another path that has been taken in order to reconcile PTS with classical logic and - at least in some of these works - to even show that classical logic is *better* suitable for PTS than intuitionistic logic, has occurred under the name of *bilateralism*. Since this approach is especially important for substantive parts of this thesis, I will dedicate the next subsection to sketch its most important ideas and trends.

Before I expand on bilateralism, though, I want to briefly mention another issue that is connected to intuitionistic logic and which is also very important for this thesis, namely the so-called *Curry-Howard correspondence*. This describes an isomorphism between derivations in the implicational fragment of intuitionistic ND and typed terms in the simply typed $\lambda$-calculus, proven in (Howard, 1980[1969]) building upon findings by Curry.[13] The general idea behind this is often captured by the slogans "formulas/propositions-as-types" and "proofs-as-programs", i.e., on this view proofs of propositions correspond to programs of the corresponding type. Thus, this establishes an important correspondence between logic and computation. As Sørensen and Urzyczyn (2006) write, it was discovered that "proof theory and the theory of computation turn out to be two sides of the same field".[14]

## 1.4 Bilateralism

There are different conceptions captured under the notion of *bilateralism*, although the differences have been kept rather concealed in the literature. Although the origin of bilateralism is Rumfitt's (2000) seminal paper in the sense that the concrete term and idea are introduced therein and spelled out thoroughly, there are some predecessors to the general idea that are frequently cited, like (Price, 1983), (Smiley, 1996), and (Humberstone, 2000). The most frequent characterization that is used for bilateralism is that it is a theory of meaning displaying a symmetry between certain notions (or often rather: conditions governing these notions), which have not been

---

property is also sometimes - and also in this thesis - referred to as *purity* of the rules (cf., e.g., Humberstone, 2011). For a list of desirable features of rules from a PTS point of view, see (Wansing, 2000).

[12] Another approach of this kind can be found in (von Plato & Siders, 2012), where they claim that if we work with the so-called *general-elimination rules*, instead of Gentzen-style rules, and if we have a broader understanding of what establishes the subformula property, then classical logic is also compatible with PTS.

[13] The precise correspondence for typed combinatory logic was stated in (Curry & Feys, 1958) but Sørensen and Urzyczyn (2006) claim in their thorough tracing of the history of these ideas that hints on this can already be found in earlier papers from the 1930s, such as (Curry, 1934).

[14] One may be tempted to see the fact that this holds for intuitionistic logic as another point in favor of this logic. However, it must be mentioned that, although extensions to, e.g., classical logic have not seemed obvious for a long time, it has been shown that the same style of correspondence can be obtained here as well, see (Griffin, 1990; Parigot, 1992, 1993).

considered being on a par by ‘conventional’ theories of meaning. The relevant notions are most often *assertion* and *denial*, or *assertibility* and *deniability*, sometimes also *affirmation* and *rejection*.[15] While the former are usually taken to describe speech acts, the latter are usually – though not always (see Ripley (2020a) for a thorough distinction) – considered to describe the corresponding internal cognitive states or attitudes. ‘Assertibility’ and ‘deniability’, on the other hand, are of a third kind, since they can be seen to describe something like properties of propositions. The symmetry between these respective concepts is often described with expressions like “both being primitive”, “not reducible to each other”, “being on a par”, and “of equal importance”. Another point to characterize bilateralism, which is often mentioned, though not as frequent or central as the former point,[16] is that in a bilateral approach the denial of $A$ is not interpreted in terms of, or *as* the assertion of the negation of $A$ but that it is the other way around: In bilateralism rejection and/or denial are usually considered as conceptually prior to negation.

Ripley (2017; 2020a) distinguishes two camps of bilateral theories of meaning in terms of “what kinds of condition on assertion and denial they appeal to” (Ripley, 2020a, p. 50): a warrant-based approach and a coherence-based approach, for the latter of which he himself argues (2013) and which was firstly devised by Restall (2005; 2013).[17] As references for the first camp, which Ripley calls the ‘orthodox’ bilateralism, (Price, 1983), (Smiley, 1996), and (Rumfitt, 2000) are given. Warrant-based bilateralism takes the relevant conditions to be the ones under which propositions can be *warrantedly* asserted or denied. Coherence-based bilateralism, on the other hand, takes the relevant conditions to be the conditions under which *collections* of propositions can be *coherently* asserted and/or denied together.

What the two approaches have in common is that they were both meant, as they were originally devised, to motivate a PTS approach using *classical* instead of intuitionistic logic. What they differ in, though, is in their ‘proof-theoretic outcome’. Rumfitt (2000) uses an ND system with signed formulas for assertion and denial, i.e., rules do not apply to propositions but to speech acts. He argues that the shortcomings that a classical ND calculus has from a PTS point of view are overcome once we consider a calculus containing introduction and elimination rules determining not only the assertion conditions for formulas containing the connective in question but also the denial conditions. Thus, he means to give a motivation why

[15]To give some examples of references using a characterization of essentially this flavor: (Francez, 2014b; Gabbay, 2017; Hjortland, 2014; Kürbis, 2016; Ripley, 2011; Rumfitt, 2000; Wansing, 2017).

[16]The following use this as an additional characterization (while also using the essential characterization that the references in fn. 15 use): (Ferreira, 2008; Francez, 2014a, 2019; Ripley, 2013; Steinberger, 2011). This is not to say that this point does not occur in other works on bilateralism but that it is not used as a *characterizing feature* of bilateralism there.

[17]In (Ripley, 2020a) this one is called the “bounds-based bilateralism”. Interestingly, Restall does not use the expression “bilateralism” at all in the cited works, only later does this term become part of his terminology, e.g., in (Restall, 2021).

the rules of classical logic lay down the meaning of the connectives.[18]

Restall (2005), opting for the coherence-based approach, does the same but coming from another direction, namely in proposing a bilateral reading of classical sequent calculus (i.e., with multiple conclusions) incorporating the speech acts of assertion and denial. In a nutshell, he proposes that having the derivation of a sequent $\Gamma \vdash \Delta$, means that the position of asserting each of the members of $\Gamma$ while simultaneously denying each of the members of $\Delta$ would be 'out of bounds'. In a recent paper, though, Restall (2021) seems convinced by Steinberger's (2011) criticism of multiple-conclusion systems as not adhering to our natural inferential practice and he considers an approach using a natural deduction system instead, which does not employ signed formulas but rather uses different positions for certain commitments from which the inference is drawn to the conclusion, namely from *assumptions*, which are ruled in, and *alternatives*, which are ruled out (see also Restall, in press).[19] The advantage of this system compared to Rumfitt's is that the pragmatic status of discharged formulas is much clearer in that they are taken as temporary suppositions for the sake of argument and not simply as assertions or denials. The latter has been criticized (Kürbis, 2019, p. 221) about Rumfitt's system because it seems to cause problems if we have to say about, e.g., a (discharged) formula signed with +, that an assertion is assumed (discharged) (see also Kürbis, in press, for a more detailed version).

What Ripley (2020a) mentions in a footnote is that there are also other kinds of bilateralism, which do not fit into either camp because they do not consider speech acts (i.e., assertion and denial) as the primary notions to act upon in the context of PTS. The kind of bilateralism which will be advocated in this thesis is exactly of this 'other kind' since what will be rather considered here are notions being on a par with proof, provability, or verification, i.e., refutation, refutability, or falsification, respectively (Wansing, 2010, 2017). The point of interest is, thus, a duality between different inferential relationships. It can be argued, then, that this gives rise to accounting for more than one derivability relation that needs to be implemented in the proof-theoretic framework. Hence, this leads to yet another way to devise proof systems, which can be claimed to establish bilateralism on a very fundamental level. The idea is, thus, to detach bilateralism from being a theory of speech acts and rather propose it being a theory of multiple kinds of inferential relationships.

[18] For critical assessments of that paper, see, e.g., (Gibbard, 2002) or (Kürbis, 2016), the former pointing out that the rules of Rumfitt's system actually yield the logic `N4` and not classical logic, the latter showing that the meaning-theoretical requirements that Rumfitt imposes on his system and which he uses to argue for the preferability of classical logic over intuitionistic logic can also be maintained in a bilateral formulation of intuitionistic logic.

[19] The motivation is still to make a case for classical logic being usable in a PTS framework, although Restall does not seem too dogmatic about anything being 'the best' logic. He also wants to show how such a system can be used for substructural logics.

## 1.5 Overview of the chapters

The origins of PTS lie in the question of what constitutes the meaning of logical connectives and its response: the rules of inference that govern the use of the connective. However, what if we go a step further and ask about the meaning of a proof as a whole? In Chapter 2 I address this question and lay out a framework to distinguish sense and denotation of proofs. Therefore, I will use proof systems with $\lambda$-term annotations, which, I argue, are crucially beneficial for this purpose. Two questions are central here. First of all, if we have two (syntactically) different derivations, does this always lead to a difference, firstly, in sense, and secondly, in denotation? The other question concerns the relation between different kinds of proof systems (here: ND vs. SC) with respect to this distinction. Do the different forms of presenting a proof necessarily correspond to a difference in how the inferential steps are given? In my proposed framework it will be possible to identify denotation as well as sense of proofs not only within one proof system but also between different kinds of proof systems. Thus, I give an account to distinguish a mere syntactic divergence from a divergence in meaning and a divergence in meaning from a divergence of proof objects analogous to Frege's distinction for singular terms and sentences.

Chapter 3 will be concerned - in direct relation to Chapter 2 - with the question what constitutes acceptable reductions. It has been argued that reduction procedures are closely connected to the question about identity of proofs and that accepting certain reductions would lead to a trivialization of identity of proofs in the sense that every derivation of the same conclusion would have to be identified. In this chapter it will be shown that the question, which reductions we accept in our system, is not only important if we see them as generating a theory of proof identity but is also decisive for the more general question whether a proof has meaningful content. There are certain reductions which would not only force us to identify proofs of different arbitrary formulas but which would render derivations in a system allowing them meaningless. To exclude such cases, a minimal criterion is proposed, which reductions have to fulfill to be acceptable. I will do so by, again, using $\lambda$-term-annotated proof systems and thus, benefiting from well-established insights on reductions from the field of type theory.

In Chapter 4 a bilateral `G3`-style sequent calculus for the bi-intuitionistic logic `2Int` will be introduced. A distinctive feature of this calculus, called `SC2Int`, is that it makes use of two kinds of sequents, one representing proofs, the other representing so-called *dual proofs*. Thus, it can be seen as a bilateralist calculus. The structural rules of `SC2Int`, in particular its cut-rules, are shown to be admissible. By giving a proof of cut-elimination the result for the corresponding ND calculus, for which a normal form theorem is proven in (Wansing, 2017), is extended.

Chapter 5 will continue with the setting of bilateralism. I will show the problems that are encountered when dealing with uniqueness of connectives in a bilateralist setting within the larger framework of PTS and suggest a solution. Since the logic `2Int` is suitable for this, I use the sequent calculus introduced in Chapter 4, displaying - just like the corresponding natural deduction system - a consequence relation for provability as well as one dual to provability. I will propose a modified characterization of uniqueness incorporating such a duality of consequence relations, with which we can maintain uniqueness in a bilateralist setting.

Finally, Chapter 6 will in a certain sense accumulate all results of the preceding chapters. Again, the Curry-Howard correspondence will be used as in Chapters 2 and 3 - this time applied to the bi-intuitionistic logic `2Int`, which is the subject of Chapters 4 and 5. The basis will be Wansing's (2016a) natural deduction system, which I will turn into a term-annotated form. In order to deal with the bilateralist aspect of having two derivability relations, we need a type theory that extends to a two-sorted typed $\lambda$-calculus. This, in relation to our proposed theory of sense and denotation of derivations from Chapter 2, will give us interesting perspectives on how to view identity and synonymy of derivations when we are situated in a bilateralist setting.

Since this thesis consists mostly of already published or submitted independent papers, every chapter will be self-contained. Hence, all the necessary definitions and requirements will be given in the respective chapter. For the same reason each chapter features a distinct introduction and conclusion within the frame of the topic and repetitions might occur.

# 2 What is the meaning of proofs?

## A Fregean distinction in proof-theoretic semantics

### 2.1 Introduction

In proof-theoretic semantics (PTS) the meaning of the logical constants is taken to be given by the rules of inference that govern their use. As a proof is constituted by applications of rules of inference, it seems reasonable to ask what the meaning of proofs as a whole would consist of on this account. What we are particularly interested in is a Fregean distinction between sense and denotation in the context of proofs.[20] This account builds up on (Tranchini, 2016), where such a distinction is proposed and used in a proof-theoretic explanation of paradoxes.

The notion of denotation is nothing new in the context of proofs. It is common in the literature on proof theory and PTS (e.g., (Kreisel, 1971, p. 6), Prawitz (1971), Martin-Löf (1975)) to distinguish between *derivations*, as linguistic objects, and *proofs*, as abstract (in the intuitionistic tradition: mental) entities. Proofs are then said to be *represented* or *denoted* by derivations, i.e., the abstract proof object is the denotation of a derivation. The notion of sense, on the other hand, has been more or less neglected. Tranchini (2016), therefore, made a proposal that for a derivation to have sense means to be made up of applications of correct inference rules. While this is an interesting approach to consider, Tranchini only determines whether a proof has sense or not but does not go further into what the sense of a proof exactly consists of, so there might be further questions worth pursuing. We will spell out an account of a distinction between sense and denotation of proofs, which can be considered a full-fledged analogy to Frege's distinction concerning singular terms and sentences.[21] Another question concerns the relation of different kinds of proof systems (intuitionistic natural deduction (ND) and sequent calculus (SC) systems will be considered) with respect to such a distinction. If we have two syntactically different derivations with the same denotation in different proof systems, do they always also differ in sense or can sense be shared over different systems?

### 2.2 Connecting structure and meaning

The basic point of departure is the simple observation that there can be different ways leading from the same premises to the same conclusion, either in different proof systems or also within one system. The focus in this matter so far has been

[20] We assume at least a basic familiarity with this idea, laid out in Frege's famous paper "Über Sinn und Bedeutung" (1892); see (Frege, 1948[1892]) for an English translation.

[21] There is some literature also in the field of proof theory concerned with this Fregean distinction, however, to our knowledge, apart from (Tranchini, 2016) this is not concerned with the sense of derivations but with the sense of sentences: see (Martin-Löf, 2021) or (Sundholm, 1994).

on normal vs. non-normal derivations in ND and correspondingly on derivations containing cut vs. cut-free derivations in SC. However, there can also simply be a change of the order of rule applications that can lead to syntactically different derivations from the same premises to the same conclusion. Does this lead to a different denotation or should we say that it is only the sense that differs in such cases, while the underlying proof stays the same?

### 2.2.1 Normal form and the denotation of derivations

One and the same proof may be linguistically represented by different derivations. We will follow the general opinion in taking proofs to be the denotation - the semantic value - of (valid) derivations. In ND a derivation in *normal form* is the most direct form of representation of its denotation, i.e., the represented proof object. For our purposes we will consider a derivation to be in normal form iff neither $\beta$- nor $\eta$-conversions (see rules below) can be applied to it. A derivation in normal form in ND corresponds to a derivation in cut-free form in SC. In intuitionistic logic derivations in non-normal form in ND (resp. with cut in SC) can be reduced to ones in normal form (resp. cut-free form). These are then thought to represent the same underlying proof, just one more indirectly than the other, because, as Prawitz (1971, p. 257f.) says, they represent the same *idea* this proof is based on. In order to make sense and denotation transparent, our approach will be to encode the derivations with $\lambda$-terms. As is well known, by the Curry-Howard-isomorphism there is a correspondence between the intuitionistic ND calculus and the simply typed $\lambda$-calculus and we can formulate the following ND-rules annotated with $\lambda$-terms together with the usual $\beta$- and $\eta$-conversions for the terms. The $\beta$-conversions correspond to the well-known *reduction* procedures, which can be formulated for every connective in ND (Prawitz, 1965, p. 36f.), while the $\eta$-conversions are usually taken to correspond to proof *expansions* (Martin-Löf, 1975, p. 101). We use $p$, $q$, $r$,... for arbitrary atomic formulas, $A$, $B$, $C$,... for arbitrary formulas, and $\Gamma$, $\Delta$,... for sets of formulas. $\Gamma$, $A$ stands for $\Gamma \cup \{A\}$. For variables in terms $x$, $y$, $z$,... is used and $r$, $s$, $t$,... for arbitrary terms.

**Term-annotated ND-rules**:

$$\dfrac{\begin{matrix}\Gamma, [x : A] \\ \vdots \\ t : B\end{matrix}}{\lambda x.t : A \supset B}\ {\supset}\mathrm{I} \qquad\qquad \dfrac{\begin{matrix}\Gamma \\ \vdots \\ s : A \supset B\end{matrix} \quad \begin{matrix}\Delta \\ \vdots \\ t : A\end{matrix}}{App(s,t) : B}\ {\supset}\mathrm{E}$$

$$\dfrac{\begin{matrix}\Gamma\\ \vdots\\ s:A\end{matrix}\quad\begin{matrix}\Delta\\ \vdots\\ t:B\end{matrix}}{\langle s,t\rangle : A\wedge B}\,{\wedge}\mathrm{I}\qquad \dfrac{\begin{matrix}\Gamma\\ \vdots\\ t:A\wedge B\end{matrix}}{fst(t) : A}\,{\wedge}\mathrm{E}_1\qquad \dfrac{\begin{matrix}\Gamma\\ \vdots\\ t:A\wedge B\end{matrix}}{snd(t) : B}\,{\wedge}\mathrm{E}_2$$

$$\dfrac{\begin{matrix}\Gamma\\ \vdots\\ s:A\end{matrix}}{\texttt{inl}s : A\vee B}\,{\vee}\mathrm{I}_1\qquad \dfrac{\begin{matrix}\Gamma\\ \vdots\\ s:B\end{matrix}}{\texttt{inr}s : A\vee B}\,{\vee}\mathrm{I}_2\qquad \dfrac{\begin{matrix}\Gamma\\ \vdots\\ r:A\vee B\end{matrix}\quad\begin{matrix}\Delta,[x:A]\\ \vdots\\ s:C\end{matrix}\quad\begin{matrix}\Theta,[y:B]\\ \vdots\\ t:C\end{matrix}}{\texttt{case}\ r\ \{x.s \mid y.t\} : C}\,{\vee}\mathrm{E}$$

$$\dfrac{\begin{matrix}\Gamma\\ \vdots\\ t:\bot\end{matrix}}{abort(t) : A}\,{\bot}\mathrm{E}$$

**$\beta$-conversions:**

$$App(\lambda x.t, s) \rightsquigarrow t[s/x]$$

$$fst(\langle s,t\rangle) \rightsquigarrow s \qquad\qquad snd(\langle s,t\rangle) \rightsquigarrow t$$

$$\texttt{case inl}r\ \{x.s \mid y.t\} \rightsquigarrow s[r/x] \qquad\qquad \texttt{case inr}r\ \{x.s \mid y.t\} \rightsquigarrow t[r/y]$$

**$\eta$-conversions:**

$$\lambda x.App(t,x) \rightsquigarrow t \text{ (if } x \text{ not free in } t)$$

$$\langle fst(t), snd(t)\rangle \rightsquigarrow t$$

$$\texttt{case}\ r\ \{t.\texttt{inl}t \mid s.\texttt{inr}s\} \rightsquigarrow r$$

We read $x : A$ as "$x$ is a proof of $A$". $t[t'/x]$ means that in term $t$ every free occurrence of $x$ is substituted with $t'$. The usual capture-avoiding requirements for variable substitution are to be observed and $\alpha$-equivalence of terms is assumed. A term that cannot be converted by either $\beta$- or $\eta$-conversion is in normal form.

Since there is a correspondence between intuitionistic SC and intuitionistic ND, for every derivation in ND there must be a derivation in SC named by the same $\lambda$-term. This correspondence is of course not one-to-one, but many-to-one, i.e., for each proof in ND there are at least potentially different derivations in SC.[22] The

[22] On the complications of such a correspondence and also on giving a term-annotated version of SC, see, e.g., (Barendregt & Ghilezan, 2000; Herbelin, 1994; Negri & von Plato, 2001; Pottinger, 1977; Prawitz, 1965; Urban, 2014; Zucker, 1974). Term-annotated sequent calculi can be found i.a. in (Troelstra & Schwichtenberg, 2000) or (Sørensen & Urzyczyn, 2006), from which our presentation is only a notational variant.

following are our respective SC-rules, where we use the propositional fragment of an intuitionistic SC with independent contexts (Negri & von Plato, 2001, p. 89). The reduction procedures remain the same as above in ND; $\beta$-reduction corresponds to the procedures needed to establish cut-elimination, while $\eta$-conversion corresponds to what may be called "identicals-elimination" (Hacking, 1979) or "identity atomization" (Došen, 2008)[23]:

**Term-annotated G0ip**:

**Logical axiom:**

$$\dfrac{}{x : A \vdash x : A}\ \mathrm{Rf}$$

**Logical rules:**

$$\dfrac{\Gamma \vdash s : A \quad \Delta \vdash t : B}{\Gamma, \Delta \vdash \langle s, t \rangle : A \wedge B}\ \wedge\mathrm{R} \qquad \dfrac{\Gamma, x : A, y : B \vdash s : C}{\Gamma, z : A \wedge B \vdash s[[fst(z)/x]snd(z)/y] : C}\ \wedge\mathrm{L}$$

$$\dfrac{\Gamma \vdash s : A}{\Gamma \vdash \mathtt{inl}s : A \vee B}\ \vee\mathrm{R}_1 \quad \dfrac{\Gamma \vdash s : B}{\Gamma \vdash \mathtt{inr}s : A \vee B}\ \vee\mathrm{R}_2 \quad \dfrac{\Gamma, x : A \vdash s : C \quad \Delta, y : B \vdash t : C}{\Gamma, \Delta, z : A \vee B \vdash \mathtt{case\ z}\ \{x.s \mid y.t\} : C}\ \vee\mathrm{L}$$

$$\dfrac{\Gamma, x : A \vdash t : B}{\Gamma \vdash \lambda x.t : A \supset B}\ \supset\mathrm{R} \qquad \dfrac{\Gamma \vdash t : A \quad \Delta, y : B \vdash s : C}{\Gamma, \Delta, x : A \supset B \vdash s[App(x, t)/y] : C}\ \supset\mathrm{L}$$

$$\dfrac{}{x : \bot \vdash abort(x) : C}\ \bot\mathrm{L}$$

**Structural rules**:
Weakening:

$$\dfrac{\Gamma \vdash t : C}{\Gamma, x : A \vdash t : C}\ \mathrm{W}$$

Contraction:

$$\dfrac{\Gamma, x : A, y : A \vdash t : C}{\Gamma, x : A \vdash t[x/y] : C}\ \mathrm{C}$$

The rule of cut

$$\dfrac{\Gamma \vdash t : D \quad \Delta, x : D \vdash s : C}{\Gamma, \Delta \vdash s[t/x] : C}\ \mathrm{cut}$$

[23]Showing that it is possible to get rid of axiomatic sequents with complex formulas and derive them from atomic axiomatic sequents. This is also part of cut-elimination but in principle those are separate procedures (Došen, 2008, p. 26).

is admissible in G0ip.

In the left operational rules as well as in the weakening rule we have the case that variables occur beneath the line that are not explicitly mentioned above the line. In these cases the variables must be either fresh or - together with the same type assignment - already occurring in the context Γ, Δ, etc. Same variables can only (but need not) be chosen for the same type, i.e., if a new type occurs in a proof, then a fresh variable must be chosen. If we would allow to chose the same variable for different types, i.e., for example to let $x : A$ and $x : B$ occur in the same derivation this would amount to assuming that arbitrarily different formulas have the same proof, which is not desirable.

### 2.2.2 Identity of proofs and equivalence of derivations

Figuring prominently in the literature on identity of proofs is a conjecture by Prawitz (1971, p. 257) that two derivations represent the same proof iff they are equivalent.[24] This shifts the question of course to asking when two derivations can be considered equivalent. Using the equational theory of the λ-calculus is one way to provide an answer here: terms on the right and the left hand side of the $\beta$- and $\eta$-conversions are considered denotationally equal (Girard, 1989, p. 16). Hence, two derivations can be considered equivalent iff they are $\beta$-$\eta$-equal (Widebäck, 2001, p. 10; Došen, 2003, p. 5; Sørensen & Urzyczyn, 2006, p. 83ff.).[25] The denotation is then seen to be referred to by the term that annotates the formula or sequent to be proven. We will call this the 'end-term' henceforth so that we can cover and compare both ND and SC at once. So, if we have two derivations with essentially different end-terms (in the sense that they are not belonging to the same equivalence class induced by $\beta$-$\eta$-conversion), we would say that they denote essentially different proofs. On the other hand, for two ND-derivations, where one reduces to the other (or both reduce to the same), e.g., via normalization, we have corresponding λ-terms, one $\beta$-reducible to the other (or both $\beta$-reducible to the same term). In this case we would say that they refer to the same proof. Prawitz (1971, p. 257) stresses that this seems evident since two derivations reducing to identical normal derivations must be seen as equivalent. Note that we can also have the case that two derivations of the same formula, which would look identical in a non-term-annotated version, here

[24] Prawitz gives credit for this conjecture to Martin-Löf. See also (Martin-Löf, 1975, p. 102) on this issue, in his terminology "definitional equality".

[25] There is some discussion about whether $\eta$-conversions are indeed identity-preserving. Martin-Löf (1975, p. 100) does not think so, for example. Prawitz (1971, p. 257) is not clearly decided but writes in the context of identity of proofs it would seem "unlikely that any interesting property of proofs is sensitive to differences created by an expansion". Widebäck (2001), relating to results in the literature on the typed λ-calculus like (Friedman, 1975) and (Statman, 1983), argues for $\beta$-$\eta$-equality to give the right account of identity of proofs and Girard (1989, p. 16) does the same, although he mentions, too, that $\eta$-equations "have never been given adequate status" compared to the $\beta$-equations.

for example of ND, are distinguished on the grounds of our term annotation, like the following two derivations:

$$\mathrm{ND}_{1p \supset (p \supset (p \wedge p))} \qquad \dfrac{\dfrac{\dfrac{[x:p]^1 \quad [y:p]^2}{\langle x,y\rangle : p \wedge p}\wedge\mathrm{I}}{\lambda x.\, \langle x,y\rangle : p \supset (p \wedge p)}\supset\mathrm{I}^1}{\lambda y.\lambda x.\, \langle x,y\rangle : p \supset (p \supset (p \wedge p))}\supset\mathrm{I}^2$$

$$\mathrm{ND}_{2p \supset (p \supset (p \wedge p))} \qquad \dfrac{\dfrac{\dfrac{[x:p]^2 \quad [y:p]^1}{\langle x,y\rangle : p \wedge p}\wedge\mathrm{I}}{\lambda y.\, \langle x,y\rangle : p \supset (p \wedge p)}\supset\mathrm{I}^1}{\lambda x.\lambda y.\, \langle x,y\rangle : p \supset (p \supset (p \wedge p))}\supset\mathrm{I}^2$$

The reason for this is that it is possible to generalize these derivations in different directions, which is made explicit by the variables. Hence, the first one can be generalized to a derivation of $B \supset (A \supset (A \wedge B))$, while the second one generalizes to $A \supset (B \supset (A \wedge B))$.[26]

So, encoding derivations with $\lambda$-terms seems like a suitable method to clarify the underlying structure of proofs. There is one kind of conversion left, though, that needs consideration, namely what we will call *permutative conversions*, or also $\gamma$-*conversions*.[27] They become relevant here because we have disjunction as part of our logical vocabulary. Prawitz (1965) was the first to introduce these conversions. In the conjunction-implication-fragment of intuitionistic propositional logic derivations in normal form satisfy the subformula property, i.e., in a normal derivation $\mathscr{D}$ of $A$ from $\Gamma$ each formula is either a subformula of $A$ or of some formula in $\Gamma$. However, with the disjunction elimination rule this property is messed up, since we get to derive a formula $C$ from $A \vee B$ which is not necessarily related to $A$ or $B$. That is why, in order to recover the subformula property, permutation conversions are introduced, which can be presented in their most general form in the following way:

$$\dfrac{\dfrac{\begin{matrix}\Gamma\\ \vdots\\ A \vee B\end{matrix} \quad \begin{matrix}\Delta,[A]\\ \vdots\\ C\end{matrix} \quad \begin{matrix}\Theta,[B]\\ \vdots\\ C\end{matrix}}{C}\vee E}{D} \quad \rightsquigarrow \quad \dfrac{\begin{matrix}\Gamma\\ \vdots\\ A \vee B\end{matrix} \quad \begin{matrix}\Delta,[A]\\ \vdots\\ \dfrac{C}{D}\end{matrix} \quad \begin{matrix}\Theta,[B]\\ \vdots\\ \dfrac{C}{D}\end{matrix}}{D}\vee E$$

Whether or not these are supposed to be taken into the same league as $\beta$- and $\eta$-conversions in matters of identity preservation of proofs is an even bigger dispute than the one mentioned concerning $\eta$-conversions. Prawitz (1971, p. 257) says that while there can be no doubt about the 'proper reductions' having no influence on the identity of the proof, "[t]here may be some doubts concerning the permutative ∨E-[...]reductions in this connection" but does not go into that matter any further.

[26]For a more detailed examination of generalization, see (Widebäck, 2001) or (Došen, 2003).

[27]It goes under various other names, as well, like *permutation/permuting conversions* or *commuting/commutative conversions*. Some also prefer "reductions" but we will go with the - to us seemingly - more neutral "conversions". The term $\gamma$-*conversions* appears in (Lindley, 2007). About these conversions in general, see, e.g., (Prawitz, 1971, pp. 251-259), (Girard, 1989, Ch. 10), (de Groote, 1999), (Francez, 2017).

Since he needs these reductions to prove his normalization theorem, it seems that he would be inclined not to have too many doubts about identity preservation under the permutative conversions. Girard (1989, p. 73), on the other hand, does not seem to be convinced, as he says - considering an example of permutation conversion - that we are forced to identify "*a priori* different deductions" in these cases. Even though he accepts these conversions for technical reasons, he does not seem to be willing to really identify the underlying proof objects. Restall (2017), however, analyzing derivations by assigning to them what he calls "proof terms" rather than $\lambda$-terms, considers the derivations above as merely distinct in representation but not in the underlying proof, which on his account is the same for both. What is more, he does so not only for technical but rather philosophical reasons, since he claims the flow of information from premises to conclusion to be essentially the same. Lindley (2007, p. 258) and Tranchini (2018, p. 1037f.) both make a point about the connection between reductions and expansions (although they speak of certain kinds of "generalized" expansions) on the one hand and ("generalized") permutative conversions on the other, claiming that performing a (generalized) expansion on the left hand side of the conversion above followed by a reduction (and possibly $\alpha$-conversion) just yields the right hand side. To conclude, if we only consider the $\supset$-$\wedge$-fragment of intuitionistic propositional logic, $\beta$-$\eta$-equality is enough, but if we consider a richer vocabulary, it seems to us at least that there are substantial reasons to include permutative conversions in our equational theory.[28] We do not aim to make a final judgment on this issue here. Rather, when we have laid out our distinction about sense and denotation of proofs below, we will consider the matter again and show why it makes no essential difference for our purposes whether we include permutative conversions or not.

## 2.3 The sense of derivations

Let us spell out at this point what exactly we will consider as the sense and also again the denotation of a derivation in our approach:

> **Definition of denotation:** The denotation of a derivation in a system with $\lambda$-term assignment is referred to by the end-term of the derivation. Identity of denotation holds modulo belonging to the same equivalence class induced by the set of $\alpha$-, $\beta$- and $\eta$-conversions of $\lambda$-terms, i.e., derivations that are denoted by terms belonging to the same equivalence class induced by these conversions are identical, they refer to the same proof object.[29]

[28] The consequence for this paper would be of course to add "$\gamma$-conversions" to the list of relevant conversions in our definitions about normal forms, identity of denotation, etc.

[29] We use the more accurate formulation of "belonging to the same equivalence class" here instead of the formulation we used before of two terms "having the same normal form". The reason for this

> **Definition of sense:** The sense of a derivation in a system with $\lambda$-term assignment consists of the set[30] of $\lambda$-terms that occur within the derivation. Only a derivation made up of applications of correct inference rules, i.e., rules that have reduction procedures, can have sense.

### 2.3.1 Change of sense due to reducibility

Concerning a distinction between sense and denotation in the context of proofs, the rare cases where this is mentioned at all deal with derivations one of which is reducible to the other or with $\lambda$-terms which are $\beta$-convertible to the same term in normal form (Girard, 1989, p. 14; Tranchini, 2016, p. 501; Restall, 2017, p. 6). Since Tranchini is the only one to spell out the part about sense in detail, we will briefly summarize his considerations. As mentioned above, in his account, for a derivation to have sense means that it is made up of applications of correct inference rules. The question to be asked, then, is of course what makes up *correct* inference rules? Tranchini's answer is that inference rules are correct if they have reduction procedures available, i.e., a procedure to eliminate any maximal formula resulting from an application of an introduction rule immediately followed by an elimination rule of the same connective. From a PTS point of view, applying reduction procedures can be seen as a way of interpreting the derivation because it aims to bring the derivation to a normal form, i.e., the form in which the derivation represents the proof it denotes most directly (Tranchini, 2016, p. 507).[31] So the reduction procedures are the instructions telling us how to identify the denotation of the derivation, which for Tranchini means that they give rise to the sense of the derivation. If we have two derivations denoting the same proof, for example, one in normal form and the other in a form that can be reduced to the former, we could say in Fregean terminology that they have the same denotation but differ in their sense because they denote the proof in different ways, one directly, the other indirectly. So, we can take as an example the following two derivations, one in normal and one in non-normal form:

$\mathrm{ND}_{p \supset p}$

$$\frac{[x : p]}{\lambda x.x : p \supset p}\supset\mathrm{I}$$

is that while these two properties coincide for most standard cases, they do not necessarily concur when it comes to Lindley's "general permutative conversions" or also to SC in general because in these cases the confluence property is not guaranteed. We want to thank one of the anonymous referees for indicating this important point.

[30] One could also consider the question whether multi-sets are an even better choice here, which would of course yield a much stronger differentiation of senses. The reason why we consider sets instead of multi-sets is that to us the distinctions brought about by multi-sets by, e.g., a variable occurrence more or less, do not seem to go hand in hand with substantial differences in how inferences are built up.

[31] Tranchini does not restrict his examination to derivations that normalize, though, but to the contrary, uses it to analyze non-normalizable derivations, like paradoxical ones.

$\mathrm{ND}_{\text{non-normal } p \supset p}$

$$\dfrac{\dfrac{\dfrac{[x : p]}{\lambda x.x : p \supset p}{\supset}\mathrm{I} \qquad \dfrac{[y : q]}{\lambda y.y : q \supset q}{\supset}\mathrm{I}}{\langle \lambda x.x, \lambda y.y \rangle : (p \supset p) \wedge (q \supset q)}{\wedge}\mathrm{I}}{\boldsymbol{fst(\langle \lambda x.x, \lambda y.y \rangle)} : p \supset p}{\wedge}\mathrm{E}$$

The latter obviously uses an unnecessary detour via the maximal formula $(p \supset p) \wedge (q \supset q)$, which is introduced by conjunction introduction and then immediately eliminated again, thus, producing different and more complex terms than the former derivation. The derivation can be easily reduced to the former, though, which can be also seen by $\beta$-reducing the term denoting the formula to be proven:

$$\boldsymbol{fst(\langle \lambda x.x, \lambda y.y \rangle)} \rightsquigarrow \lambda x.x$$

We can also give an example analogous to the one above, where a non-normal term (highlighted in bold) in SC is created by using the cut rule:[32]

$\mathrm{SC}_{\vdash\, (p \wedge p) \supset (p \vee p)}$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{z : p \vdash z : p}\mathrm{Rf}}{z : p, x : p \vdash z : p}\mathrm{W}}{y : p \wedge p \vdash fst(y) : p}{\wedge}\mathrm{L}}{y : p \wedge p \vdash \mathtt{inl}\, fst(y) : p \vee p}{\vee}\mathrm{R}}{\vdash \lambda y.\mathtt{inl}\, fst(y) : (p \wedge p) \supset (p \vee p)}{\supset}\mathrm{R}$$

$\mathrm{SC}_{\mathrm{cut} \vdash\, (p \wedge p) \supset (p \vee p)}$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{z : p \vdash z : p}\mathrm{Rf}}{z : p, x : p \vdash z : p}\mathrm{W}}{y : p \wedge p \vdash fst(y) : p}{\wedge}\mathrm{L} \quad \dfrac{\dfrac{\dfrac{}{z : p \vdash z : p}\mathrm{Rf}}{x : p, z : p \vdash z : p}\mathrm{W}}{y : p \wedge p \vdash snd(y) : p}{\wedge}\mathrm{L}}{y : p \wedge p, y : p \wedge p \vdash \langle fst(y), snd(y) \rangle : p \wedge p}{\wedge}\mathrm{R}}{y : p \wedge p \vdash \langle fst(y), snd(y) \rangle : p \wedge p}\mathrm{C} \quad \dfrac{\dfrac{\dfrac{}{z : p \vdash z : p}\mathrm{Rf}}{z : p, x : p \vdash z : p}\mathrm{W}}{y : p \wedge p \vdash fst(y) : p}{\wedge}\mathrm{L}}{y : p \wedge p \vdash \boldsymbol{fst\ \langle fst(y), snd(y) \rangle} : p}\mathbf{cut}}{y : p \wedge p \vdash \mathtt{inl}\boldsymbol{fst\ \langle fst(y), snd(y) \rangle} : p \vee p}{\vee}\mathrm{R}}{\vdash \boldsymbol{\lambda y.}\mathtt{inl}\boldsymbol{fst\ \langle fst(y), snd(y) \rangle} : (p \wedge p) \supset (p \vee p)}{\supset}\mathrm{R}$$

[32] Note however, that the connection between the application of cut and the resulting non-normal term is necessary but not sufficient, i.e., there can be applications of cut not creating a non-normal term. A non-normal term is produced if both occurrences of the cut formula in the premises are principal.

$$\boldsymbol{\lambda y}.\mathtt{inl}\boldsymbol{fst}\,\langle \boldsymbol{fst(y)}, \boldsymbol{snd(y)}\rangle \rightsquigarrow \lambda y.\mathtt{inl}fst(y)$$

In this case again the two derivations are essentially the same because the latter can be reduced to the former by eliminating the application of the cut rule. Again, the proof object they represent is thus the same, only the way of making the inference, represented by the different terms occurring within the derivation, differs, i.e., the sense is different.

### 2.3.2 Change of sense due to rule permutations

So far we only considered the case in which there is an identity of denotation but a difference in sense of derivations due to one being represented by a $\lambda$-term in non-normal form reducible to one in normal form. However, we want to show that this is not the only case where we can make such a distinction. This is also the reason why our approach differs from Tranchini's (who works solely in an ND system) in how we grasp the notion of sense of a derivation. Following Tranchini, the derivation having sense at all depends on there being reduction procedures available for the rules that are applied in it. Since we are also interested in a comparison of sense-and-denotation relations between ND and SC systems, our approach requires that there are reduction procedures available for the created *terms*. Thereby we will be able to cover both systems at once.

Encoding the proof systems with $\lambda$-terms also makes the connection between changing the order of the rule applications and the sense-and-denotation distinction transparent, which is the other case we want to cover. In ND *with* disjunction rules, it is possible to have rule permutations producing derivations with end-terms identifiable by means of the permutative conversions. In SC, however, there are more cases of rule permutations possible. When the left disjunction rule is involved, this also leads to different - though $\gamma$-equal - terms; with the left conjunction or implication rule the end-term remains completely unchanged. Consider, e.g., the following three derivations in SC of the same sequent $\vdash ((q\wedge r)\vee p)\supset((p\vee q)\wedge(p\vee r))$:

$SC_{1\vdash((q\wedge r)\vee p)\supset((p\vee q)\wedge(p\vee r))}$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{q\vdash q}\,\text{Rf}}{q\vdash p\vee q}\,\vee\mathbf{R}}{q,r\vdash p\vee q}\,\text{W}}{q\wedge r\vdash p\vee q}\,\wedge\mathbf{L} \quad \dfrac{\dfrac{\dfrac{\dfrac{}{r\vdash r}\,\text{Rf}}{r\vdash p\vee r}\,\vee\mathbf{R}}{q,r\vdash p\vee r}\,\text{W}}{q\wedge r\vdash p\vee r}\,\wedge\mathbf{L}}{q\wedge r, q\wedge r\vdash(p\vee q)\wedge(p\vee r)}\,\wedge\text{R}}{q\wedge r\vdash(p\vee q)\wedge(p\vee r)}\,\text{C} \quad \dfrac{\dfrac{\dfrac{\dfrac{}{p\vdash p}\,\text{Rf}}{p\vdash p\vee q}\,\vee\text{R} \quad \dfrac{\dfrac{}{p\vdash p}\,\text{Rf}}{p\vdash p\vee r}\,\vee\text{R}}{p,p\vdash(p\vee q)\wedge(p\vee r)}\,\wedge\text{R}}{p\vdash(p\vee q)\wedge(p\vee r)}\,\text{C}}{(q\wedge r)\vee p\vdash(p\vee q)\wedge(p\vee r)}\,\vee\text{L}}{\vdash((q\wedge r)\vee p)\supset((p\vee q)\wedge(p\vee r))}\,\supset\text{R}$$

$\text{SC}_{2 \vdash ((q \wedge r) \vee p) \supset ((p \vee q) \wedge (p \vee r))}$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{q \vdash q}\,\text{Rf}}{q, r \vdash q}\,\text{W}}{q \wedge r \vdash q}\,\wedge\mathbf{L}}{q \wedge r \vdash p \vee q}\,\vee\mathbf{R} \quad \dfrac{\dfrac{\dfrac{\dfrac{}{r \vdash r}\,\text{Rf}}{q, r \vdash r}\,\text{W}}{q \wedge r \vdash r}\,\wedge\mathbf{L}}{q \wedge r \vdash p \vee r}\,\vee\mathbf{R}}{q \wedge r, q \wedge r \vdash (p \vee q) \wedge (p \vee r)}\,\wedge\text{R}}{q \wedge r \vdash (p \vee q) \wedge (p \vee r)}\,\text{C} \quad \dfrac{\dfrac{\dfrac{\dfrac{}{p \vdash p}\,\text{Rf}}{p \vdash p \vee q}\,\vee\text{R} \quad \dfrac{\dfrac{}{p \vdash p}\,\text{Rf}}{p \vdash p \vee r}\,\vee\text{R}}{p, p \vdash (p \vee q) \wedge (p \vee r)}\,\wedge\text{R}}{p \vdash (p \vee q) \wedge (p \vee r)}\,\text{C}}{(q \wedge r) \vee p \vdash (p \vee q) \wedge (p \vee r)}\,\vee\text{L}}{\vdash ((q \wedge r) \vee p) \supset ((p \vee q) \wedge (p \vee r))}\,\supset\text{R}$$

$\text{SC}_{3 \vdash ((q \wedge r) \vee p) \supset ((p \vee q) \wedge (p \vee r))}$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{q \vdash q}\,\text{Rf}}{q \vdash p \vee q}\,\vee\mathbf{R}}{q, r \vdash p \vee q}\,\text{W}}{q \wedge r \vdash p \vee q}\,\wedge\mathbf{L} \quad \dfrac{\dfrac{}{p \vdash p}\,\text{Rf}}{p \vdash p \vee q}\,\vee\text{R}}{(q \wedge r) \vee p \vdash p \vee q}\,\underline{\vee\text{L}} \quad \dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{r \vdash r}\,\text{Rf}}{r \vdash p \vee r}\,\vee\mathbf{R}}{q, r \vdash p \vee r}\,\text{W}}{q \wedge r \vdash p \vee r}\,\wedge\mathbf{L} \quad \dfrac{\dfrac{}{p \vdash p}\,\text{Rf}}{p \vdash p \vee r}\,\vee\text{R}}{(q \wedge r) \vee p \vdash p \vee r}\,\underline{\vee\text{L}}}{(q \wedge r) \vee p, (q \wedge r) \vee p \vdash (p \vee q) \wedge (p \vee r)}\,\underline{\wedge\text{R}}}{(q \wedge r) \vee p \vdash (p \vee q) \wedge (p \vee r)}\,\text{C}}{\vdash ((q \wedge r) \vee p) \supset ((p \vee q) \wedge (p \vee r))}\,\supset\text{R}$$

The difference between $\text{SC}_1$ and $\text{SC}_2$ (highlighted in bold) is that the order of applying the right disjunction rule and the left conjunction rule is permuted. The difference between $\text{SC}_1$ and $\text{SC}_3$ (highlighted with underlining) is that the order of applying the right conjunction rule and the left disjunction rule is permuted. The order of applying the right disjunction rule and the left conjunction rule stays fixed this time. Encoded with $\lambda$-terms, though, we see that in the first case, comparing $\text{SC}_1$ and $\text{SC}_2$, the permutation of rule applications produces exactly the same end-term. Both derivations have the same end-term, namely:

$$\boldsymbol{\lambda u.}\texttt{case u}\ \boldsymbol{\{v.}\ \langle \texttt{inr}\boldsymbol{fst(v)}, \texttt{inr}\boldsymbol{snd(v)}\rangle\ \mid \boldsymbol{x.}\ \langle \texttt{inl}\boldsymbol{x}, \texttt{inl}\boldsymbol{x}\rangle \boldsymbol{\}}$$

$\text{SC}_{1 \vdash ((q \wedge r) \vee p) \supset ((p \vee q) \wedge (p \vee r))}$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{y : q \vdash y : q}\,\text{Rf}}{y : q \vdash \texttt{inr}y : p \vee q}\,\vee\mathbf{R}}{y : q, z : r \vdash \texttt{inr}y : p \vee q}\,\text{W}}{v : q \wedge r \vdash \texttt{inr}fst(v) : p \vee q}\,\wedge\mathbf{L} \quad \dfrac{\dfrac{\dfrac{\dfrac{}{z : r \vdash z : r}\,\text{Rf}}{z : r \vdash \texttt{inr}z : p \vee r}\,\vee\mathbf{R}}{y : q, z : r \vdash \texttt{inr}z : p \vee r}\,\text{W}}{v : q \wedge r \vdash \texttt{inr}snd(v) : p \vee r}\,\wedge\mathbf{L}}{v : q \wedge r, v : q \wedge r \vdash \langle \texttt{inr}fst(v), \texttt{inr}snd(v)\rangle : (p \vee q) \wedge (p \vee r)}\,\wedge\text{R}}{v : q \wedge r \vdash \langle \texttt{inr}fst(v), \texttt{inr}snd(v)\rangle : (p \vee q) \wedge (p \vee r)}\,\text{C} \quad \dfrac{\dfrac{\dfrac{\dfrac{}{x : p \vdash x : p}\,\text{Rf}}{x : p \vdash \texttt{inl}x : p \vee q}\,\vee\text{R} \quad \dfrac{\dfrac{}{x : p \vdash x : p}\,\text{Rf}}{x : p \vdash \texttt{inl}x : p \vee r}\,\vee\text{R}}{x : p, x : p \vdash \langle \texttt{inl}x, \texttt{inl}x\rangle : (p \vee q) \wedge (p \vee r)}\,\wedge\text{R}}{x : p \vdash \langle \texttt{inl}x, \texttt{inl}x\rangle : (p \vee q) \wedge (p \vee r)}\,\text{C}}{u : (q \wedge r) \vee p \vdash \texttt{case u}\ \{v.\ \langle \texttt{inr}fst(v), \texttt{inr}snd(v)\rangle\ \mid x.\ \langle \texttt{inl}x, \texttt{inl}x\rangle\} : (p \vee q) \wedge (p \vee r)}\,\vee\text{L}}{\vdash \boldsymbol{\lambda u.}\texttt{case u}\ \boldsymbol{\{v.}\ \langle \texttt{inr}\boldsymbol{fst(v)}, \texttt{inr}\boldsymbol{snd(v)}\rangle\ \mid \boldsymbol{x.}\ \langle \texttt{inl}\boldsymbol{x}, \texttt{inl}\boldsymbol{x}\rangle \boldsymbol{\}} : ((q \wedge r) \vee p) \supset ((p \vee q) \wedge (p \vee r))}\,\supset\text{R}$$

$SC_{2 \vdash ((q \wedge r) \vee p) \supset ((p \vee q) \wedge (p \vee r))}$

$$
\begin{prooftree}
\AxiomC{}
\RightLabel{Rf}
\UnaryInfC{$y : q \vdash y : q$}
\RightLabel{W}
\UnaryInfC{$y : q, z : r \vdash y : q$}
\RightLabel{$\wedge$L}
\UnaryInfC{$v : q \wedge r \vdash fst(v) : q$}
\RightLabel{$\vee$R}
\UnaryInfC{$v : q \wedge r \vdash \mathtt{inr}fst(v) : p \vee q$}
\AxiomC{}
\RightLabel{Rf}
\UnaryInfC{$z : r \vdash z : r$}
\RightLabel{W}
\UnaryInfC{$y : q, z : r \vdash z : r$}
\RightLabel{$\wedge$L}
\UnaryInfC{$v : q \wedge r \vdash snd(v) : r$}
\RightLabel{$\vee$R}
\UnaryInfC{$v : q \wedge r \vdash \mathtt{inr}snd(v) : p \vee r$}
\RightLabel{$\wedge$R}
\BinaryInfC{$v : q \wedge r, v : q \wedge r \vdash \langle \mathtt{inr}fst(v), \mathtt{inr}snd(v) \rangle : (p \vee q) \wedge (p \vee r)$}
\RightLabel{C}
\UnaryInfC{$v : q \wedge r \vdash \langle \mathtt{inr}fst(v), \mathtt{inr}snd(v) \rangle : (p \vee q) \wedge (p \vee r)$}
\AxiomC{}
\RightLabel{Rf}
\UnaryInfC{$x : p \vdash x : p$}
\RightLabel{$\vee$R}
\UnaryInfC{$x : p \vdash \mathtt{inl}x : p \vee q$}
\AxiomC{}
\RightLabel{Rf}
\UnaryInfC{$x : p \vdash x : p$}
\RightLabel{$\vee$R}
\UnaryInfC{$x : p \vdash \mathtt{inl}x : p \vee r$}
\RightLabel{$\wedge$R}
\BinaryInfC{$x : p, x : p \vdash \langle \mathtt{inl}x, \mathtt{inl}x \rangle : (p \vee q) \wedge (p \vee r)$}
\RightLabel{C}
\UnaryInfC{$x : p \vdash \langle \mathtt{inl}x, \mathtt{inl}x \rangle : (p \vee q) \wedge (p \vee r)$}
\RightLabel{$\vee$L}
\BinaryInfC{$u : (q \wedge r) \vee p \vdash \mathtt{case\ u}\ \{v.\, \langle \mathtt{inr}fst(v), \mathtt{inr}snd(v) \rangle \mid x.\, \langle \mathtt{inl}x, \mathtt{inl}x \rangle\} : (p \vee q) \wedge (p \vee r)$}
\RightLabel{$\supset$R}
\UnaryInfC{$\vdash \boldsymbol{\lambda u.}\mathtt{case\ u}\ \boldsymbol{\{v.\, \langle \mathtt{inr}fst(v), \mathtt{inr}snd(v) \rangle \mid x.\, \langle \mathtt{inl}x, \mathtt{inl}x \rangle\}} : ((q \wedge r) \vee p) \supset ((p \vee q) \wedge (p \vee r))$}
\end{prooftree}
$$

Considering the second comparison between $SC_1$ and $SC_3$ the situation is different: here the permutation of rule applications leads to a different end-term. In the end-term for $SC_1$ and $SC_2$ the pairing operation is embedded within the case expression, whereas in the end-term for $SC_3$ the case expression is embedded within the pairing:

$$\lambda u.\, \langle \mathtt{case\ u}\ \{v.\mathtt{inr}fst(v) \mid x.\mathtt{inl}x\}, \mathtt{case\ u}\ \{v.\mathtt{inr}snd(v) \mid x.\mathtt{inl}x\} \rangle$$

$SC_{3 \vdash ((q \wedge r) \vee p) \supset ((p \vee q) \wedge (p \vee r))}$

$$
\begin{prooftree}
\AxiomC{}
\RightLabel{Rf}
\UnaryInfC{$y : q \vdash y : q$}
\RightLabel{$\vee$R}
\UnaryInfC{$y : q \vdash \mathtt{inr}y : p \vee q$}
\RightLabel{W}
\UnaryInfC{$y : q, z : r \vdash \mathtt{inr}y : p \vee q$}
\RightLabel{$\wedge$L}
\UnaryInfC{$v : q \wedge r \vdash \mathtt{inr}fst(v) : p \vee q$}
\AxiomC{}
\RightLabel{Rf}
\UnaryInfC{$x : p \vdash x : p$}
\RightLabel{$\vee$R}
\UnaryInfC{$x : p \vdash \mathtt{inl}x : p \vee q$}
\RightLabel{$\vee$L}
\BinaryInfC{$u : (q \wedge r) \vee p \vdash \mathtt{case\ u}\ \{v.\mathtt{inr}fst(v) \mid x.\mathtt{inl}x\} : p \vee q$}
\AxiomC{}
\RightLabel{Rf}
\UnaryInfC{$z : r \vdash z : r$}
\RightLabel{$\vee$R}
\UnaryInfC{$z : r \vdash \mathtt{inr}z : p \vee r$}
\RightLabel{W}
\UnaryInfC{$y : q, z : r \vdash \mathtt{inr}z : p \vee r$}
\RightLabel{$\wedge$L}
\UnaryInfC{$v : q \wedge r \vdash \mathtt{inr}snd(v) : p \vee r$}
\AxiomC{}
\RightLabel{Rf}
\UnaryInfC{$x : p \vdash x : p$}
\RightLabel{$\vee$R}
\UnaryInfC{$x : p \vdash \mathtt{inl}x : p \vee r$}
\RightLabel{$\vee$L}
\BinaryInfC{$u : (q \wedge r) \vee p \vdash \mathtt{case\ u}\ \{v.\mathtt{inr}snd(v) \mid x.\mathtt{inl}x\} : p \vee r$}
\RightLabel{$\wedge$R}
\BinaryInfC{$u : (q \wedge r) \vee p, u : (q \wedge r) \vee p \vdash \langle \mathtt{case\ u}\ \{v.\mathtt{inr}fst(v) \mid x.\mathtt{inl}x\}, \mathtt{case\ u}\ \{v.\mathtt{inr}snd(v) \mid x.\mathtt{inl}x\} \rangle : (p \vee q) \wedge (p \vee r)$}
\RightLabel{C}
\UnaryInfC{$u : (q \wedge r) \vee p \vdash \langle \mathtt{case\ u}\ \{v.\mathtt{inr}fst(v) \mid x.\mathtt{inl}x\}, \mathtt{case\ u}\ \{v.\mathtt{inr}snd(v) \mid x.\mathtt{inl}x\} \rangle : (p \vee q) \wedge (p \vee r)$}
\RightLabel{$\supset$R}
\UnaryInfC{$\vdash \lambda u.\, \langle \mathtt{case\ u}\ \{v.\mathtt{inr}fst(v) \mid x.\mathtt{inl}x\}, \mathtt{case\ u}\ \{v.\mathtt{inr}snd(v) \mid x.\mathtt{inl}x\} \rangle : ((q \wedge r) \vee p) \supset ((p \vee q) \wedge (p \vee r))$}
\end{prooftree}
$$

When we take a look at how the term-annotated rules must be designed in order to have a correspondence to the respective rules in ND, we see why some permutations of rule applications lead to different end-terms, while others do not; and why SC is in general more flexible in this respect than ND. In SC the left conjunction rule as well as the left implication rule are substitution operations, i.e., they can change their place in the order without affecting the basic term structure because only in the inner term structure terms are substituted with other terms.[33] In ND, on the other hand, there are no substitution operations used in the term assignment, i.e., for each rule application a new basic term structure is created.

How is this related to the distinction between sense and denotation? In cases like $SC_1$ vs. $SC_2$ the way the inference is given differs, which can also be seen in different terms annotating the formulas occurring within the derivation: with otherwise identical terms in the two derivations **inry** and **inrz** only occur in $SC_1$, while **fst(v)** and **snd(v)** only occur in $SC_2$. However, the resulting end-term stays the same,

[33] For $\supset$L the only exception is when an application of this rule is permuted with an application of $\vee$L, which creates a different, though $\gamma$-convertible term.

thus, we would describe the difference between these derivations as a difference in sense but not in denotation. In other cases, when disjunction elimination or the left disjunction rule is involved, permutation of rule applications can lead to a different end-term, as we see above in $SC_1$ vs. $SC_3$. Whether this corresponds to a difference in denotation depends on whether we accept $\gamma$-conversions to be identity-preserving. What all cases have in common, though, is that rule permutation always leads to a difference in sense of the given derivations because the sets of terms occurring within the derivations differ from each other.

### 2.3.3 Philosophical motivation

Let us have a look at how the Fregean conception of sense is received in the literature in order to show the philosophical motivation for adopting such a definition of sense for derivations. According to Dummett (1973, p. 91), Fregean sense is to be considered as a procedure to determine its denotation.[34] Girard (1989, p. 2), in a passage about sense and denotation and the relation between proofs and programs, mentions that the sense is determined by a "sequence of instructions" and when we see in this context terms as representing programs and "the purpose of a program [...] to calculate [...] its denotation" (ibid., p. 17), then it seems plausible to view the terms occurring within the derivation, decorating the intermediate steps in the construction of the complex end-term that decorates the conclusion, as the sense of that derivation. Tranchini holds the reduction procedures to be the sense because these 'instructions' lead to the term in normal form. However, in our framework - because we do not only consider normal vs. non-normal cases - it seems more plausible to look at the exact terms occurring within the derivations and view them as representing the steps in the process of construction encoding how the derivation is built up and leading us to the denotation, the end-term. For us it is therefore only a necessary requirement for the derivation to have sense to contain only terms for which reduction procedures are available but it does not *make up* the sense. In the case of rule permutation we can then say that the proof is essentially the same but the way it is given to us, the way of inference, differs: i.e., the sense differs. This can be read off from the set of terms that occur within the derivation: they end up building the same end-term, but the way it is built differs, the *procedures to determine the denotation* differ. Thus, this allows us to compare differences in sense within one proof system as well as over different proof systems.

Troelstra and Schwichtenberg (2000, p. 74), e.g., give an example of two derivations in SC producing the same end-term in different ways to show that just from the variables and the end-term we cannot read off how the derivation is built up:[35]

[34] This idea of sense as procedures also occurs in more recent publications like Muskens (2005) or Duží, Jespersen, and Materna (2010).

[35] For simplicity we omit the weakening steps that would strictly seen have to precede the appli-

$$\mathrm{SC}_{1 \vdash (s \wedge p) \supset ((q \wedge r) \supset (p \wedge q))}$$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{x : p \vdash x : p}\,\mathrm{Rf} \qquad \dfrac{}{y : q \vdash y : q}\,\mathrm{Rf}}{x : p, y : q \vdash \langle x, y\rangle : p \wedge q}\,{\wedge}\mathrm{R}}{x : p, z : q \wedge r \vdash \langle x, fst(z)\rangle : p \wedge q}\,{\wedge}\mathrm{L}}{u : s \wedge p, z : q \wedge r \vdash \langle snd(u), fst(z)\rangle : p \wedge q}\,{\wedge}\mathrm{L}}{u : s \wedge p \vdash \lambda z.\, \langle snd(u), fst(z)\rangle : (q \wedge r) \supset (p \wedge q)}\,{\supset}\mathrm{R}}{\vdash \lambda u.\lambda z.\, \langle snd(u), fst(z)\rangle : (s \wedge p) \supset ((q \wedge r) \supset (p \wedge q))}\,{\supset}\mathrm{R}$$

$$\mathrm{SC}_{2 \vdash (s \wedge p) \supset ((q \wedge r) \supset (p \wedge q))}$$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{x : p \vdash x : p}\,\mathrm{Rf} \qquad \dfrac{}{y : q \vdash y : q}\,\mathrm{Rf}}{x : p, y : q \vdash \langle x, y\rangle : p \wedge q}\,{\wedge}\mathrm{R}}{u : s \wedge p, y : q \vdash \langle snd(u), y\rangle : p \wedge q}\,{\wedge}\mathrm{L}}{u : s \wedge p, z : q \wedge r \vdash \langle snd(u), fst(z)\rangle : p \wedge q}\,{\wedge}\mathrm{L}}{u : s \wedge p \vdash \lambda z.\, \langle snd(u), fst(z)\rangle : (q \wedge r) \supset (p \wedge q)}\,{\supset}\mathrm{R}}{\vdash \lambda u.\lambda z.\, \langle snd(u), fst(z)\rangle : (s \wedge p) \supset ((q \wedge r) \supset (p \wedge q))}\,{\supset}\mathrm{R}$$

The senses of these derivations would be the following:

Sense of $\mathrm{SC}_1$:

$$\{x, y, z, u, \langle x, y\rangle, \underline{\langle x, fst(z)\rangle}, \langle snd(u), fst(z)\rangle, \lambda z.\, \langle snd(u), fst(z)\rangle, \lambda u.\lambda z.\, \langle snd(u), fst(z)\rangle\}$$

Sense of $\mathrm{SC}_2$:

$$\{x, y, z, u, \langle x, y\rangle, \underline{\langle snd(u), y\rangle}, \langle snd(u), fst(z)\rangle, \lambda z.\, \langle snd(u), fst(z)\rangle, \lambda u.\lambda z.\, \langle snd(u), fst(z)\rangle\}$$

The two sets only differ with regard to the underlined terms, otherwise they are identical. Thus, they only differ in the order in which the two left conjunction rules are applied. For the resulting end-term this is inessential, but we can see that when taking the sense, and not only the end-terms, i.e., the denotation, into account, it is indeed possible to read off the structure of the derivations. As noted above (examples on p. 17), the term annotation of the calculi makes this structure of derivations explicit so that we can differentiate between derivations which would otherwise look identical. As several authors point out,[36] this is a desirable feature

cations of the ∧L-rule.

[36] See (Sørensen & Urzyczyn, 2006, p. 82), (Pfenning, 2000, p. 93) and for the latter point

if one is not only interested in mere provability but wants to study the structure of the derivations in question and also, for simplicity, if one wants to compare proof systems of ND and SC with each other. Since we are interested in both of these points, it seems the right choice for our purposes to consider the annotated versions of the calculi and that is also why these annotated versions are indeed needed for our notions of sense and denotation. Of course, one could argue that the underlying structure is still the same in the non-annotated versions and can be made explicit by other means, too, like showing the different generalizations of the derivations, but still, we do not see how in these calculi our notions could be easily applied.

Another issue that needs to be considered is the one of *identity of senses*, i.e., *synonymy*. Therefore, we want to extend our definition of sense given above with an addition:

> If a sense-representing set can be obtained from another by uniformly replacing (respecting the usual capture-avoiding conventions) any occurrence of a variable, bound or free, by another variable of the same type, they express the same sense.

What we ensure with this point is just that it does not (and should not) matter which variables one chooses for which proposition as long as one does it consistently. So, it does not make a difference whether we have

$\mathrm{ND}_{1 p \supset (q \supset p)}$ Sense$_1$: $\{x, \lambda z.x, \lambda x.\lambda z.x\}$

$$\dfrac{\dfrac{[x : p]}{\lambda z.x : q \supset p}\,{\supset}\mathrm{I}}{\lambda x.\lambda z.x : p \supset (q \supset p)}\,{\supset}\mathrm{I}$$

or

$\mathrm{ND}_{2 p \supset (q \supset p)}$ Sense$_2$: $\{y, \lambda z.y, \lambda y.\lambda z.y\}$

$$\dfrac{\dfrac{[y : p]}{\lambda z.y : q \supset p}\,{\supset}\mathrm{I}}{\lambda y.\lambda z.y : p \supset (q \supset p)}\,{\supset}\mathrm{I}$$

Sense$_1$ and Sense$_2$ represent the same sense. Or to give another example (pointed to by one of the anonymous referees) where we have free variables occurring within the derivation but not appearing in the end-term: If one would replace all occurrences of the free variable $y$ by the variable $w$ in derivation $\mathrm{SC}_{1 \vdash (s \wedge p) \supset ((q \wedge r) \supset (p \wedge q))}$ (see above), then this would make no difference to the sense according to our definition since the sense-representing sets would be obtained from replacing $y$ by $w$.

(Troelstra & Schwichtenberg, 2000, p. 73).

This also fits the Fregean criterion of two sentences' identical sense, as Sundholm (1994, p. 304) depicts it within a broader analysis: two propositions express the same sense if it is not possible to hold different *epistemic attitudes* towards them, i.e., "if one holds the one true, one also *must* hold the other one true, and *vice versa*". Whereas, if we have two sentences which only differ in two singular terms, referring to the same object but differing in sense, we can easily hold the one sentence to be true, while thinking the other is false, if we do not know that they are referring to the same object. With proofs it is the same: Looking at $\mathrm{ND}_{1p \supset (q \supset p)}$ and $\mathrm{ND}_{2p \supset (q \supset p)}$ we may not know whether the derivation is valid or not, we *do* know, however, that if one is a valid derivation, then so is the other. With derivations differing in sense this is not so straightforward.

For Frege this point of considering cases where intensionality is directed towards sentences was crucial to develop his notion of sense, so the question arises how we can explain cases of intensionality directed towards *proofs* with our notions of sense and denotation. Let us suppose we have two denotationally-identical proofs which are represented by two different derivations $\mathscr{D}$ and $\mathscr{D}'$. In this case it could happen that a (rational) person believes that derivation $\mathscr{D}$ is valid but does not believe that derivation $\mathscr{D}'$ is valid. How can we account for that? One explanation would be of course to point to the difference in linguistic representation. After all, it can just be the case that one way of writing down a proof is more accessible to the person than another (they may not be familiar with a certain proof system, for example). This would amount to letting the linguistic representation, the signs, collapse with the sense of a derivation. However, then we would have no means to distinguish this case from cases in which we want to argue that it is not justified for a rational person to have different propositional attitudes towards propositions which are about derivations differing insignificantly from each other, like in the cases of $\mathrm{ND}_{1p \supset (q \supset p)}$ and $\mathrm{ND}_{2p \supset (q \supset p)}$ above. For Frege (Frege, 1948[1892], p. 212, 218) the referent of an expression in an intensional context is not its *customary referent*, i.e., the object it refers to or the truth value in the case of sentences, but its *customary sense*. Here the situation is the same: What is referred to in such a setting, when speaking about the attitudes of a person towards propositions about derivations, is not the proof objects (which are identical in our situation) but their senses, which are in this context represented by the sets of terms encoding the steps of construction. It seems plausible, then, to say that when the construction steps differ in two derivations, a person can have different attitudes towards propositions about them, because the different construction steps may lead to this person grasping the one derivation, while not understanding the other.

## 2.4 Analogy to Frege's cases

Let us finally compare how our conception of sense and denotation in the context of proofs fits the distinction Frege came up with for singular terms and sentences. We can have the following two cases with Frege's distinction: firstly, there can be different signs corresponding to exactly one sense (and then of course also only one denotation) (Frege, 1948[1892], p. 211). In the case of singular terms an example would be "Gottlob's brother" and "the brother of Gottlob". The sense, the way the denoted individual object is given to us, is the same because there is only a minor grammatical difference between the two expressions. More frequently, this occurs in comparing different languages, though, taking singular terms which express exactly the same sense only using different words, like "the capital of France" and "die Hauptstadt Frankreichs". In the case of sentences an example would be changing from an active to a passive construction without changing the emphasis of the sentence; an example from Frege is the following: "$M$ gave document $A$ to $N$", "Document $A$ was given to $N$ by $M$" (Frege, 1979, p. 141). In the case of proofs, finally, an example would be the following case:

$\mathrm{ND}_{(p \vee p) \supset (p \wedge p)}$

$$\dfrac{\dfrac{\dfrac{[y:p\vee p]^3 \quad [x:p]^1 \quad [x:p]^1}{\texttt{case y } \{x.x \mid x.x\} : p}\,{\scriptstyle \vee E^1} \qquad \dfrac{[y:p\vee p]^3 \quad [x:p]^2 \quad [x:p]^2}{\texttt{case y } \{x.x \mid x.x\} : p}\,{\scriptstyle \vee E^2}}{\langle \texttt{case y } \{x.x \mid x.x\}, \texttt{case y } \{x.x \mid x.x\}\rangle : p \wedge p}\,{\scriptstyle \wedge I}}{\lambda y.\, \langle \texttt{case y } \{x.x \mid x.x\}, \texttt{case y } \{x.x \mid x.x\}\rangle : (p \vee p) \supset (p \wedge p)}\,{\scriptstyle \supset I^3}$$

$\mathrm{SC}_{\vdash (p \vee p) \supset (p \wedge p)}$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{x:p \vdash x:p}\,{\scriptstyle Rf} \quad \dfrac{}{x:p \vdash x:p}\,{\scriptstyle Rf}}{y:p\vee p \vdash \texttt{case y } \{x.x \mid x.x\} : p}\,{\scriptstyle \vee L} \qquad \dfrac{\dfrac{}{x:p \vdash x:p}\,{\scriptstyle Rf} \quad \dfrac{}{x:p \vdash x:p}\,{\scriptstyle Rf}}{y:p\vee p \vdash \texttt{case y } \{x.x \mid x.x\} : p}\,{\scriptstyle \vee L}}{y:p\vee p, y:p\vee p \vdash \langle \texttt{case y } \{x.x \mid x.x\}, \texttt{case y } \{x.x \mid x.x\}\rangle : p \wedge p}\,{\scriptstyle \wedge R}}{y:p\vee p \vdash \langle \texttt{case y } \{x.x \mid x.x\}, \texttt{case y } \{x.x \mid x.x\}\rangle : p \wedge p}\,{\scriptstyle C}}{\vdash \lambda y.\, \langle \texttt{case y } \{x.x \mid x.x\}, \texttt{case y } \{x.x \mid x.x\}\rangle : (p \vee p) \supset (p \wedge p)}\,{\scriptstyle \supset R}$$

Sense:

$$\{x, y, \texttt{case y } \{x.x \mid x.x\}, \langle \texttt{case y } \{x.x \mid x.x\}, \texttt{case y } \{x.x \mid x.x\}\rangle\,, \\ \lambda y.\, \langle \texttt{case y } \{x.x \mid x.x\}, \texttt{case y } \{x.x \mid x.x\}\rangle\}$$

Or to give another example:

$\mathrm{ND}_{p \supset (p \supset (p \wedge p))}$

$$\dfrac{\dfrac{\dfrac{[x:p]^2 \quad [y:p]^1}{\langle x, y\rangle : p \wedge p}\,{\wedge}\mathrm{I}}{\lambda y.\,\langle x, y\rangle : p \supset (p \wedge p)}\,{\supset}\mathrm{I}^1}{\lambda x.\lambda y.\,\langle x, y\rangle : p \supset (p \supset (p \wedge p))}\,{\supset}\mathrm{I}^2$$

$\mathrm{SC}_{\vdash p \supset (p \supset (p \wedge p))}$

$$\dfrac{\dfrac{\dfrac{\dfrac{}{x:p \vdash x:p}\,\mathrm{Rf} \qquad \dfrac{}{y:p \vdash y:p}\,\mathrm{Rf}}{x:p, y:p \vdash \langle x, y\rangle : p \wedge p}\,{\wedge}\mathrm{R}}{x:p \vdash \lambda y.\,\langle x, y\rangle : p \supset (p \wedge p)}\,{\supset}\mathrm{R}}{\vdash \lambda x.\lambda y.\,\langle x, y\rangle : p \supset (p \supset (p \wedge p))}\,{\supset}\mathrm{R}$$

Sense: $\{x, y, \langle x, y\rangle, \lambda y.\,\langle x, y\rangle, \lambda x.\lambda y.\,\langle x, y\rangle\}$

In these cases derivations can consist of different signs, namely by having one representation in SC and one in ND, which do not differ in sense nor in denotation, since they both contain exactly the same terms and produce the same end-term. This comparison between different proof systems seems to fit nicely with Frege's (Frege, 1948[1892], p. 211) comment on "the same sense ha[ving] different expressions in different languages". However, as we have seen above with the examples $\mathrm{ND}_{1p \supset (q \supset p)}$ and $\mathrm{ND}_{2p \supset (q \supset p)}$, this case can also occur within the same proof system. One could wonder whether there should not be a differentiation between the senses of the derivations in the first example since it seems that different rules are applied: in $\mathrm{SC}_{\vdash (p \vee p) \supset (p \wedge p)}$ we have an application of contraction, which we do not have in $\mathrm{ND}_{(p \vee p) \supset (p \wedge p)}$. This would also question whether our definition of sense distinguishes and identifies the right amount of cases. We do believe that this is the case, though, because in the first example, where there is an application of the contraction rule in SC, there is also a multiple assumption discharge in the ND-derivation, which is generally seen as the corresponding procedure, just as cases of vacuous discharge of assumptions in ND correspond to the application of weakening in SC. So, just as in different languages of course not exactly the same expressions are used, here too, the rules differ from ND to SC but since the *corresponding* procedures are used, one can argue that the sense does not differ for that reason.

Another case that can occur according to Frege (ibid.) is that we have one denotation, i.e., one object a sign refers to, but different senses. An example for this would be his famous "morning star" and "evening star" comparison, where both expressions refer to the same object, the planet Venus, but the denoted object is given differently. On the sentence level this would amount to exchanging singular terms in a sentence by ones which have the same denotation: "The morning star is the planet Venus" and "The evening star is the planet Venus". The denotation of the sentence - with Frege: its truth value - thus stays the same, only the sense of it differs, the information is conveyed differently to us. For our proof cases we can say that this case is given when we have syntactically different derivations, be it in one or in different proof systems, which have end-terms belonging to the same equivalence class induced by the set of $\alpha$-, $\beta$- and $\eta$-conversions. Thus, examples would be corresponding proofs in ND and SC, which share the same end-term, but contain different terms occurring within the derivations. The reason for this to happen seems that in SC often more variables are necessary than in ND. If we compare derivations *within ND*, one definite case in which we have the same denotation but a different sense is between equivalent but syntactically distinct derivations, e.g., non-normal and normal derivations, one reducible to the other. Another case up for debate would be the one with rule permutations due to disjunction elimination. *Within SC* we can have two cases: one due to rule permutation, one due to applications of cut. For the first case, where the inference could be given in a different way, although ending on the same term, we gave examples above. However, it is worth mentioning that our distinction still captures the usual distinction, the second case, where it is said that two derivations, one containing cut and the other one in cut-free form (as a result of cut-elimination applied to the former), have the same denotation but differ in sense:

$\mathrm{SC}_{\vdash\,(p\,\wedge\,p)\,\supset\,(p\,\vee\,p)}$

$$
\begin{prooftree}
\AxiomC{}
\RightLabel{Rf}
\UnaryInfC{$z : p \vdash z : p$}
\RightLabel{W}
\UnaryInfC{$z : p, x : p \vdash z : p$}
\RightLabel{$\wedge$L}
\UnaryInfC{$y : p \wedge p \vdash fst(y) : p$}
\RightLabel{$\vee$R}
\UnaryInfC{$y : p \wedge p \vdash \mathtt{inl}fst(y) : p \vee p$}
\RightLabel{$\supset$R}
\UnaryInfC{$\vdash \lambda y.\mathtt{inl}fst(y) : (p \wedge p) \supset (p \vee p)$}
\end{prooftree}
$$

Sense: $\{z, x, y, fst(y), \mathtt{inl}fst(y), \lambda y.\mathtt{inl}fst(y)\}$

$\mathrm{SC}_{\mathrm{cut}\vdash\,(p\,\wedge\,p)\,\supset\,(p\,\vee\,p)}$

$$
\begin{prooftree}
\AxiomC{}
\RightLabel{Rf}
\UnaryInfC{$z : p \vdash z : p$}
\RightLabel{W}
\UnaryInfC{$z : p, x : p \vdash z : p$}
\RightLabel{$\wedge$L}
\UnaryInfC{$y : p \wedge p \vdash fst(y) : p$}
\AxiomC{}
\RightLabel{Rf}
\UnaryInfC{$z : p \vdash z : p$}
\RightLabel{$\vee$R}
\UnaryInfC{$z : p \vdash \mathtt{inl}z : p \vee p$}
\RightLabel{cut}
\BinaryInfC{$y : p \wedge p \vdash \mathtt{inl}fst(y) : p \vee p$}
\RightLabel{$\supset$R}
\UnaryInfC{$\vdash \lambda y.\mathtt{inl}fst(y) : (p \wedge p) \supset (p \vee p)$}
\end{prooftree}
$$

Sense: $\{z, x, y, fst(y), \underline{\mathtt{inl}z}, \mathtt{inl}fst(y), \lambda y.\mathtt{inl}fst(y)\}$

As mentioned above (fn 32), cut does not need to create a non-normal term, as it is the case here, but still any application of cut will necessarily change the sense of a derivation as opposed to its cut-free form.

Finally, cases that need to be avoided in a formal language according to Frege (Frege, 1948[1892], p. 211) would be to have one sign, corresponding to different senses, or on the other hand, one sense corresponding to different denotations. As he mentions, these cases of course occur in natural languages but should not happen in formal ones, so it should also not be possible in our present context, for sure. Fortunately, this cannot happen in the context of our annotated proof systems, either, since the signs (taken to be the derivation as it is written down) always express at most one sense in our annotated system, and likewise the sense always yields a unique denotation since the end-term is part of the sense-denoting set.[37]

## 2.5 Conclusion

The context in which Frege considered sense and denotation was the context of identity. Likewise, we argued in this paper, if we use term-annotated calculi, we can also say something about proof identity: identity of proofs over different calculi or within the same calculus consists in having end-terms that belong to the same equivalence class induced by the set of $\alpha$-, $\beta$- and $\eta$-conversions. In ND this can happen when we have the same proof in normal and non-normal form, in SC this can happen when we have the same proof using cut and in cut-free form but also when there are forms of rule permutations where an application of the $\wedge$L-rule or the $\supset$L-rule switches place with another rule. Including disjunction in our language creates for both calculi the additional question of whether rule permutations including disjunction elimination (resp. the left disjunction rule) lead to a different proof, or whether these proofs should be identified. We are more interested in sense, however, and here we can conclude that what in all these cases changes is the sense of the derivation in question. Finally, considering the question of identity of sense, i.e., synonymy, and trying to follow Frege's conception on this matter, too, we can say the following: if two derivations are supposed to be *identical in sense*, this means that the way the inference is given is essentially the same, so the set of terms building up the end-term must be the same. The end-term itself does not necessarily tell us anything about the structure of the proof. Sense, on the other hand, is more fine-grained in that the set of terms occurring within the derivation reflects how the derivation is built up. Especially in SC, where we can have different orders of rule applications leading up to the same end-term, the sense gives us means to distinguish on a more fine-grained level.

[37] Another question would be whether there can be signs without any sense at all. Frege (1948[1892], p. 211) dismisses this case, as well, with a remark that we need at least the requirement that our expressions are "grammatically well-formed". Tranchini (2016) gives a good analogy pointing to the notorious connective `tonk` playing this role in the case of proofs.

# 3 What are acceptable reductions?

## Perspectives from proof-theoretic semantics and type theory

### 3.1 Introduction

What are acceptable reductions[38] in the context of proofs and why is it important to distinguish these from ‘bad’ ones? As Schroeder-Heister and Tranchini (2017, p. 574) argue, from a philosophical point of view, or more specifically a standpoint of proof-theoretic semantics (Schroeder-Heister, 2022), reduction procedures are closely connected to the question about identity of proofs: If we take proofs to be abstract entities represented by (natural deduction) derivations, then derivations belonging to the same equivalence class induced by the reflexive, symmetric, and transitive closure of reducibility can be said to represent the same proof object.[39] As they show, accepting certain reductions, more specifically accepting the so-called *Ekman-reduction* (see below), would lead to a trivialization of identity of proofs in the sense that every derivation of the same conclusion would have to be identified. They suggest such a trivialization as a criterion to disallow reductions. I will argue that the question, which reductions we accept in our system, is not only important if we see them as generating a theory of proof identity but is also decisive for the more general question whether a proof has meaningful content, i.e., it does not only matter to the question about the *denotation* of proofs but also to the question about their *sense*. Therefore, we need to be careful: We cannot just accept any reduction, i.e., any procedure eliminating some kind of detour in a derivation.

An example of a reduction not belonging to the usual reductions is Ekman-reduction, as it is presented in (Ekman, 1994, 1998) and extensively discussed by Schroeder-Heister and Tranchini (2017; 2018) and Tennant (2021):

$$\dfrac{B \to A \qquad \dfrac{A \to B \qquad \begin{array}{c}\mathscr{D}\\ \vdots \\ A\end{array}}{B}\,{\scriptstyle \to E}}{A}\,{\scriptstyle \to E} \qquad \rightsquigarrow \qquad \begin{array}{c}\mathscr{D}\\ \vdots \\ A\end{array}$$

What we want in light of such a non-standard kind of reduction are criteria determining which reductions can be allowed in our system and which should be dismissed. It is advantageous for such an approach to exploit the so-called *Curry-Howard-correspondence* (see, e.g., Sørensen & Urzyczyn, 2006) and examine proof systems annotated with $\lambda$-terms. These make the structure of our derivations ex-

[38]Note that I am focusing strictly on *reductions* in this paper, i.e., procedures that cut out what is in some way considered a detour of a derivation, not conversions in general, like expansions or permutations. The latter are certainly also of great interest for proof-theoretic semantics but would extend the scope of this paper.

[39]This goes back to Prawitz (1971, pp. 257-261), who credits the idea to Per Martin-Löf; many others have defended this view since.

plicit and facilitate to show what is wrong with potential reductions and why they should not be admitted in our system. The question, then, shifts to asking which reduction procedures for *terms* can be allowed. The $\lambda$-calculus and some well-known properties thereof can provide us with directions as to what could be (un)desirable features of reductions.

Annotating Ekman-reduction with terms then shows something else - besides Schroeder-Heister and Tranchini's point - that is essentially problematic about this reduction. Indeed, I want to show that allowing it would be equal to allowing a reduction for `tonk`, i.e., a reduction for a derivation consisting of a `tonk`-introduction rule followed by its elimination rule (see below, Section 3.3.1). It is generally agreed upon, though, that there cannot be a sensible reduction for this connective. By creating a `tonk`-reduction in the same fashion as it would be done for other connectives, the consequences of allowing this reduction are made explicit and it can be shown that those would be the same for Ekman-reduction: Not only would these reductions induce an equivalence relation relating different terms in normal form of the *same* conclusion, they would also allow to reduce a term of one type to a term of an *arbitrary* other. If we take reductions as generating identity between proofs, then that would also force us to identify proofs of arbitrarily different formulas. But even if we reject this assumption (some researchers do not find this theory of proof identity very compelling, see Section 3.4), I will argue that allowing such reductions would render derivations in such a system *meaningless*.

## 3.2 Reduction procedures in natural deduction and $\lambda$-calculus

The reductions for the connectives of, say, minimal propositional logic, corresponding to $\beta$-reductions in $\lambda$-calculus, are meant to eliminate unnecessary detours of the following form: There is a formula, called *maximal formula*, which is both the conclusion of an application of an introduction rule of a connective as well as the major premise of an applied elimination rule governing the same connective. It can be shown for those connectives (see below for $\rightarrow$) that in these cases the maximal formula (below $A \rightarrow B$) can be eliminated without losing anything essential because, as Prawitz (1971, p. 251) argues, this procedure is just a way to make the *inversion principle*[40] explicit. It can and has been argued, however, that there are more reductions than the ones for the connectives that are usually considered (see, e.g., Tennant, 1995). One of those, presented in (Ekman, 1994, 1998), will be discussed in this paper.

[40] The principle, to which (at least) the rules governing the connectives of minimal logic adhere, saying that nothing new is obtained by an elimination immediately following an introduction of the major premise of the elimination rule (Prawitz, 1971, p. 246).

$$\dfrac{\dfrac{\begin{array}{c}[A]\\ \vdots\ \mathscr{D}\\ B\end{array}}{A \to B}{\to}I \qquad \begin{array}{c}\mathscr{D}'\\ \vdots\\ A\end{array}}{B}{\to}E \qquad \rightsquigarrow \qquad \begin{array}{c}\mathscr{D}'\\ \vdots\\ [A]\\ \vdots\ \mathscr{D}\\ B\end{array}$$

As mentioned above, in using a term-annotated proof system we are implementing the Curry-Howard correspondence, which takes the view of proofs as programs and formulas as types as a basis and which states a close correspondence between these notions in the simply typed $\lambda$-calculus and natural deduction (ND) systems of intuitionistic logic. For our purposes here it suffices to consider the $\to$-fragment of intuitionistic logic and correspondingly the system $\lambda^{\to}$. We use $\rho$, $\sigma$, $\tau$,... for arbitrary atomic formulas, $A$, $B$, $C$,... for arbitrary formulas, and $\Gamma$, $\Delta$,... for sets of formulas. The concatenation $\Gamma, A$ stands for $\Gamma \cup \{A\}$. For term variables, $x$, $y$, $z$,... are used and $r$, $s$, $t$,... for arbitrary terms. Furthermore, we use '$\equiv$' to denote *syntactic identity* between terms, types, or derivations. The following are our term-annotated ND-rules with the corresponding $\beta$-reduction:

$$\dfrac{\begin{array}{c}[x : A]\\ \vdots\\ t : B\end{array}}{\lambda x.t : A \to B}{\to}I \qquad \dfrac{\begin{array}{c}\Gamma\\ \vdots\\ s : A \to B\end{array} \quad \begin{array}{c}\Delta\\ \vdots\\ t : A\end{array}}{App(s,t) : B}{\to}E \qquad App(\lambda x.t, s) \rightsquigarrow_{\beta} t[s/x]$$

We read $t : A$ as "term $t$ is of type $A$" or, in the 'proof-reading', "$t$ is a proof of formula $A$". Such an expression is also called a *type assignment statement* with the $\lambda$-term being the *subject* and the type the *predicate*. Thus, we use a type-system *à la Curry* here, in which the terms are not typed, in the sense that the types are part of the term's structure, but are *assigned* types according to type assignment rules, which in our case are simply the rules above. With $t : A$ we express that $A$ is the *principal type* of term $t$, i.e., the most general type that can be assigned to $t$.[41] Substitution is expressed by $t[s/x]$, meaning that in term $t$ every free occurrence of $x$ is substituted with $s$. The usual capture-avoiding requirements for variable substitution are to be observed. I will follow standard terminology of type theory here and call a term of the form $App(\lambda x.t, s)$ a *$\beta$-redex* and the corresponding term $t[s/x]$ its *contractum*.[42] Replacing an occurrence of a $\beta$-redex contained in a term by its contractum is called a *$\beta$-contraction* and if there is a finite (possibly empty) series of $\beta$-contractions changing term $t$ to $t'$, we say that $t$ *$\beta$-reduces* to $t'$ and write $t \rightsquigarrow_{\beta} t'$. The reduction relation $\rightsquigarrow_{\beta}$ is reflexive and transitive and closed under $\alpha$-conversion, i.e., renaming of bound variables. A term that contains no $\beta$-redexes

[41] For example, for the term $\lambda x.x$, its types could be $p \to p$, $q \to q$, $(p \to q) \to (p \to q)$, etc., while its principal type would be $A \to A$.

[42] For the following definitions (with only slightly differing formulations and notations), see (Barendregt, 1992), (Girard, 1989), and (Hindley & Seldin, 2008). I will use the same terminology (without the '$\beta$-') for any reduction procedures, whether or not they will be found acceptable in the course of the paper. For reduction procedures in general, see also (Baader & Nipkow, 1998).

is said to be *$\beta$-normal* or in *$\beta$-normal form ($\beta$-nf)* and $t'$ is the $\beta$-nf of $t$ if $t \rightsquigarrow_\beta t'$ and $t'$ is $\beta$-normal.

In general (if we consider more connectives than $\rightarrow$), the correspondence to the introduction and elimination rules in $\lambda$-calculus is that each connective has its own *constructor*, an operator constructing canonical objects of particular types, and a *destructor*, specifying the use of these objects in computations. A $\beta$-redex consists of a destructor applied to the constructor of the same connective, so Curry-Howard correspondence always gives us an analogy between a $\beta$-redex and a proof detour consisting of an elimination immediately following the introduction of the same connective (Sørensen & Urzyczyn, 2006, p. 87f.).

It is important to stress that reduction procedures can be interpreted in two different (yet certainly closely related) ways (see, e.g., Barendregt, 1992, Ch. 2.3; Girard, 1989, pp. 18-20; Hindley & Seldin, 2008, pp. 11-18). One interpretation is to see reductions as inducing an identity relation, i.e., on this view, the relation applies equally in both directions. We will speak of *$\beta$-equality* of terms in these cases and use $=_\beta$ to express this relation. It is just like, e.g., $App(\lambda x.x^2 + 7, 2) \rightsquigarrow_\beta 11$ expresses the fact that 4+7=11.

Another interpretation of reductions is to see them as *directed* computations, calculations, or executions corresponding to the idea of program evaluation. On this view, the asymmetry between redex and contractum must be stressed: In our example '4+7' can be interpreted as '11' by doing a calculation. A reduction procedure is seen as an evaluation that is run on a term and thereby interprets this term in a different way. The non-symmetric $\beta$-reducibility relation implies the symmetric relation of $\beta$-equality but not the other way around (Hindley & Seldin, 2008, p. 16). Hence, if $t \rightsquigarrow_\beta t'$, then $t =_\beta t'$; but not: if $t =_\beta t'$, then $t \rightsquigarrow_\beta t'$. Just like 4+7 evaluates to 11 but not the other way around: 11 is fully evaluated; it is already in normal form, i.e., we do not *reduce* it to 4+7.

One of the most important results in $\lambda$-calculus, which will also be important for this paper, is the so-called *Church-Rosser Theorem* stating the *confluence property* for $\beta$-reduction:

**Church-Rosser Theorem:** If a term can be reduced to two syntactically different terms, then there is a term to which these two can be reduced. Put formally, if $t \rightsquigarrow_\beta t'$ and $t \rightsquigarrow_\beta t''$, then there is a term $s$ such that $t' \rightsquigarrow_\beta s$ and $t'' \rightsquigarrow_\beta s$.

Likewise, this property holds for $\beta$-equality: if two syntactically different terms are $\beta$-equal, then there is a term to which they both can be reduced in finitely many steps, i.e., if $t =_\beta t'$ and $t \not\equiv t'$, then there is a term $s$ such that $t \rightsquigarrow_\beta s$ and $t' \rightsquigarrow_\beta s$. One corollary of this is the uniqueness of $\beta$-normal forms for terms (provided they have a normal form). Another important corollary for our purposes is that two terms $t$ and $t'$ that are in $\beta$-nf *and* syntactically distinct cannot be $\beta$-equal, which means that the relation $=_\beta$ is non-trivial: not all terms are $\beta$-equal (Hindley & Seldin, 2008, p. 17).[43]

## 3.3 What distinguishes 'good' from 'bad' reductions?

### 3.3.1 Problematic reductions

**`Tonk`-reduction** The reduction procedure considered above for $\to$ for eliminating maximal formulas, that arise from applying an elimination rule immediately after the corresponding introduction rule, works equally well for our other 'well-behaved' connectives (Prawitz, 1965). A comparison with the notorious connective `tonk` might help to see, however, why this is not the case for every connective. `Tonk` was introduced by Prior (1960) as an ad absurdum-attack on the idea of proof-theoretic semantics[44]: If it was only the rules giving the meaning of a connective, no other metaphysically underlying concept, then what would stop the proof-theoretic semanticist from accepting the following rules?

$$\frac{A}{A \texttt{ tonk } B}\,\texttt{tonk}I \qquad\qquad \frac{A \texttt{ tonk } B}{B}\,\texttt{tonk}E$$

Applying these immediately after each other gives us a derivation from arbitrary $A$ to arbitrary $B$, i.e., our system would trivialize. Additionally, there is no real way to make out a reasonable reduction procedure in this case, which is, of course, due to the fact that `tonk` violates the inversion principle. This has been one of the ways to give a reason why `tonk` can be considered inadmissible.[45] It should be noted that there are approaches to `tonk`, which do not even consider it inadmissible in principle but which rather question our underlying assumptions about logical consequence on the grounds of which we dismiss `tonk`.[46] Yet, I want to emphasize that even with an argumentation that accepts `tonk`, to my knowledge, there is still no way of giving an acceptable reduction procedure for this connective.

Can this be made explicit with term annotations? Leaving $\lambda$-calculus we can still give term-annotated rules and a corresponding reduction for non-standard connectives, like for a Liar-connective $L$ for example, as it has been proposed in (Schroeder-Heister, 2012b):

[43] This also implies the consistency of the simply typed $\lambda$-calculus (Barendregt, 1992, Ch. 2.3; Girard, 1989, p. 23).

[44] Although this specific term has been introduced only later in 1991 by Schroeder-Heister, as he mentions in (2022), the general idea has been prevalent much longer.

[45] Since the `tonk`-rules do not adhere to the inversion principle, they are not in *harmony*. On this notion as a criterion for acceptable connectives, see, e.g., (Dummett, 1991; Francez & Dyckhoff, 2012; Read, 2010; Tennant, 1978; Tranchini, 2015).

[46] Cook (2005) and Ripley (2015), e.g., argue like this in claiming that if we do not assume a transitive consequence relation, then an extension with `tonk` would not yield inconsistency. See also Wansing (2006), however, who shows that the problems of `tonk` avoided in a non-transitive system can be recreated by other `tonk`-like connectives.

$$\frac{t : L \to \bot}{lt : L}\,\text{L}I \qquad \frac{t : L}{l't : L \to \bot}\,\text{L}E \qquad l'lt \rightsquigarrow_{\text{L}} t$$

We have $l$ here serving as a constructor for the introduction rule and $l'$ as a destructor in the elimination rule. We can do the same for `tonk`, annotating the rules with a constructor $k$ and a destructor $k'$:

$$\frac{t : A}{kt : A \ \texttt{tonk}\ B}\,\texttt{tonk}I \qquad \frac{t : A \ \texttt{tonk}\ B}{k't : B}\,\texttt{tonk}E$$

Just like for $L$ a non-normal term for `tonk` would then be constructed by applying the destructor to the constructor, which is, as for the usual connectives, the result of a derivation containing the conclusion of the introduction rule as the major premise of the elimination rule, i.e.:

$$\dfrac{\dfrac{t : A}{kt : A \ \texttt{tonk}\ B}\,\texttt{tonk}I}{k'kt : B}\,\texttt{tonk}E$$

The usual reduction would be to reduce the term for the conclusion of the elimination rule to the one of the premise of the introduction rule, so analogous to the Liar-reduction: $k'kt \rightsquigarrow_{\texttt{tonk}} t$. However, $t$ is assigned type $A$, while $k'kt$ is assigned $B$. So, if we would accept this reduction, it would mean to accept a reduction relating terms of arbitrarily different types. In the following I want to show that what is wrong with Ekman-reduction is essentially the same as in the case of `tonk`-reduction and on this basis identify what could be a good criterion for reductions of proofs in terms of type theory.

**Ekman-reduction** Again, Ekman-reduction has the following form:

$$\dfrac{B \to A \qquad \dfrac{A \to B \qquad \begin{matrix}\mathscr{D}\\ \vdots \\ A\end{matrix}}{B}\,{\to}E}{A}\,{\to}E \qquad \rightsquigarrow \qquad \begin{matrix}\mathscr{D}\\ \vdots \\ A\end{matrix}$$

The motivation for Ekman to consider this reduction was to give a counterexample to Tennant's (1982) proof-theoretic characterization of paradoxes. According to this, a paradoxical derivation is one that yields a non-normalizable derivation of $\bot$. Tennant considers several examples, like versions of the Liar paradox, Curry's paradox or Russell's paradox, which all have this feature in common and of course, contain some special rules for the respective paradoxical connectives. Ekman gave an example of a derivation of $\bot$, though, not containing any other rules than the usual ones for implication but which still, if we accept Ekman-reduction that is, could not be brought into normal form because as with Tennant's examples the reduction sequences are looping. Thus, he concluded that Tennant's criterion does not capture a genuinely *paradoxical* feature of the derivations considered, since with

Ekman-reduction we could get such a derivation, as well, without containing any paradoxical elements. There have been attempts to show that this can be avoided by using a different representation of the rules, e.g., in (von Plato, 2000) by using general elimination rules in ND showing that such a derivation can be brought to a normal form or in (Tennant, 2021) with rules in sequent calculus showing that we get a cut-free derivation in such a system. This does not stand in opposition to what I am focusing on in this paper, though. Note that these 'solutions' to the so-called *Ekman-paradox* do concede that Ekman-reduction is permissible, since only by using it, we get into this infinite loop of reduction sequences. If we reject Ekman-reduction for independent reasons, for which I will argue in this paper, then Ekman-paradox is no problem either.

The problem with this reduction, which Schroeder-Heister and Tranchini (2017) point out and neatly prove, is the following: if we allow Ekman-reduction (plus always assuming for now that proofs related via reductions can be identified), then we would be forced to identify every derivation of a formula with every other derivation of the same formula, i.e., there would be no basis to distinguish different derivations other than their obvious syntactic difference. We would have to commit to them all representing one and the same proof. It is on these grounds of proof identity that Schroeder-Heister and Tranchini argue that Ekman-reduction should not be counted as an acceptable reduction. I want to show now that annotating Ekman-reduction with terms shows something else that is essentially problematic with this reduction.

In our term-annotated system the derivation to which Ekman-reduction is applied is the following:[47]

$$\dfrac{y : B \to A \qquad \dfrac{x : A \to B \qquad \begin{array}{c}\mathscr{D}\\ \vdots \\ t : A\end{array}}{App(x,t) : B}\,{\scriptstyle \to E}}{App(y, App(x,t)) : A}\,{\scriptstyle \to E}$$

So the Ekman-reduction procedure for terms would be:

**Ekman-reduction:** $App(y, App(x,t)) \rightsquigarrow_{Ekman} t$

In the specific case above, this reduction seems fine. However, the problem with it, as opposed to the known $\beta$-reductions, is that it is too unspecific concerning the term structure. With the same terms (since in Curry-style the types are not part of the syntactical structure of the terms) the following derivation could be constructed:[48]

[47] Note that Schroeder-Heister and Tranchini also consider a more general form of this reduction in their paper in that $A \to B$ and $B \to A$ are not assumptions but are derived formulas. Since this would not change the results here, I will stick to the original form, though.

[48] Still, there are reasons why Curry-style typing is preferable to Church-style, see Section 3.5.

$$\dfrac{y : B \to A \qquad \dfrac{x : (A \to A) \to B \qquad \begin{matrix}\mathscr{D} \\ \vdots \\ t : A \to A\end{matrix}}{App(x,t) : B}\ {\to}E}{App(y, App(x,t)) : A}\ {\to}E$$

The derivation is fine but the reduction would be problematic. In this case it is clear that $A \not\equiv A \to A$, i.e., it cannot be the case that the term for $A$ and the term for $A \to A$ constructed out of the same type context in $\mathscr{D}$ are syntactically the same. Using a concrete example, we can show why allowing this reduction can create a problem. Consider the following derivation:

$$\dfrac{y : \tau \to \rho \qquad \dfrac{x : (\sigma \to \sigma) \to \tau \qquad \dfrac{[z : \sigma]}{\lambda z.z : \sigma \to \sigma}\ {\to}I}{App(x, \lambda z.z) : \tau}\ {\to}E}{App(y, App(x, \lambda z.z)) : \rho}\ {\to}E$$

So, $App(y, App(x, \lambda z.z))$ would Ekman-reduce to $\lambda z.z$. However, no type assigned to $\lambda z.z$ can be atomic, i.e., $\lambda z.z : \rho$ is impossible. The problems arising for these reductions relate to questions of so-called *type preservation*, *typechecking*, and *type reconstruction*, which I will discuss in the next section.

### 3.3.2 Subject reduction and type reconstruction

The problem of `tonk`- and Ekman-reduction seems to be that, unlike the $\beta$-reductions, they are not type preserving. Let us briefly take a look at this property and its significance for reduction procedures. Sometimes the expressions *subject reduction* and *type preservation* are used synonymously. However, type preservation describes a broader concept than subject reduction, since the latter only says that types are preserved when terms (i.e., "subjects") are *reduced*, whereas type preservation can also be used to describe a property of subject *expansions*. So, we will distinguish this terminology here.[49] The subject reduction theorem for the proof system with $\lambda$-terms we consider states the following (Sørensen & Urzyczyn, 2006, p. 59):

**Subject Reduction Theorem:** If $\Gamma \vdash t : A$ and $t \rightsquigarrow_\beta t'$, then $\Gamma \vdash t' : A$.

Subject expansion, on the other hand, does not hold for this system in general, i.e., it is not the case that if $t \rightsquigarrow_\beta t'$ and $t' : A$, then $t : A$, meaning that the set of types assigned to a term is not invariant under conversion in general.[50]

The examples given in the previous section clearly show that subject reduction does not hold for `tonk`- and Ekman-reduction, i.e., it is not the case that whenever $t : A$ and $t \rightsquigarrow_{Ekman/\texttt{tonk}} t'$, then $t' : A$. We can also say that the contractum does not

[49] If in the following a reduction is stated (not) to be type preserving, this means that it enjoys (no) subject reduction, i.e., reductions are to be understood in the one-directed sense without looking at the other direction of expansions.

[50] See, e.g., (Barendregt, 1992, p. 41); (Hindley & Seldin, 2008, p. 170) for counterexamples).

*typecheck* at every type the redex typechecks at. Typechecking is something that needs to be considered in Curry-style type systems (see, e.g., Sørensen & Urzyczyn, 2006, p. 60) and is about deciding whether or not $\Gamma \vdash t : A$ holds, for a given context $\Gamma$, a term $t$ and a type $A$. We can express typechecking in the following form, then:

**Typechecking:** $t$ *typechecks* at $A$ iff $\Gamma \vdash t : A$ holds, for a given context $\Gamma$, a term $t$ and a type $A$.

As can be seen above, there are cases with Ekman-reduction (and for `tonk` it is even more obvious) in which it is *impossible* to assign $t'$ the type assigned to $t$. If we understand types like labels telling us the combinations that can safely be made with a term, then we can understand subject reduction as saying that a term will not become 'less safe' during a reduction, i.e., when performing a computation on a term, this term cannot turn from a well-typed into an ill-typed one (Hindley & Seldin, 2008, p. 168). Subject reduction thus establishes the correctness of our system of type assignment (Sørensen & Urzyczyn, 2006, p. 59). It seems, therefore, that maintaining subject reduction would certainly be a desirable feature for reduction procedures.

So, is subject reduction a good criterion to measure the acceptability of reduction procedures? To deal with this question we need to determine whether non-type-preserving reductions necessarily lead to trivialization of the system. In other words, are there systems which can contain reduction procedures that are not type preserving but yet do not trivialize the reducibility relation? Although it looks like a promising criterion for reductions, it actually seems to be the case that failure of subject reduction need not necessarily cause trivialization. To wit, it does not seem impossible that there could be a type theory with a reduction that is not type preserving without relating terms of arbitrary types but, e.g., only of types which are equivalent (i.e., interderivable formulas).[51]

The actual problem with the `tonk`-/Ekman-reductions, though, leading to trivialization of the system, can be identified when looking at type reconstruction for their redexes. Type reconstruction is used to decide the typability of terms (Sørensen & Urzyczyn, 2006, p. 60):

**Type Reconstruction:** Given term $t$, decide if there is a context $\Gamma$ and a type $A$, such that $\Gamma \vdash t : A$.

[51] An anonymous reviewer pressed the point here that it seemed unnecessary to look for a weaker criterion than subject reduction since with it we do get rid of the problematic cases. However, this would only amount to give a sufficient criterion, not a necessary one. Indeed, if a reduction enjoys the property of subject reduction, then it will be deemed acceptable and the criterion I will propose here will not stand against that. But that does not exclude the possibility that there might be reductions that are acceptable without having this property (see Section 3.3.4).

This can be achieved using a type reconstruction algorithm, which is simply based on the type assignment rules that are used. Since these are just given by the annotated inference rules of our system, they are syntax-oriented. This means that we should be able to figure out the *principal types* of terms, i.e., figure out the derivation by reconstructing bottom-up the term using the type assignment rules. To give an example of a successful type reconstruction, let us consider the one for the redex $App(\lambda x.t, s)$ resulting from our $\to$-rules starting with assigning it an arbitrary type $B$. We write '?' whenever this part of the type is syntactically undetermined in this step of the reconstruction. Two occurrences of '?' in the same step mean that, although their structure is undetermined, they must be filled in by the same type symbol. In the next step we are to use a 'fresh' type symbol for '?'.

**Type reconstruction for $App(\lambda x.t, s)$:**

$$\dfrac{\lambda x.t :? \to B \qquad \begin{matrix}\mathscr{D}' \\ \vdots \\ s :?\end{matrix}}{App(\lambda x.t, s) : B}\,{\scriptstyle \to E} \qquad\qquad \dfrac{\dfrac{\begin{matrix}[x : A] \\ \vdots\ \mathscr{D} \\ t : B\end{matrix}}{\lambda x.t : A \to B}\,{\scriptstyle \to I} \qquad \begin{matrix}\mathscr{D}' \\ \vdots \\ s : A\end{matrix}}{App(\lambda x.t, s) : B}\,{\scriptstyle \to E}$$

As we can see, the type reconstruction proceeds in such a way that *we have to* assign contractum $t$ the same type as the redex. The structure of the redex and the connected type assignment rules lead to an exactly determined type reconstruction, which cannot 'go wrong' concerning the relation between types of redex and contractum.

### 3.3.3 Criterion for acceptable reductions

The problem with allowing a reduction such as the one for `tonk`, however, can be shown by a type reconstruction of the non-normal term $k'kt$, assuming $k'kt \leadsto_{\texttt{tonk}} t$ as a reduction, as motivated above. If we assign $k'kt$ an arbitrary type $B$, then the only information this gives us for $kt$ is that its type must be of the form "? `tonk` $B$". Consequently, $t$ can be assigned an arbitrary type. This means that the types of redex and contractum are arbitrarily independent of each other, which is exactly the core of the problem with a reduction for `tonk`.

**Type reconstruction for $k'kt$:**

$$\dfrac{kt :? \ \texttt{tonk}\ B}{k'kt : B}\,{\scriptstyle tonkE} \qquad\qquad \dfrac{\dfrac{t : A}{kt : A\ \texttt{tonk}\ B}\,{\scriptstyle tonkI}}{k'kt : B}\,{\scriptstyle tonkE}$$

With type reconstruction it also becomes evident that the same problem as with `tonk` prevails with Ekman-reduction. We are doing a type reconstruction for the redex again. If we assign $App(y, App(x, t))$ an arbitrary type $A$, then we can reconstruct bottom-up the following derivation in which a new type variable is used whenever it is independent from the ones already used (skipping the step-by-step illustration with '?'):

**Type reconstruction for** $App(y, App(x, t))$**:**

$$\dfrac{y : B \to A \qquad \dfrac{x : C \to B \qquad \overset{\mathscr{D}}{\vdots}\; t : C}{App(x,t) : B}\ {\to}E}{App(y, App(x,t)) : A}\ {\to}E$$

Again, such a reduction allows reducing a term of one type to one of an *arbitrary* other; one that is arbitrarily unrelated in the type reconstruction from the type of the term that is reduced.

This arbitrariness cannot arise with the standard $\beta$-reductions and, importantly, there are also other non-standard reductions which are well-behaved with respect to this feature, i.e., this is not simply to say that $\beta$-reductions are the only acceptable reductions. For instance, if we compare `tonk`-reduction to the Liar-reduction given above, of course, they look very similar. But in the Liar case type reconstruction quickly shows that this reduction is well-behaved, while the `tonk`-reduction is not.

So, what we are actually asking for is what I will call a 'weak' subject reduction:

**Weak Subject Reduction:**

(i) If $\Gamma \vdash t : A$ and $t \leadsto t'$, then $\Gamma \vdash t' : A$, or

(ii) if $\Gamma \vdash t : A$, $t \leadsto t'$ and $\Gamma \nvdash t' : A$, then it is not the case that $\Gamma \vdash t' : B$ for arbitrary $B$. $B$ is considered *arbitrary* iff the rules of type assignment do not determine the type reconstruction of $t$ in a way that $B$ is related to $A$.

This is what I propose to demand as a criterion for a reduction to be acceptable: it should enjoy the property of weak subject reduction. To reformulate it in other words, what we demand is that for the case that 'full' subject reduction, i.e., clause (i), fails, $\Gamma \vdash t' : B$ holds only for those $B$, which the rules of type assignment relate to $A$ in the type reconstruction of $t$. This criterion ensures that whenever subject reduction holds, weak subject reduction holds as well, i.e., failure of weak subject reduction also implies failure of 'full' subject reduction. That is important because it means that not meeting this desideratum only rules out the 'bad' reductions. The ones for which 'full' subject reduction holds, which, as we said above, is deemed to be sufficient for 'acceptable' reductions, cannot be ruled out by that criterion. Also, note, that failure of weak subject reduction does not necessarily mean that there is something wrong with the rules of type assignment in question. In the case of Ekman-reduction there is nothing wrong with the rules, since the only rules used are the ones for $\to$ and those are fine for the $\beta$-reduction. It rather shows that the reduction generated on grounds of these rules is misbehaved: it may work for specific types but it cannot be generalized in the same way 'proper' reductions can.

Our way of *checking* whether weak subject reduction holds or not is, then, via type reconstructions in the way described above: We conduct a type reconstruction for a term that would count as non-normal under this reduction, i.e., a redex, choosing 'fresh' types whenever the type assignment rules allow this. If the resulting types of redex and contractum occurring in this reconstruction are of arbitrarily different, unrelated types, then weak subject reduction fails and this means that this reduction should be rejected. Thus, we do not only have a clear criterion of what distinguishes acceptable from unacceptable reductions but also a fairly simple way of testing this by the respective type reconstruction. That can be considered an advantage when comparing it to Schroeder-Heister and Tranchini's way of showing how Ekman-reduction leads to unacceptable consequences, which they do by giving a very well-thought-out example of certain derivations leading to these consequences. This is a very clever and sophisticated way, for sure, but one has to be able to come up with these examples in the first place. Here, on the other hand, we have a systematic procedure of checking whether a reduction is acceptable or not.

### 3.3.4 Type theory of core logic - another problematic case?

In the following I want to give a concrete example of a reduction which is not type preserving, i.e., does not enjoy 'full' subject reduction, but still does not necessarily have to be dismissed as a 'bad' reduction.[52] The reduction is presented by Ripley (2020b) as part of an interesting typed term calculus for Tennant's Core Logic, i.e., an intuitionistic relevant logic. The calculus, called *Core Type Theory*, is interesting because it displays some very unusual features, while at the same time it is - at least in some respects - quite well-behaved.

According to Ripley the system maintains a similar correspondence to the implication-negation-fragment of core logic as the one established by the Curry-Howard-correspondence between the simply typed λ-calculus and intuitionistic logic. In the proof system that he presents, next to formulas and connectives we have ☹, which is related to negation but should not be considered as something like $\bot$. ☹ is neither a formula nor a connective, i.e., it cannot be used to form any complex formulas, but rather, it is understood as a "structural marker that interacts with the connective rules" in a way specified by the proof system (Ripley, 2020b, p. 112). One of the things to note is that in core type theory in addition to the usual case where terms are of certain types (served by the formulas), here terms can also have ☹ instead of

[52]Another example could be found in (Wansing, 1993), where a type theory for Nelson's logic with strong negation, N4, is given, which identifies terms of type $A$ with terms of type $\sim\sim A$. A concrete reduction is not formulated there but it is likely that it would have similar features as the one discussed in this section, since in such a system there would have to be rules of type assignment for $\sim$ which in some way relate $A$ and $\sim\sim A$.

a type, in which case Ripley speaks of *exceptional terms.* Also, Ripley uses Church-style typing, i.e., the types and ☹ are part of the syntax of the terms.

Ignoring differences in notation, redex and contractum are defined in the same way as above, i.e., $App(\lambda x.t, s)$ as redex and the corresponding term $t[s/x]$ as contractum.[53] Relevant for our purpose is that the reduction procedure fails to be type preserving because it can happen that a typed term, on which we perform a reduction procedure, has an exceptional term as contractum. Also, the system is not confluent, which means that normal forms are not unique, i.e., two syntactically distinct terms in normal form do not necessarily belong to two distinct equivalence classes generated by this reduction. The system is indeed trivializing in the sense that *if* we assume proof identity via the equivalence relation induced by its reduction procedure, then every term would have to be identified with every other term (Ripley, 2020b, p. 128). What must be stressed, however, is that there is only one non-type-preserving direction that is possible, namely from typed to exceptional terms. We cannot go from terms of one type to terms of another type or from exceptional terms to typed terms. This is one of the preservation properties this system still has. Another is that the reduction can never lead to new free variables, i.e., the set of free variables in a redex is a (possibly proper) superset of the free variables in its contractum (Ripley, 2020b, p. 116). This is the same as in the simply typed $\lambda$-calculus.

Since this is no system of type assignment but a typed system à la Church, the issue of type reconstruction can actually not be raised (Sørensen & Urzyczyn, 2006, p. 66). However, it seems rather unproblematic to convert Ripley's system into a system in which types and ☹ are assigned to terms according to the inference rules that are given. The rules and reduction for $\to$ (he additionally considers rules for negation) would, then, look like this:

$$\dfrac{\begin{array}{c}[x:A]\\ \vdots\\ t:B\end{array}}{\lambda x.t : A \to B}\,{\to}I \qquad \dfrac{\begin{array}{c}[x:A]\\ \vdots\\ t:\frownie\end{array}}{\lambda x.t : A \to B}\,{\to}I! \qquad \dfrac{\begin{array}{c}\Gamma\\ \vdots\\ s:A\to B\end{array} \quad \begin{array}{c}\Delta\\ \vdots\\ t:A\end{array} \quad \begin{array}{c}[y:B]\\ \vdots\\ r:C\end{array}}{r[App(s,t)/y]:C}\,{\to}E$$

$$App(\lambda x.t, s) \rightsquigarrow_\beta t[s/x]$$

Since the inference rules are not as determined as our standard rules,[54] it is clear that type reconstruction[55] cannot be conducted in such a way that it yields a

[53] It may be noted here that in the definition of those, the differences between Curry- and Church-style typing are blurred somehow because the terms used in the definition are not typed, which would be the usual thing to be done in Church-style, see (Hindley, 1997, p. 26) or (Troelstra & Schwichtenberg, 2000, p. 13). Sørensen and Urzyczyn (2006), who Ripley refers to in this context, leave out the types in their definition, as well, however, they say themselves that their 'Church-style' is actually "halfway between the Curry style and the 'orthodox' Church style" (2006, p. 66).

[54] As can be seen, we have two $\to$-introduction rules.

[55] If we want to be very precise, we would have to speak of "*hat* reconstruction", "rules of *hat*

determinate result as with the standard rules. Importantly, however, neither does it result in complete arbitrariness of the kind we have seen with Ekman-reduction or `tonk`. What can happen indeed, is that due to the two →-I rules, we have *two possible* paths in the type reconstruction, but that's it:[56]

$$\dfrac{\lambda x.t :? \to B \qquad \begin{array}{c}\mathscr{D}'\\ \vdots \\ s :?\end{array}}{App(\lambda x.t, s) : B}\to E \qquad\qquad \dfrac{\dfrac{\begin{array}{c}[x : A]\\ \vdots\ \mathscr{D} \\ \color{red}{t : B/\frownie}\end{array}}{\lambda x.t : A \to B}\to I/\!\to I! \qquad \begin{array}{c}\mathscr{D}'\\ \vdots \\ s : A\end{array}}{App(\lambda x.t, s) : B}\to E$$

The two paths are marked by the step in red. Everything else will be exactly the same, though. Since the reduction is $App(\lambda x.t, s) \rightsquigarrow_\beta t[s/x]$, it can happen that the redex reduces to a contractum which does not have the same type (neither does it have *another* type, though, because ☹ is no type at all). So, this means that the reduction in this system is not type preserving, i.e., subject reduction fails. However, *weak* subject reduction holds since it cannot reduce to an arbitrary type. The contractum will be either of the same type as the redex or it will be assigned ☹, which *is* related to $B$ by the type assignment rules: thus, to this extent the type reconstruction is determined.

Therefore, we have a reduction in this system which is at least partially well-behaved. On the one hand, confluence and subject reduction fail and if we would like the equivalence relation induced by reductions to give us proof identity, the reduction in this type theory would certainly not be suitable, since it trivializes identity of terms. On the other hand, type reconstruction can be conducted in an ordered manner without the possibility of yielding arbitrary results. Thus, the cases in which subject reduction fails are not completely arbitrary concerning the types, since it is not possible, as opposed to Ekman- and `tonk`-reduction, that a well-typed term reduces to a term of an arbitrarily different type. An anonymous reviewer raised doubts about the acceptability of this system because the identity of terms would be trivialized by the reductions, demanding that disallowing this should rather be our minimal criterion for the acceptability of a system. Note here that we must distinguish between the equivalence relation induced by the reductions and the reduction relation itself. While the former is certainly too permissive to be interesting for a philosophical interpretation of the proof theory, the latter can still be recognized to be at least so well-behaved that it does not lead to an Ekman-`tonk`-ish kind of trivialization, which is the kind we are worried about for reasons to be discussed in the following section.

assignment", etc. since this is Ripley's terminology for including both types and ☹. For simplicity, though, and because the criterion of weak subject reduction would still be met under such a reformulation, we will stick to the usual terminology.

[56] Note that although in Core Logic we have a generalized form of the elimination rule for →, the instance of this rule here is the usual Modus Ponens since Ripley defines a redex being of the form $App(\lambda x.t, s)$.

## 3.4 Philosophical implications: Reduction procedures and meaning of proofs

One of Prawitz's most important conjectures in this context is that, since the reductions induce an equivalence relation and two derivations should be considered to represent the same proof iff they are equivalent, proofs relating via these reductions are identical in nature.[57] This means in general that one and the same proof may be linguistically represented by different derivations and that in natural deduction a derivation in *normal form* is the most direct form of representation of its denotation, i.e., the represented proof object.[58]

Failure of weak subject reduction means to have reductions that relate terms of arbitrarily different types, i.e., proofs of arbitrarily different formulas. If we consider reductions to induce identity of proofs, a feature that ultimately results in having to identify proofs of arbitrarily different formulas would certainly be undesirable. However, there is no necessity to subscribe to this identity theory of proofs. There are other views on theories about identity of proofs on the market,[59] of course, or it is also possible to argue like Tennant (2021, p. S599), who seems to be a bit of an agnostic when it comes to this question. He indicates, though, in response to the proposal made in (Schroeder-Heister & Tranchini, 2017) to discard Ekman-reduction because it leads to a trivialization of proof identity, that we do not know enough about identity of proofs to use it as a criterion for other conceptions. However, reductions can also be conceived of as calculations, evaluations, or interpretations of the given program, as discussed in Section 3.2. I will argue here that if we go for the latter conception of reductions, failure of weak subject reduction still remains a problem, even if it is not problematic for the identity of proofs anymore. While this would be a problem for the *denotation* of proofs, I want to show that the arbitrariness is a problematic feature also concerning the *sense* of proofs.

Tranchini (2016) argues that only proofs which contain connectives for which

[57] What is left undecided in Prawitz's remarks (1971, p. 257) is whether the $\beta$-reductions are the only conversions preserving identity of proofs or whether expansion operations (corresponding to $\eta$-expansions) and permutative conversions for $\vee$-elimination and $\exists$-elimination should be considered, as well. He seems to lean towards accepting at least the expansions, while Martin-Löf (1975, pp. 100f.) discards both kinds of operations for identity preservation. Girard (1989, pp. 16, 73), on the other hand, includes $\eta$-expansions but is highly sceptical w.r.t. the permutative conversions when it comes to the question of identifying the 'real objects' represented by the ND derivations. Since I am concerned only with *reductions* here and not conversions in general, I will leave this issue as it is.

[58] This Fregean formulation can be found in (Tranchini, 2016), where this is meant to explicate Prawitz's and Dummett's conceptions on these matters.

[59] Examples for other approaches to proof identity would be Straßburger's (e.g. 2019) based on graphical proof practice, like proof nets for linear logic or what he calls a 'combinatorial' approach for classical logic. Another one is Wansing's (in press) approach, where a notion of identity between derivations is defined on the basis of taking both the notion of proofs as well as *disproofs* as primitive. On this account, proofs of certain formulas can be identified with disproofs of other, in specific ways related formulas and based on this a bilateralist notion of synonymy between formulas is defined.

reduction procedures are available can have sense. He bases his argumentation on the Prawitzian tradition that derivations in normal form can be identified with the proof objects, i.e., their denotation, and the fact that the reductions are the instruments with which we can bring a derivation to its normal form. If reductions for terms are considered to be decisive for the meaning of proofs, it seems that we should be clear about the question of the present paper: What are the conditions of acceptable reduction procedures? In (Ayhan, 2021b) the general assumption from Tranchini, that the connectives appearing in a derivation need to have acceptable reductions in order for the derivation to have sense at all, is retained and based on this an approach with λ-term-annotated proof systems is motivated to spell out what the sense of derivations consists in. It is argued that in a term-annotated setting the denotation of derivations is represented by the end-term[60] of the derivation in normal form, since this term encodes the ultimate proof. The *sense* of a derivation, on the other hand, consists in the *set of terms occurring within the derivation* because those terms encode the intermediate steps in the construction of the complex end-term encoding the conclusion (Ayhan, 2021b, p. 578). Thus, these terms reflect the operations used in the derivation, i.e., they reflect the way that is taken to get to the denotation. Since they determine how the end-term is built up, they can be seen as encoding a procedure, which, finally, yields the end-term. This seems in accordance with what, e.g., Dummett (1973, pp. 232, 323, 636) (a "procedure" to determine the denotation), Girard (1989, p. 2) ("a sequence of *instructions*") or Horty (2007, pp. 66-69) ("senses as procedures") say about Fregean sense (Girard even in the context of relating this to the "proofs as programs" conception).[61]

According to Frege, what is crucial, is that the signs (here: the syntax) uniquely determine the sense and the sense uniquely determines the denotation. What can happen, though, is of course the classic example of 'Hesperus' and 'Phosphorus', where there is the same underlying denotation but different senses attached to different syntax (i.e., different words). In the context of proofs this would be indicated by different terms used within the derivations, ending, however, on the same end-term or being reducible to the same end-term. The following example illustrates this with a derivation in non-normal form reducing to the other, which is in normal form, since $App(\lambda y.\lambda x.x, \lambda y.y) \rightsquigarrow_\beta \lambda x.x$:

$$\dfrac{\dfrac{\dfrac{[x:p]}{\lambda x.x : p \to p}\,{\to}I}{\lambda y.\lambda x.x : (q \to q) \to (p \to p)}\,{\to}I \qquad \dfrac{[y:q]}{\lambda y.y : q \to q}\,{\to}I}{App(\lambda y.\lambda x.x, \lambda y.y) : p \to p}\,{\to}E \qquad\qquad \dfrac{[x:p]}{\lambda x.x : p \to p}\,{\to}I$$

[60] The term decorating the formula that is proven.

[61] Of course, there are other approaches on the Fregean sense like Evans' (1982), on whose conception the interpretation given here could not be considered since lacking denotation would mean lacking sense as well. However, I do not find that interpretation very convincing, especially not in this context, but also not in general (see also (Fitch & Nelson, 2018) on the problems of this conception).

The relation of these conceptions and the (un)acceptability of reduction procedures is the following now. Whether or not we see the reductions as generating identity, or 'merely' in this directed way as calculations, makes a difference concerning the denotation but *not* concerning the sense. We could use the theory described here but only equate terms over $\alpha$-conversion, for example. The derivations above, one reducing to the other, would not be identified anymore in this case but the senses would remain unchanged. They would not be identified because the denotation is referred to by the end-terms and if we do not assume identity over $\beta$-reductions, then these terms could not be identified, i.e., they would point to different proof objects. The senses, though, consist in both cases (whether or not we assume $\beta$-equality for the end-terms) of the terms occurring within the derivations, i.e., they are different from each other in both cases but each for itself does not change by that assumption about the denotation of the proofs. It would still hold that the sense determines the reference, in that there cannot be one sense leading to different denotations, and, importantly, that the syntax determines the sense. This can only be claimed to hold, however, if the rules (the syntax) determine the type reconstruction for the redexes of reductions (the sense) to the extent that types of redex and contractum are not arbitrarily unrelated. Otherwise, it cannot be said that the syntax determines the sense. That needs to be the case, though, since meaning must be governed by rules. It cannot be arbitrarily generated. So, even if we do not accept the assumption that reductions generate equivalence relations over which proofs can be identified, it still makes sense to disallow reductions which render the derivations they are related to meaningless.

## 3.5 Remarks on possible objections and Church- vs. Curry-style typing

An objection that may be raised against the present approach in general is that it does not actually tell us anything interesting or important about reductions but rather, that it shows a weakness of the underlying assumption that using such term-annotated proof systems is a good way to go. The argument could be delivered along the following lines: The version of giving Ekman-reduction in a term-annotated form, as presented in this paper, is too generalized to capture what the reduction in original form meant to express. In the original form the assumptions $A \rightarrow B$ and $B \rightarrow A$ are clearly essential for having such a reduction but these disappear completely in the reduction for terms considered here. Thus, it is inadmissible to take this reduction, lacking essential features, as a correspondence for the original Ekman-reduction and dismiss it on this basis.

Further, it could be argued that, in order to capture the original Ekman-reduction appropriately, a restriction on the types should be implemented. Such a restriction could be: $App(y, App(x,t)) \rightsquigarrow_{Ekman} t$ iff $t$ typechecks at every type $App(y, App(x,t))$ typechecks at.[62] What you get thereby is basically subject reduction *by definition.* It still might seem a bit more generalized than the original Ekman-reduction because we do not demand that the types are the same but the only way they could differ now is in the variables used for the atomic formulas, i.e., the principal types are always the same. Of course, with such a restriction we would not have the undesirable arbitrariness in type reconstruction and thus, such a reduction would be well-behaved.

However, such a restriction must be rejected as a 'saviour' for Ekman-reduction, since implementing it would entirely beg the question of what we wanted to devise here. Firstly, in exactly the same way `tonk`-reduction could be restored: By demanding that $k'kt \rightsquigarrow_{\texttt{tonk}} t$ iff $t$ typechecks at every type $k'kt$ typechecks at. I do not see, though, that this is what we are philosophically interested in when we want to investigate the nature of reductions. It could not be said anymore, as Prawitz did, that the reductions make the inversion principle explicit if what we are doing is to restrict them by definition to cases in which the inversion principle is maintained. What we are interested in, concerning the question "What are acceptable reductions?", is not to decide case by case whether it makes a difference to eliminate a certain detour or not, but to have some generalized form about which we can make such a judgment.

What must be considered, secondly, when asking why such a restriction should be rejected in our approach, is that by using it we would abandon basics of Curry-style typing and de facto do Church-style typing instead.[63] One could claim now, thereby raising another objection, that indeed, it simply would be better to use Church-style typing. The related objection would be to say that these features of type reconstruction just show that Curry-style typing has a severe disadvantage over Church-style, namely a looseness of the connection between terms and types, which makes it less beneficial for an approach like the present. In other words, what all of this shows, is not that there is something wrong with certain reductions, but that our typing system is not helpful for this question and that we should rather use a type system à la Church (against which the first objection could not be raised anymore, either).

However, if we are interested in proofs from a philosophical, rather than merely

[62] Note that it does not suffice here to demand that there is a type $A$ such that both $App(y, App(x,t))$ and $t$ typecheck at $A$. To see why this is not enough, consider the example at the end of Section 3.3.1. There is a type such that $App(y, App(x, \lambda z.z))$ and $\lambda z.z$ both typecheck at, namely $\sigma \to \sigma$ (if we had $\sigma \to \sigma$ instead of $\rho$ in the example, this would clearly work out).

[63] Whereas in Curry-style the syntax of the terms is independent of types, in Church-style types are part of the syntax of terms. This means that each variable is *uniquely* typed and therefore, e.g., $\lambda x^A.x^A$ is a term of type $A \to A$ but not of, let's say, $B \to B$. In Curry-style, on the other hand, $\lambda x.x$ is a term of type $A \to A$ *for every* $A$.

technical, point of view, then Curry-style typing is preferable to Church-style. In Church-style you *will* get invariance of types under conversion but just because of the definition of the language, not because it is an interesting property. All terms are typed, i.e., so are the $\beta$-redexes. Thus, $\beta$-reduction is restricted w.r.t. types and type changes are prevented (Hindley, 1997, p. 26). It is exactly these features of Church-style language, then, which prevent us from asking philosophically interesting questions. Because the types are part of the syntax of the terms and the typing rules are just part of the definition of the language, they cannot be used to answer questions about a more primitive, underlying language like "Is this term typeable/meaningful?", "Can this term be assigned this or that type?",... If you think of these questions as applied to proofs, these are philosophically interesting questions, though. With Curry-style we have a language in which those can be asked, in Church-style they are prevented simply by the definition of the language.

## 3.6 Conclusion

What cannot be provided by our analysis here is an exhaustive list of properties that reductions need to have in order to count as 'good', because this seems to depend on the role one wants a reduction to fulfill, which differs in the literature. All we can do is to draw a distinct line of what *unacceptable* reductions are: reductions which do not enjoy the property of *weak subject reduction*, that is, which yield the possibility of type reconstructions in which redex and contractum are arbitrarily independent of one another. Further, it can be claimed that if 'full' subject reduction fails, this does not necessarily need to lead to the exclusion of such a reduction. It is a reason to be careful about identifying terms via the reduction, though. Within the framework I outlined in this paper we have three kinds of reductions: firstly, the ones that are clearly well-behaved, like $\beta$-reductions. They have what seem to be very desirable features, like having the Church-Rosser-property, preserving types, etc. Secondly, we have reductions which are clearly not well-behaved. Those would be Ekman-reduction, or `tonk`-reduction, or any reduction which does not allow for a meaningful system because it arbitrarily connects terms of different types. Thirdly, we have reductions in between those two categories. These would be reductions like the example we saw in Section 3.3.4; ones which may lack desirable features but are still well-behaved enough that they need not necessarily be excluded. Whether or not one wants to accept them, then depends on the underlying philosophical theory (e.g., about identity of proofs) one is subscribing to.

What I showed in this paper is that the question of what makes up acceptable reductions is neither trivial nor easy to answer in a positive way. Thus, I make do with a negative answer, just like Schroeder-Heister and Tranchini do in their elaborate analysis of the topic in saying that acceptable reductions are not to yield an equivalence relation that trivializes the identity of proofs. While I agree with their analysis, I aimed at going a step further and show that even if one does not agree with the underlying assumption that reductions induce an identity relation for proofs, there are certain reductions, like Ekman-reduction, which still have to be considered problematic. The main point is that having to identify all proofs of the same formula is surely undesirable but it is all about the denotation. However, if we have to commit to a notion of reductions according to which terms of a certain type reduce to terms of arbitrarily unrelated types, then such a system cannot be considered rule-generated anymore, and thus, not meaningful.

# 4 A cut-free sequent calculus for the bi-intuitionistic logic 2Int

## 4.1 Introduction

The purpose of this paper is to introduce a bi-intuitionistic sequent calculus and to give proofs of admissibility for its structural rules. Since I will ponder over the philosophical problems and implications of this calculus in a different paper (Ayhan, 2021a), I only want to make some brief comments on these matters here. The calculus I will present, called SC2Int, is a sequent calculus for the bi-intuitionistic logic 2Int, which Wansing presents in (2016a). There, he also gives a natural deduction system for this logic, N2Int, to which SC2Int is equivalent in terms of what is derivable. I will spell out below what this amounts to exactly. What is important is that these calculi represent a kind of bilateralist reasoning, since they do not only internalize processes of verification or provability but also the dual processes in terms of falsification or what is called *dual provability.* In (Wansing, 2017) a normal form theorem for N2Int is stated and proven. Here, I want to prove a cut-elimination theorem for SC2Int, i.e., if successful, this would extend the results existing so far.

## 4.2 The calculus SC2Int

The language $\mathscr{L}_{2Int}$ of 2Int, as given by Wansing, is defined in Backus-Naur form as follows:
$A ::= p \mid \bot \mid \top \mid (A \wedge A) \mid (A \vee A) \mid (A \rightarrow A) \mid (A \prec A)$.
As can be seen, we have a non-standard connective in this language, namely the operator of co-implication $\prec$, which acts as a dual to implication, just like conjunction and disjunction can be seen as dual connectives. With that, we are in the realms of so-called *bi-intuitionistic logic*, which is a conservative extension of intuitionistic logic with co-implication.[64] We read $A \prec B$ as 'B co-implies A'.
The general design of SC2Int resembles the intuitionistic sequent calculus G3ip. The distinguishing features of this calculus consist in the shared contexts for all the logical rules, the axiom (in our calculus the reflexivity rules) being restricted to atomic formulas and the admissibility of *all* structural rules (see (Negri & von Plato, 2001, p. 28-30) for more information about the origins of this calculus). Another distinguishing feature is the repetition of $A \rightarrow B$ in the left premise of the left

[64]Note that there is also a use of *bi-intuitionistic logic* in the literature to refer to a specific system, namely BiInt, also called *Heyting-Brouwer logic*, e.g., in (Goré, 2000; Kowalski & Ono, 2017; Postniece, 2010; Rauszer, 1974). Co-implication is there to be understood to internalize the preservation of non-truth from the conclusion to the premises in a valid inference. The system 2Int, which is treated here, uses the same language as BiInt, but the meaning of co-implication differs (Wansing, 2016a, 2016c, 2017, p. 30f.).

introduction rule for implication, which is necessary for the proof of admissibility of contraction. Here, this happens in $\to L^a$ as well as with $A \prec B$ in $\prec L^c$.

We will use $p, q, r, \ldots$ for atomic formulas, $A, B, C, \ldots$ for arbitrary formulas, and $\Gamma, \Delta, \Gamma', \ldots$ for multisets of formulas. Sequents are of the form $(\Gamma; \Delta) \vdash^* C$ (with $\Gamma$ and $\Delta$ being finite, possibly empty multisets and $* \in \{+, -\}$). Thus, we have a calculus in which a duality of derivability relations is considered, not only the one of verification but also the one of falsification.[65] The formulas in $\Gamma$ can then be understood as *assumptions*, while the formulas in $\Delta$ can be understood as *counterassumptions*. `SC2Int` is equivalent to `N2Int` in that we have a proof in `N2Int` of $A$ from the pair $(\Gamma; \Delta)$ of assumptions $\Gamma$ and counterassumptions $\Delta$, iff the sequent $(\Gamma; \Delta) \vdash^+ A$ is derivable in `SC2Int` and we have a dual proof of $A$ from the pair $(\Gamma; \Delta)$ of assumptions $\Gamma$ and counterassumptions $\Delta$, iff the sequent $(\Gamma; \Delta) \vdash^- A$ is derivable in `SC2Int`.

In contrast to `G3ip`, there will be no distinction between axioms and logical rules but within the logical rules the zero-premise rules, which comprise $Rf^+$, $Rf^-$, $\bot L^a$, $\top L^c, \bot R^-$, and $\top R^+$, are distinguished from the non-zero-premise rules due to the special role of the former for the admissibility proofs below. Each of the logical rules has a *context* designated by $\Gamma$ and $\Delta$, *active formulas* designated by $A$ and $B$ and a *principal formula*, which is the one introduced on the left or right side of $\vdash^*$. Within the right introduction rules we need to distinguish whether the derivability relation expresses verification or falsification by using the superscripts + and -. Within the left introduction rules this is not necessary, but what is needed here is distinguishing an introduction of the principal formula into the *assumptions* from an introduction into the *counterassumptions*. The former are indexed by superscript $a$, while the latter are indexed by superscript $c$. The set of $R^+$ and $L^a$ rules are the *proof rules*; the set of $R^-$ and $L^c$ rules are the *dual proof rules*.

## SC2Int

For $* \in \{+, -\}$:

$$\frac{}{(\Gamma, p; \Delta) \vdash^+ p}\, Rf^+ \qquad \frac{}{(\Gamma; \Delta, p) \vdash^- p}\, Rf^-$$

$$\frac{}{(\Gamma, \bot; \Delta) \vdash^* C}\, \bot L^a \qquad \frac{}{(\Gamma; \Delta, \top) \vdash^* C}\, \top L^c$$

$$\frac{}{(\Gamma; \Delta) \vdash^- \bot}\, \bot R^- \qquad \frac{}{(\Gamma; \Delta) \vdash^+ \top}\, \top R^+$$

$$\frac{(\Gamma; \Delta) \vdash^+ A \qquad (\Gamma; \Delta) \vdash^+ B}{(\Gamma; \Delta) \vdash^+ A \wedge B}\, \wedge R^+ \qquad \frac{(\Gamma, A, B; \Delta) \vdash^* C}{(\Gamma, A \wedge B; \Delta) \vdash^* C}\, \wedge L^a$$

[65] In `N2Int` this is indicated by using single lines for verification and double lines for falsification.

$$\frac{(\Gamma;\Delta) \vdash^- A}{(\Gamma;\Delta) \vdash^- A \wedge B}\,{\scriptstyle \wedge R_1^-} \qquad \frac{(\Gamma;\Delta) \vdash^- B}{(\Gamma;\Delta) \vdash^- A \wedge B}\,{\scriptstyle \wedge R_2^-}$$

$$\frac{(\Gamma;\Delta, A) \vdash^* C \qquad (\Gamma;\Delta, B) \vdash^* C}{(\Gamma;\Delta, A \wedge B) \vdash^* C}\,{\scriptstyle \wedge L^c}$$

$$\frac{(\Gamma;\Delta) \vdash^+ A}{(\Gamma;\Delta) \vdash^+ A \vee B}\,{\scriptstyle \vee R_1^+} \qquad \frac{(\Gamma;\Delta) \vdash^+ B}{(\Gamma;\Delta) \vdash^+ A \vee B}\,{\scriptstyle \vee R_2^+}$$

$$\frac{(\Gamma, A;\Delta) \vdash^* C \qquad (\Gamma, B;\Delta) \vdash^* C}{(\Gamma, A \vee B;\Delta) \vdash^* C}\,{\scriptstyle \vee L^a}$$

$$\frac{(\Gamma;\Delta) \vdash^- A \qquad (\Gamma;\Delta) \vdash^- B}{(\Gamma;\Delta) \vdash^- A \vee B}\,{\scriptstyle \vee R^-} \qquad \frac{(\Gamma;\Delta, A, B) \vdash^* C}{(\Gamma;\Delta, A \vee B) \vdash^* C}\,{\scriptstyle \vee L^c}$$

$$\frac{(\Gamma, A;\Delta) \vdash^+ B}{(\Gamma;\Delta) \vdash^+ A \to B}\,{\scriptstyle \to R^+} \qquad \frac{(\Gamma, A \to B;\Delta) \vdash^+ A \qquad (\Gamma, B;\Delta) \vdash^* C}{(\Gamma, A \to B;\Delta) \vdash^* C}\,{\scriptstyle \to L^a}$$

$$\frac{(\Gamma;\Delta) \vdash^+ A \qquad (\Gamma;\Delta) \vdash^- B}{(\Gamma;\Delta) \vdash^- A \to B}\,{\scriptstyle \to R^-} \qquad \frac{(\Gamma, A;\Delta, B) \vdash^* C}{(\Gamma;\Delta, A \to B) \vdash^* C}\,{\scriptstyle \to L^c}$$

$$\frac{(\Gamma;\Delta) \vdash^+ A \qquad (\Gamma;\Delta) \vdash^- B}{(\Gamma;\Delta) \vdash^+ A \prec B}\,{\scriptstyle \prec R^+} \qquad \frac{(\Gamma, A;\Delta, B) \vdash^* C}{(\Gamma, A \prec B;\Delta) \vdash^* C}\,{\scriptstyle \prec L^a}$$

$$\frac{(\Gamma;\Delta, B) \vdash^- A}{(\Gamma;\Delta) \vdash^- A \prec B}\,{\scriptstyle \prec R^-} \qquad \frac{(\Gamma;\Delta, A \prec B) \vdash^- B \qquad (\Gamma;\Delta, A) \vdash^* C}{(\Gamma;\Delta, A \prec B) \vdash^* C}\,{\scriptstyle \prec L^c}$$

Note that the rules for $\wedge L^a$, $\vee L^c$, $\to L^c$ and $\prec L^a$ could also be given in the form of two rules, each with only one active formula $A$ or $B$, as it is for example done in Gentzen's original calculus for the left conjunction rule. We need this single rule formulation, however, in order to get the invertibility of these rules (see Lemma 4.3 below), which is important for the proof of admissibility of contraction. As said above, the structural rules do not have to be taken as primitive in the calculus but can be shown to be admissible. We want to consider rules for weakening, contraction and cut. Due to the dual nature of the calculus, we need two rules for each of these rules:

$$\frac{(\Gamma;\Delta) \vdash^* C}{(\Gamma, A;\Delta) \vdash^* C}\,{\scriptstyle W^a} \qquad \frac{(\Gamma;\Delta) \vdash^* C}{(\Gamma;\Delta, A) \vdash^* C}\,{\scriptstyle W^c}$$

$$\frac{(\Gamma, A, A;\Delta) \vdash^* C}{(\Gamma, A;\Delta) \vdash^* C}\,{\scriptstyle C^a} \qquad \frac{(\Gamma;\Delta, A, A) \vdash^* C}{(\Gamma;\Delta, A) \vdash^* C}\,{\scriptstyle C^c}$$

$$\frac{(\Gamma;\Delta) \vdash^+ D \qquad (\Gamma', D;\Delta') \vdash^* C}{(\Gamma, \Gamma';\Delta, \Delta') \vdash^* C}\,{\scriptstyle Cut^a} \qquad \frac{(\Gamma;\Delta) \vdash^- D \qquad (\Gamma';\Delta', D) \vdash^* C}{(\Gamma, \Gamma';\Delta, \Delta') \vdash^* C}\,{\scriptstyle Cut^c}$$

## 4.3 Proving admissibility of the structural rules

### 4.3.1 Preliminaries

The proofs of admissibility of the structural rules and especially of cut-elimination are conducted analogously to the respective proofs in (Negri & von Plato, 2001, pp. 30-40) for `G3ip`. The proofs will use induction on **weight of formulas** and **height of derivations**.

**Definition 4.1**
*The weight w(A) of a formula A is defined inductively by*
$w(\bot) = w(\top) = 0$,
$w(p) = 1$ *for atoms* $p$,
$w(A \# B) = w(A) + w(B) + 1$ *for* $\# \in \{\wedge, \vee, \rightarrow, \prec\}$.

**Definition 4.2**
*A derivation in* ***SC2Int*** *is either an instance of a zero-premise rule, or an application of a logical rule to derivations concluding its premises. The* height of a derivation *is the greatest number of successive applications of rules in it, where zero-premise rules have height 0.*

First, I will show that the reflexivity rules can be generalized to instances with arbitrary formulas, not only atomic formulas.

**Lemma 4.1**
*The sequents* $(\Gamma, C; \Delta) \vdash^+ C$ *and* $(\Gamma; \Delta, C) \vdash^- C$ *are derivable for an arbitrary formula* $C$ *and arbitrary context* $(\Gamma; \Delta)$.

*Proof.* The proof is by induction on weight of $C$. If $w(C) \leq 1$, we have the 19 cases listed below. Note that for some of the derivations there is more than one possibility to derive the desired sequent and also some of the conclusions of zero-premise rules are conclusions of more than one of those rules. I will just show one exemplary derivation for each case, since this is enough for the proof.

$C = \bot$. Then $(\Gamma, C; \Delta) \vdash^+ C$ is an instance of $\bot L^a$ and $(\Gamma; \Delta, C) \vdash^- C$ is an instance of $\bot R^-$.

$C = \top$. Then $(\Gamma, C; \Delta) \vdash^+ C$ is an instance of $\top R^+$ and $(\Gamma; \Delta, C) \vdash^- C$ is an instance of $\top L^c$.

$C = p$ for some atom $p$. Then $(\Gamma, C; \Delta) \vdash^+ C$ is an instance of $Rf^+$ and $(\Gamma; \Delta, C) \vdash^- C$ is an instance of $Rf^-$.

$C = \bot \wedge \bot$. Then $(\Gamma, C; \Delta) \vdash^+ C$ and $(\Gamma; \Delta, C) \vdash^- C$ are derived by

$$\dfrac{\overline{(\Gamma, \bot, \bot; \Delta) \vdash^+ \bot \wedge \bot}\ {\scriptstyle \bot L^a}}{(\Gamma, \bot \wedge \bot; \Delta) \vdash^+ \bot \wedge \bot}\ {\scriptstyle \wedge L^a} \quad \text{and} \quad \dfrac{\overline{(\Gamma; \Delta, \bot \wedge \bot) \vdash^- \bot}\ {\scriptstyle \bot R^-}}{(\Gamma; \Delta, \bot \wedge \bot) \vdash^- \bot \wedge \bot}\ {\scriptstyle \wedge R^-}$$

$C = \bot \vee \bot$. Then $(\Gamma, C; \Delta) \vdash^+ C$ and $(\Gamma; \Delta, C) \vdash^- C$ are derived by

$$\dfrac{\dfrac{}{(\Gamma,\bot;\Delta)\vdash^{+}\bot\vee\bot}\,{\bot L^{a}}\qquad\dfrac{}{(\Gamma,\bot;\Delta)\vdash^{+}\bot\vee\bot}\,{\bot L^{a}}}{(\Gamma,\bot\vee\bot;\Delta)\vdash^{+}\bot\vee\bot}\,{\vee L^{a}}\quad\text{and}\quad\dfrac{\dfrac{}{(\Gamma;\Delta,\bot\vee\bot)\vdash^{-}\bot}\,{\bot R^{-}}\qquad\dfrac{}{(\Gamma;\Delta,\bot\vee\bot)\vdash^{-}\bot}\,{\bot R^{-}}}{(\Gamma;\Delta,\bot\vee\bot)\vdash^{-}\bot\vee\bot}\,{\vee R^{-}}$$

$C=\bot\to\bot$. Then $(\Gamma,C;\Delta)\vdash^{+}C$ and $(\Gamma;\Delta,C)\vdash^{-}C$ are derived by

$$\dfrac{\dfrac{}{(\Gamma,\bot\to\bot,\bot;\Delta)\vdash^{+}\bot}\,{\bot L^{a}}}{(\Gamma,\bot\to\bot;\Delta)\vdash^{+}\bot\to\bot}\,{\to R^{+}}\quad\text{and}\quad\dfrac{\dfrac{}{(\Gamma,\bot;\Delta,\bot)\vdash^{-}\bot\to\bot}\,{\bot L^{a}}}{(\Gamma;\Delta,\bot\to\bot)\vdash^{-}\bot\to\bot}\,{\to L^{c}}$$

$C=\bot\prec\bot$. Then $(\Gamma,C;\Delta)\vdash^{+}C$ and $(\Gamma;\Delta,C)\vdash^{-}C$ are derived by

$$\dfrac{\dfrac{}{(\Gamma,\bot;\Delta,\bot)\vdash^{+}\bot\prec\bot}\,{\bot L^{a}}}{(\Gamma,\bot\prec\bot;\Delta)\vdash^{+}\bot\prec\bot}\,{\prec L^{a}}\quad\text{and}\quad\dfrac{\dfrac{}{(\Gamma;\Delta,\bot\prec\bot,\bot)\vdash^{-}\bot}\,{\bot R^{-}}}{(\Gamma;\Delta,\bot\prec\bot)\vdash^{-}\bot\prec\bot}\,{\prec R^{-}}$$

$C=\bot\wedge\top$. Then $(\Gamma,C;\Delta)\vdash^{+}C$ and $(\Gamma;\Delta,C)\vdash^{-}C$ are derived by

$$\dfrac{\dfrac{}{(\Gamma,\bot,\top;\Delta)\vdash^{+}\bot\wedge\top}\,{\bot L^{a}}}{(\Gamma,\bot\wedge\top;\Delta)\vdash^{+}\bot\wedge\top}\,{\wedge L^{a}}\quad\text{and}\quad\dfrac{\dfrac{}{(\Gamma;\Delta,\bot\wedge\top)\vdash^{-}\bot}\,{\bot R^{-}}}{(\Gamma;\Delta,\bot\wedge\top)\vdash^{-}\bot\wedge\top}\,{\wedge R^{-}_{1}}$$

$C=\bot\vee\top$. Then $(\Gamma,C;\Delta)\vdash^{+}C$ and $(\Gamma;\Delta,C)\vdash^{-}C$ are derived by

$$\dfrac{\dfrac{}{(\Gamma,\bot\vee\top;\Delta)\vdash^{+}\top}\,{\top R^{+}}}{(\Gamma,\bot\vee\top;\Delta)\vdash^{+}\bot\vee\top}\,{\vee R^{+}_{2}}\quad\text{and}\quad\dfrac{\dfrac{}{(\Gamma;\Delta,\bot,\top)\vdash^{-}\bot\vee\top}\,{\top L^{c}}}{(\Gamma;\Delta,\bot\vee\top)\vdash^{-}\bot\vee\top}\,{\vee L^{c}}$$

$C=\bot\to\top$. Then $(\Gamma,C;\Delta)\vdash^{+}C$ and $(\Gamma;\Delta,C)\vdash^{-}C$ are derived by

$$\dfrac{\dfrac{}{(\Gamma,\bot\to\top,\bot;\Delta)\vdash^{+}\top}\,{\top R^{+}}}{(\Gamma,\bot\to\top;\Delta)\vdash^{+}\bot\to\top}\,{\to R^{+}}\quad\text{and}\quad\dfrac{\dfrac{}{(\Gamma,\bot;\Delta,\top)\vdash^{-}\bot\to\top}\,{\top L^{c}}}{(\Gamma;\Delta,\bot\to\top)\vdash^{-}\bot\to\top}\,{\to L^{c}}$$

$C=\bot\prec\top$. Then $(\Gamma,C;\Delta)\vdash^{+}C$ and $(\Gamma;\Delta,C)\vdash^{-}C$ are derived by

$$\dfrac{\dfrac{}{(\Gamma,\bot;\Delta,\top)\vdash^{+}\bot\prec\top}\,{\bot L^{a}}}{(\Gamma,\bot\prec\top;\Delta)\vdash^{+}\bot\prec\top}\,{\prec L^{a}}\quad\text{and}\quad\dfrac{\dfrac{}{(\Gamma;\Delta,\bot\prec\top,\top)\vdash^{-}\bot}\,{\top L^{c}}}{(\Gamma;\Delta,\bot\prec\top)\vdash^{-}\bot\prec\top}\,{\prec R^{-}}$$

$C=\top\wedge\bot$. Then $(\Gamma,C;\Delta)\vdash^{+}C$ and $(\Gamma;\Delta,C)\vdash^{-}C$ are derived by

$$\dfrac{\dfrac{}{(\Gamma,\top,\bot;\Delta)\vdash^{+}\top\wedge\bot}\,{\bot L^{a}}}{(\Gamma,\top\wedge\bot;\Delta)\vdash^{+}\top\wedge\bot}\,{\wedge L^{a}}\quad\text{and}\quad\dfrac{\dfrac{}{(\Gamma;\Delta,\top\wedge\bot)\vdash^{-}\bot}\,{\bot R^{-}}}{(\Gamma;\Delta,\top\wedge\bot)\vdash^{-}\top\wedge\bot}\,{\wedge R^{-}_{2}}$$

$C=\top\vee\bot$. Then $(\Gamma,C;\Delta)\vdash^{+}C$ and $(\Gamma;\Delta,C)\vdash^{-}C$ are derived by

$$\dfrac{\dfrac{}{(\Gamma,\top\vee\bot;\Delta)\vdash^{+}\top}\,{\top R^{+}}}{(\Gamma,\top\vee\bot;\Delta)\vdash^{+}\top\vee\bot}\,{\vee R^{+}_{1}}\quad\text{and}\quad\dfrac{\dfrac{}{(\Gamma;\Delta,\top,\bot)\vdash^{-}\top\vee\bot}\,{\top L^{c}}}{(\Gamma;\Delta,\top\vee\bot)\vdash^{-}\top\vee\bot}\,{\vee L^{c}}$$

$C=\top\to\bot$. Then $(\Gamma,C;\Delta)\vdash^{+}C$ and $(\Gamma;\Delta,C)\vdash^{-}C$ are derived by

$$\dfrac{\dfrac{}{(\Gamma,\top\to\bot;\Delta)\vdash^{+}\top}\,{\top R^{+}}\qquad\dfrac{}{(\Gamma,\bot;\Delta)\vdash^{+}\top\to\bot}\,{\bot L^{a}}}{(\Gamma,\top\to\bot;\Delta)\vdash^{+}\top\to\bot}\,{\to L^{a}}\quad\text{and}\quad\dfrac{\dfrac{}{(\Gamma;\Delta,\top\to\bot)\vdash^{+}\top}\,{\top R^{+}}\qquad\dfrac{}{(\Gamma;\Delta,\top\to\bot)\vdash^{-}\bot}\,{\bot R^{-}}}{(\Gamma;\Delta,\top\to\bot)\vdash^{-}\top\to\bot}\,{\to R^{-}}$$

$C=\top\prec\bot$. Then $(\Gamma,C;\Delta)\vdash^{+}C$ and $(\Gamma;\Delta,C)\vdash^{-}C$ are derived by

$$\dfrac{\dfrac{}{(\Gamma,\top\prec\bot;\Delta)\vdash^{+}\top}\,{\top R^{+}}\qquad\dfrac{}{(\Gamma,\top\prec\bot;\Delta)\vdash^{-}\bot}\,{\bot R^{-}}}{(\Gamma,\top\prec\bot;\Delta)\vdash^{+}\top\prec\bot}\,{\prec R^{+}}\quad\text{and}\quad\dfrac{\dfrac{}{(\Gamma;\Delta,\top\prec\bot)\vdash^{-}\bot}\,{\bot R^{-}}\qquad\dfrac{}{(\Gamma;\Delta,\top)\vdash^{-}\top\prec\bot}\,{\top L^{c}}}{(\Gamma;\Delta,\top\prec\bot)\vdash^{-}\top\prec\bot}\,{\prec L^{c}}$$

$C=\top\wedge\top$. Then $(\Gamma,C;\Delta)\vdash^{+}C$ and $(\Gamma;\Delta,C)\vdash^{-}C$ are derived by

$$\dfrac{\dfrac{}{(\Gamma,\top\wedge\top;\Delta)\vdash^{+}\top}\,{\top R^{+}}\qquad\dfrac{}{(\Gamma,\top\wedge\top;\Delta)\vdash^{+}\top}\,{\top R^{+}}}{(\Gamma,\top\wedge\top;\Delta)\vdash^{+}\top\wedge\top}\,{\wedge R^{+}}\quad\text{and}\quad\dfrac{\dfrac{}{(\Gamma;\Delta,\top)\vdash^{-}\top\wedge\top}\,{\top L^{c}}\qquad\dfrac{}{(\Gamma;\Delta,\top)\vdash^{-}\top\wedge\top}\,{\top L^{c}}}{(\Gamma;\Delta,\top\wedge\top)\vdash^{-}\top\wedge\top}\,{\wedge L^{c}}$$

$C=\top\vee\top$. Then $(\Gamma,C;\Delta)\vdash^{+}C$ and $(\Gamma;\Delta,C)\vdash^{-}C$ are derived by

$$\dfrac{\dfrac{}{(\Gamma, \top \vee \top; \Delta) \vdash^{+} \top}\,{\scriptstyle \top R^{+}}}{(\Gamma, \top \vee \top; \Delta) \vdash^{+} \top \vee \top}\,{\scriptstyle \vee R^{+}} \quad \text{and} \quad \dfrac{\dfrac{}{(\Gamma; \Delta, \top, \top) \vdash^{-} \top \vee \top}\,{\scriptstyle \top L^{c}}}{(\Gamma; \Delta, \top \vee \top) \vdash^{-} \top \vee \top}\,{\scriptstyle \vee L^{c}}$$

$C = \top \to \top$. Then $(\Gamma, C; \Delta) \vdash^{+} C$ and $(\Gamma; \Delta, C) \vdash^{-} C$ are derived by

$$\dfrac{\dfrac{}{(\Gamma, \top \to \top, \top; \Delta) \vdash^{+} \top}\,{\scriptstyle \top R^{+}}}{(\Gamma, \top \to \top; \Delta) \vdash^{+} \top \to \top}\,{\scriptstyle \to R^{+}} \quad \text{and} \quad \dfrac{\dfrac{}{(\Gamma, \top; \Delta, \top) \vdash^{-} \top \to \top}\,{\scriptstyle \top L^{c}}}{(\Gamma; \Delta, \top \to \top) \vdash^{-} \top \to \top}\,{\scriptstyle \to L^{c}}$$

$C = \top \prec \top$. Then $(\Gamma, C; \Delta) \vdash^{+} C$ and $(\Gamma; \Delta, C) \vdash^{-} C$ are derived by

$$\dfrac{\dfrac{}{(\Gamma, \top; \Delta, \top) \vdash^{+} \top \prec \top}\,{\scriptstyle \top L^{c}}}{(\Gamma, \top \prec \top; \Delta) \vdash^{+} \top \prec \top}\,{\scriptstyle \prec L^{a}} \quad \text{and} \quad \dfrac{\dfrac{}{(\Gamma; \Delta, \top \prec \top, \top) \vdash^{-} \top}\,{\scriptstyle \top L^{c}}}{(\Gamma; \Delta, \top \prec \top) \vdash^{-} \top \prec \top}\,{\scriptstyle \prec R^{-}}$$

The inductive hypothesis is that $(\Gamma, C; \Delta) \vdash^{+} C$ and $(\Gamma; \Delta, C) \vdash^{-} C$ are derivable for all formulas $C$ with $w(C) \leq n$, and we have to show that $(\Gamma, D; \Delta) \vdash^{+} D$ and $(\Gamma; \Delta, D) \vdash^{-} D$ are derivable for formulas $D$ of weight $\leq n+1$. There are four cases:

$D = A \wedge B$. By the definition of weight and our inductive hypothesis, $w(A) \leq n$ and $w(B) \leq n$.
We can derive $(\Gamma, A \wedge B; \Delta) \vdash^{+} A \wedge B$ by

$$\dfrac{\dfrac{(\Gamma, A, B; \Delta) \vdash^{+} A}{(\Gamma, A \wedge B; \Delta) \vdash^{+} A}\,{\scriptstyle \wedge L^{a}} \qquad \dfrac{(\Gamma, A, B; \Delta) \vdash^{+} B}{(\Gamma, A \wedge B; \Delta) \vdash^{+} B}\,{\scriptstyle \wedge L^{a}}}{(\Gamma, A \wedge B; \Delta) \vdash^{+} A \wedge B}\,{\scriptstyle \wedge R^{+}}$$

and $(\Gamma; \Delta, A \wedge B) \vdash^{-} A \wedge B$ by

$$\dfrac{\dfrac{(\Gamma; \Delta, A) \vdash^{-} A}{(\Gamma; \Delta, A) \vdash^{-} A \wedge B}\,{\scriptstyle \wedge R_{1}^{-}} \qquad \dfrac{(\Gamma; \Delta, B) \vdash^{-} B}{(\Gamma; \Delta, B) \vdash^{-} A \wedge B}\,{\scriptstyle \wedge R_{2}^{-}}}{(\Gamma; \Delta, A \wedge B) \vdash^{-} A \wedge B}\,{\scriptstyle \wedge L^{c}}$$

$(\Gamma; \Delta, A) \vdash^{-} A$ and $(\Gamma; \Delta, B) \vdash^{-} B$ are derivable by the inductive hypothesis and since the context is arbitrary, so are $(\Gamma', A; \Delta) \vdash^{+} A$ and $(\Gamma'', B; \Delta) \vdash^{+} B$, for $\Gamma' = \Gamma, B$ and $\Gamma'' = \Gamma, A$.

$D = A \vee B$. As before, $w(A) \leq n$ and $w(B) \leq n$.
We can derive $(\Gamma, A \vee B; \Delta) \vdash^{+} A \vee B$ by

$$\dfrac{\dfrac{(\Gamma, A; \Delta) \vdash^{+} A}{(\Gamma, A; \Delta) \vdash^{+} A \vee B}\,{\scriptstyle \vee R_{1}^{+}} \qquad \dfrac{(\Gamma, B; \Delta) \vdash^{+} B}{(\Gamma, B; \Delta) \vdash^{+} A \vee B}\,{\scriptstyle \vee R_{2}^{+}}}{(\Gamma, A \vee B; \Delta) \vdash^{+} A \vee B}\,{\scriptstyle \vee L^{a}}$$

and $(\Gamma; \Delta, A \vee B) \vdash^{-} A \vee B$ by

$$\dfrac{\dfrac{(\Gamma; \Delta, A, B) \vdash^{-} A}{(\Gamma; \Delta, A \vee B) \vdash^{-} A}\,{\scriptstyle \vee L^{c}} \qquad \dfrac{(\Gamma; \Delta, A, B) \vdash^{-} B}{(\Gamma; \Delta, A \vee B) \vdash^{-} B}\,{\scriptstyle \vee L^{c}}}{(\Gamma; \Delta, A \vee B) \vdash^{-} A \vee B}\,{\scriptstyle \vee R^{-}}$$

Again, by inductive hypothesis we get the derivability of $(\Gamma, A; \Delta) \vdash^{+} A$ and $(\Gamma, B; \Delta) \vdash^{+} B$ and since the context is arbitrary, $(\Gamma; \Delta', A) \vdash^{-} A$ and $(\Gamma; \Delta'', B) \vdash^{-} B$ are derivable, for $\Delta' = \Delta, B$ and $\Delta'' = \Delta, A$.

$D = A \to B$. As before, $w(A) \leq n$ and $w(B) \leq n$.
We can derive $(\Gamma, A \to B; \Delta) \vdash^+ A \to B$ by

$$\dfrac{\dfrac{(\Gamma, A, A \to B; \Delta) \vdash^+ A \qquad (\Gamma, A, B; \Delta) \vdash^+ B}{(\Gamma, A, A \to B; \Delta) \vdash^+ B}\, {\scriptstyle \to L^a}}{(\Gamma, A \to B; \Delta) \vdash^+ A \to B}\, {\scriptstyle \to R^+}$$

and $(\Gamma; \Delta, A \to B) \vdash^- A \to B$ by

$$\dfrac{\dfrac{(\Gamma, A; \Delta, B) \vdash^+ A \qquad (\Gamma, A; \Delta, B) \vdash^- B}{(\Gamma, A; \Delta, B) \vdash^- A \to B}\, {\scriptstyle \to R^-}}{(\Gamma; \Delta, A \to B) \vdash^- A \to B}\, {\scriptstyle \to L^c}$$

The case of $(\Gamma, A, B; \Delta) \vdash^+ B$ was already mentioned in the case of conjunction and with the same reasoning $(\Gamma', A; \Delta) \vdash^+ A$ for $\Gamma' = \Gamma, A \to B$, $(\Gamma, A; \Delta') \vdash^+ A$ for $\Delta' = \Delta, B$ as well as $(\Gamma'; \Delta, B) \vdash^- B$ for $\Gamma' = \Gamma, A$ are derivable.

$D = A \prec B$. As before, $w(A) \leq n$ and $w(B) \leq n$.
We can derive $(\Gamma, A \prec B; \Delta) \vdash^+ A \prec B$ by

$$\dfrac{\dfrac{(\Gamma, A; \Delta, B) \vdash^+ A \qquad (\Gamma, A; \Delta, B) \vdash^- B}{(\Gamma, A; \Delta, B) \vdash^+ A \prec B}\, {\scriptstyle \prec R^+}}{(\Gamma, A \prec B; \Delta) \vdash^+ A \prec B}\, {\scriptstyle \prec L^a}$$

and $(\Gamma; \Delta, A \prec B) \vdash^- A \prec B$ by

$$\dfrac{\dfrac{(\Gamma; \Delta, B, A \prec B) \vdash^- B \qquad (\Gamma; \Delta, A, B) \vdash^- A}{(\Gamma; \Delta, B, A \prec B) \vdash^- A}\, {\scriptstyle \prec L^c}}{(\Gamma; \Delta, A \prec B) \vdash^- A \prec B}\, {\scriptstyle \prec R^-}$$

With the same reasoning as above $(\Gamma; \Delta', B) \vdash^- B$ is derivable for $\Delta' = \Delta, A \prec B$ and all other cases are already dealt with above. □

### 4.3.2 Admissibility of weakening

I will now start with the proof of admissibility of weakening by induction on height of derivation. The general procedure when proving admissibility of a rule with this is to prove it for applications of this rule to conclusions of zero-premise rules and then generalize by induction on the number of applications of the rule to arbitrary derivations. Thus, we can assume that there is only one instance - as the last step in the derivation - of the rule in question.

**Theorem 4.1** (Height-preserving weakening)
*If $(\Gamma; \Delta) \vdash^* C$ is derivable with a height of derivation at most $n$, then $(\Gamma, D; \Delta) \vdash^* C$ and $(\Gamma; \Delta, D) \vdash^* C$ are derivable with a height of derivation at most $n$ for arbitrary $D$.*

*Proof.* If $n = 0$, then $(\Gamma; \Delta) \vdash^* C$ is a zero-premise rule, which means that one of the following six cases holds. $C$ is an atom and 1) a formula in $\Gamma$ with $* = +$ or 2) a formula in $\Delta$ with $* = -$. Otherwise it can be the case that 3) $C$ is $\top$ with

$* = +$ or 4) $C$ is $\bot$ with $* = -$. Lastly, it could be that 5) $\bot$ is a formula in $\Gamma$ or 6) $\top$ a formula in $\Delta$. In either case, $(\Gamma, D; \Delta) \vdash^* C$ *and* $(\Gamma; \Delta, D) \vdash^* C$ are conclusions of the respective zero-premise rules. Our inductive hypothesis is now that height-preserving weakening is admissible up to derivations of height $\leq n$. Let $(\Gamma; \Delta) \vdash^* C$ be derivable with a height of derivation at most $n + 1$.
If the last rule applied is $\wedge L^a$, then $\Gamma = \Gamma', A \wedge B$ and the last step is

$$\frac{(\Gamma', A, B; \Delta) \vdash^* C}{(\Gamma', A \wedge B; \Delta) \vdash^* C} \wedge L^a$$

So $(\Gamma', A, B; \Delta) \vdash^* C$ is derivable in $\leq n$ steps. By inductive hypothesis, also $(\Gamma', A, B, D; \Delta) \vdash^* C$ and $(\Gamma', A, B; \Delta, D) \vdash^* C$ are derivable in $\leq n$ steps. Thus, the application of $\wedge L^a$ gives a derivation of $(\Gamma', A \wedge B, D; \Delta) \vdash^* C$ and $(\Gamma', A \wedge B; \Delta, D) \vdash^* C$ in $\leq n + 1$ steps.
If the last rule applied is $\wedge L^c$, then $\Delta = \Delta', A \wedge B$ and the last step is

$$\frac{(\Gamma; \Delta', A) \vdash^* C \qquad (\Gamma; \Delta', B) \vdash^* C}{(\Gamma; \Delta', A \wedge B) \vdash^* C} \wedge L^c$$

So $(\Gamma; \Delta', A) \vdash^* C$ and $(\Gamma; \Delta', B) \vdash^* C$ are derivable in $\leq n$ steps. By inductive hypothesis, also $(\Gamma, D; \Delta', A) \vdash^* C$, $(\Gamma; \Delta', A, D) \vdash^* C$, $(\Gamma, D; \Delta', B) \vdash^* C$ and $(\Gamma; \Delta', B, D) \vdash^* C$ are derivable in $\leq n$ steps. Thus, the application of $\wedge L^c$ to the first and the third premise and to the second and the fourth premise gives a derivation of $(\Gamma, D; \Delta', A \wedge B) \vdash^* C$ and $(\Gamma; \Delta', A \wedge B, D) \vdash^* C$, respectively, in $\leq n + 1$ steps.
If the last rule applied is $\wedge R^+$, then $C = A \wedge B$ and the last step is

$$\frac{(\Gamma; \Delta) \vdash^+ A \qquad (\Gamma; \Delta) \vdash^+ B}{(\Gamma; \Delta) \vdash^+ A \wedge B} \wedge R^+$$

So $(\Gamma; \Delta) \vdash^+ A$ and $(\Gamma; \Delta) \vdash^+ B$ are derivable in $\leq n$ steps. By inductive hypothesis, also $(\Gamma, D; \Delta) \vdash^+ A$, $(\Gamma; \Delta, D) \vdash^+ A$, $(\Gamma, D; \Delta) \vdash^+ B$ and $(\Gamma; \Delta, D) \vdash^+ B$ are derivable in $\leq n$ steps. Thus, the application of $\wedge R^+$ to the first and the third premise and to the second and the fourth premise gives a derivation of $(\Gamma, D; \Delta) \vdash^+ A \wedge B$ and $(\Gamma; \Delta, D) \vdash^+ A \wedge B$, respectively, in $\leq n + 1$ steps.
If the last rule applied is $\wedge R_1^-$, then $C = A \wedge B$ and the last step is

$$\frac{(\Gamma; \Delta) \vdash^- A}{(\Gamma; \Delta) \vdash^- A \wedge B} \wedge R_1^-$$

So $(\Gamma; \Delta) \vdash^- A$ is derivable in $\leq n$ steps. By inductive hypothesis, also $(\Gamma, D; \Delta) \vdash^- A$ and $(\Gamma; \Delta, D) \vdash^- A$ are derivable in $\leq n$ steps. Thus, the application of $\wedge R_1^-$ gives a derivation of $(\Gamma, D; \Delta) \vdash^- A \wedge B$ and $(\Gamma; \Delta, D) \vdash^- A \wedge B$ in $\leq n + 1$ steps.

For the other logical rules the same can be shown with similar steps. □

Now I want to show one other result related to weakening because we will need this later in our proof for the admissibility of the cut rules, namely that for the special case that the weakening formula is $\top$ for $W^a$ and respectively $\bot$ for $W^c$, the weakening rules are invertible, i.e.:

$$\frac{(\Gamma, \top; \Delta) \vdash^* C}{(\Gamma; \Delta) \vdash^* C}\, W^{\top}_{inv} \qquad \frac{(\Gamma; \Delta, \bot) \vdash^* C}{(\Gamma; \Delta) \vdash^* C}\, W^{\bot}_{inv}$$

**Lemma 4.2** (Special case of inverted weakening)
*If* $(\Gamma, \top; \Delta) \vdash^* C$ *or* $(\Gamma; \Delta, \bot) \vdash^* C$ *are derivable with a height of derivation at most* $n$*, then so is* $(\Gamma; \Delta) \vdash^* C$.

*Proof.* If $n = 0$, then exactly the same reasoning as for Theorem 4.1 can be applied here.
Now we assume height-preserving invertibility for these two special cases of weakening up to height $n$, and let $(\Gamma, \top; \Delta) \vdash^* C$ and $(\Gamma; \Delta, \bot) \vdash^* C$ be derivable with a height of derivation $\leq n + 1$. The proof works correspondingly to the proof of height-preserving weakening above, I will show it for the case of the $\to L^c$-rule this time, just to choose one that is not familiar in 'usual' calculi, but it works similar for all logical connectives and their rules.
If the last rule applied is $\to L^c$, then we have $\Delta = \Delta', A \to B$ and the last step is

$$\frac{(\Gamma, A, \top; \Delta', B) \vdash^* C}{(\Gamma, \top; \Delta', A \to B) \vdash^* C}\, {\to}L^c \qquad \text{or respectively} \qquad \frac{(\Gamma, A; \Delta', B, \bot) \vdash^* C}{(\Gamma; \Delta', A \to B, \bot) \vdash^* C}\, {\to}L^c$$

So, $(\Gamma, A, \top; \Delta', B) \vdash^* C$ and $(\Gamma, A; \Delta', B, \bot) \vdash^* C$ are derivable in $\leq n$ steps. Then by inductive hypothesis, $(\Gamma, A; \Delta', B) \vdash^* C$ is derivable in $\leq n$ steps. If we apply $\to L^c$ to this, this gives us $(\Gamma; \Delta', A \to B) \vdash^* C$ in $\leq n + 1$ steps. □

### 4.3.3 Admissibility of contraction

Before we can prove the admissibility of the contraction rules, we need to prove the following lemma about the invertibility of premises and conclusions of the logical rules for the left introduction of formulas. Note that for $\to L^a$ and $\prec L^c$ the invertibility only holds for the right premises.[66]

**Lemma 4.3** (Inversion) $(i_1)$ *If* $(\Gamma, A \wedge B; \Delta) \vdash^* C$ *is derivable with a height of derivation at most* $n$*, then* $(\Gamma, A, B; \Delta) \vdash^* C$ *is derivable with a height of derivation at most* $n$*.*

$(i_2)$ *If* $(\Gamma; \Delta, A \wedge B) \vdash^* C$ *is derivable with a height of derivation at most* $n$*, then* $(\Gamma; \Delta, A) \vdash^* C$ *and* $(\Gamma; \Delta, B) \vdash^* C$ *are derivable with a height of derivation at most* $n$*.*

$(ii_1)$ *If* $(\Gamma, A \vee B; \Delta) \vdash^* C$ *is derivable with a height of derivation at most* $n$*, then* $(\Gamma, A; \Delta) \vdash^* C$ *and* $(\Gamma, B; \Delta) \vdash^* C$ *are derivable with a height of derivation at most* $n$*.*

[66] Negri and von Plato (2001, p. 33) give a counterexample for the implication rule. The analogous counterexamples for SC2Int would be the derivability of the sequents $(\bot \to \bot; \emptyset) \vdash^+ \bot \to \bot$ and $(\emptyset; \top \prec \top) \vdash^- \top \prec \top$.

($ii_2$) *If* $(\Gamma; \Delta, A \vee B) \vdash^* C$ *is derivable with a height of derivation at most* $n$*, then* $(\Gamma; \Delta, A, B) \vdash^* C$ *is derivable with a height of derivation at most* $n$*.*

($iii_1$) *If* $(\Gamma, A \to B; \Delta) \vdash^* C$ *is derivable with a height of derivation at most* $n$*, then* $(\Gamma, B; \Delta) \vdash^* C$ *is derivable with a height of derivation at most* $n$*.*

($iii_2$) *If* $(\Gamma; \Delta, A \to B) \vdash^* C$ *is derivable with a height of derivation at most* $n$*, then* $(\Gamma, A; \Delta, B) \vdash^* C$ *is derivable with a height of derivation at most* $n$*.*

($iv_1$) *If* $(\Gamma, A \prec B; \Delta) \vdash^* C$ *is derivable with a height of derivation at most* $n$*, then* $(\Gamma, A; \Delta, B) \vdash^* C$ *is derivable with a height of derivation at most* $n$*.*

($iv_2$) *If* $(\Gamma; \Delta, A \prec B) \vdash^* C$ *is derivable with a height of derivation at most* $n$*, then* $(\Gamma; \Delta, A) \vdash^* C$ *is derivable with a height of derivation at most* $n$*.*

*Proof.* The proof is by induction on $n$.
1.) If $(\Gamma, A \# B; \Delta) \vdash^* C$ with $\# \in \{\wedge, \vee, \to, \prec\}$ is the conclusion of a zero-premise rule, then so are $(\Gamma, A, B; \Delta) \vdash^* C$, $(\Gamma, A; \Delta) \vdash^* C$, $(\Gamma, B; \Delta) \vdash^* C$, $(\Gamma; \Delta, B) \vdash^* C$ since $A \# B$ is neither atomic nor $\bot$ nor $\top$.
Now we assume height-preserving inversion up to height $n$, and let $(\Gamma, A \# B; \Delta) \vdash^* C$ be derivable with a height of derivation $\leq n + 1$.

($i_1$) Either $A \wedge B$ is principal in the last rule or not. If $A \wedge B$ is the principal formula, the premise $(\Gamma, A, B; \Delta) \vdash^* C$ has a derivation of height $n$. If $A \wedge B$ is not principal in the last rule, then there must be one or two premises $(\Gamma', A \wedge B; \Delta') \vdash^* C'$, $(\Gamma'', A \wedge B; \Delta'') \vdash^* C''$ with a height of derivation $\leq n$. Then by inductive hypothesis, also $(\Gamma', A, B; \Delta') \vdash^* C'$, $(\Gamma'', A, B; \Delta'') \vdash^* C''$ are derivable with a height of derivation $\leq n$. Now the last rule can be applied to these premises to conclude $(\Gamma, A, B; \Delta) \vdash^* C$ in at most $n + 1$ steps.

($ii_1$) Either $A \vee B$ is principal in the last rule or not. If $A \vee B$ is the principal formula, the premises $(\Gamma, A; \Delta) \vdash^* C$ and $(\Gamma, B; \Delta) \vdash^* C$ have a derivation of height $\leq n$. If $A \vee B$ is not principal in the last rule, then there must be one or two premises $(\Gamma', A \vee B; \Delta') \vdash^* C'$, $(\Gamma'', A \vee B; \Delta'') \vdash^* C''$ with a height of derivation $\leq n$. Then by inductive hypothesis, also $(\Gamma', A; \Delta') \vdash^* C'$, $(\Gamma', B; \Delta') \vdash^* C'$ and $(\Gamma'', A; \Delta') \vdash^* C''$, $(\Gamma'', B; \Delta'') \vdash^* C''$ are derivable with a height of derivation $\leq n$. Now the last rule can be applied to the first and third premise to conclude $(\Gamma, A; \Delta) \vdash^* C$ and to the second and fourth premise to conclude $(\Gamma, B; \Delta) \vdash^* C$ in at most $n + 1$ steps.

($iii_1$) Either $A \to B$ is principal in the last rule or not. If $A \to B$ is the principal formula, the premise $(\Gamma, B; \Delta) \vdash^* C$ has a derivation of height $\leq n$. If $A \to B$ is not principal in the last rule, then there must be one or two premises $(\Gamma', A \to B; \Delta') \vdash^* C'$, $(\Gamma'', A \to B; \Delta'') \vdash^* C''$ with a height of derivation

$\leq n$. Then by inductive hypothesis, also $(\Gamma', B; \Delta') \vdash^* C'$, $(\Gamma'', B; \Delta'') \vdash^* C''$ are derivable with a height of derivation $\leq n$. Now the last rule can be applied to these premises to conclude $(\Gamma, B; \Delta) \vdash^* C$ in at most $n+1$ steps.

$(iv_1)$ Either $A \prec B$ is principal in the last rule or not. If $A \prec B$ is the principal formula, then the premise $(\Gamma, A; \Delta, B) \vdash^* C$ has a derivation of height $n$. If $A \prec B$ is not principal in the last rule, then there must be one or two premises $(\Gamma', A \prec B; \Delta') \vdash^* C'$, $(\Gamma'', A \prec B; \Delta'') \vdash^* C''$ with a height of derivation $\leq n$. Then by inductive hypothesis, also $(\Gamma', A; \Delta', B) \vdash^* C'$, $(\Gamma'', A; \Delta'', B) \vdash^* C''$ are derivable with a height of derivation $\leq n$. Now the last rule can be applied to these premises to conclude $(\Gamma, A; \Delta, B) \vdash^* C$ in at most $n+1$ steps.

2.) If $(\Gamma; \Delta, A \# B) \vdash^* C$ with $\# \in \{\wedge, \vee, \rightarrow, \prec\}$ is the conclusion of a zero-premise rule, then so are $(\Gamma; \Delta, A) \vdash^* C$, $(\Gamma; \Delta, B) \vdash^* C$, $(\Gamma; \Delta, A, B) \vdash^* C$, $(\Gamma, A; \Delta) \vdash^* C$ since $A \# B$ is neither atomic nor $\perp$ nor $\top$.
Now we assume height-preserving inversion up to height $n$, and let $(\Gamma; \Delta, A \# B) \vdash^* C$ be derivable with a height of derivation $\leq n+1$.

$(i_2)$ Either $A \wedge B$ is principal in the last rule or not. If $A \wedge B$ is the principal formula, the premises $(\Gamma; \Delta, A) \vdash^* C$ and $(\Gamma; \Delta, B) \vdash^* C$ have a derivation of height $\leq n$. If $A \wedge B$ is not principal in the last rule, then there must be one or two premises $(\Gamma'; \Delta', A \wedge B) \vdash^* C'$, $(\Gamma''; \Delta'', A \wedge B) \vdash^* C''$ with a height of derivation $\leq n$. Then by inductive hypothesis, also $(\Gamma'; \Delta', A) \vdash^* C'$, $(\Gamma'; \Delta', B) \vdash^* C'$, $(\Gamma''; \Delta'', A) \vdash^* C''$, $(\Gamma''; \Delta'', B) \vdash^* C''$ are derivable with a height of derivation $\leq n$. Now the last rule can be applied to the first and third premise to conclude $(\Gamma; \Delta, A) \vdash^* C$ and to the second and fourth premise to conclude $(\Gamma; \Delta, B) \vdash^* C$ in at most $n+1$ steps.

$(ii_2)$ Either $A \vee B$ is principal in the last rule or not. If $A \vee B$ is the principal formula, the premise $(\Gamma; \Delta, A, B) \vdash^* C$ has a derivation of height $n$. If $A \vee B$ is not principal in the last rule, then there must be one or two premises $(\Gamma'; \Delta', A \vee B) \vdash^* C'$, $(\Gamma''; \Delta'', A \vee B) \vdash^* C''$ with a height of derivation $\leq n$. Then by inductive hypothesis, also $(\Gamma'; \Delta', A, B) \vdash^* C'$, $(\Gamma''; \Delta'', A, B) \vdash^* C''$ are derivable with a height of derivation $\leq n$. Now the last rule can be applied to these premises to conclude $(\Gamma; \Delta, A, B) \vdash^* C$ in at most $n+1$ steps.

$(iii_2)$ Either $A \rightarrow B$ is principal in the last rule or not. If $A \rightarrow B$ is the principal formula, the premise $(\Gamma, A; \Delta, B) \vdash^* C$ has a derivation of height $n$. If $A \rightarrow B$ is not principal in the last rule, then there must be one or two premises $(\Gamma'; \Delta', A \rightarrow B) \vdash^* C'$, $(\Gamma''; \Delta'', A \rightarrow B) \vdash^* C''$ with a height of derivation $\leq n$. Then by inductive hypothesis, also $(\Gamma', A; \Delta', B) \vdash^* C'$, $(\Gamma'', A; \Delta'', B) \vdash^* C''$ are derivable with a height of derivation $\leq n$. Now the last rule can be applied to these premises to conclude $(\Gamma, A; \Delta, B) \vdash^* C$ in at most $n+1$ steps.

($iv_2$) Either $A \prec B$ is principal in the last rule or not. If $A \prec B$ is the principal formula, the premise $(\Gamma; \Delta, A) \vdash^* C$ has a derivation of height $\leq n$. If $A \prec B$ is not principal in the last rule, then there must be one or two premises $(\Gamma'; \Delta', A \prec B) \vdash^* C'$, $(\Gamma''; \Delta'', A \prec B) \vdash^* C''$ with a height of derivation $\leq n$. Then by inductive hypothesis, also $(\Gamma'; \Delta', A) \vdash^* C'$, $(\Gamma''; \Delta'', A) \vdash^* C''$ are derivable with a height of derivation $\leq n$. Now the last rule can be applied to these premises to conclude $(\Gamma; \Delta, A) \vdash^* C$ in at most $n + 1$ steps.

□

Next, I will prove the admissibility of the contraction rules in SC2Int.

**Theorem 4.2** (Height-preserving contraction)
*If $(\Gamma, D, D; \Delta) \vdash^* C$ is derivable with a height of derivation at most $n$, then $(\Gamma, D; \Delta) \vdash^* C$ is derivable with a height of derivation at most $n$ and if $(\Gamma; \Delta, D, D) \vdash^* C$ is derivable with a height of derivation at most $n$, then $(\Gamma; \Delta, D) \vdash^* C$ is derivable with a height of derivation at most $n$.*

*Proof.* The proof is again by induction on the height of derivation $n$.
If $(\Gamma, D, D; \Delta) \vdash^* C$ (resp. $(\Gamma; \Delta, D, D) \vdash^* C$) is the conclusion of a zero-premise rule, then either C is an atom and contained in the antecedent, in the assumptions for $\vdash^+$ or in the counterassumptions for $\vdash^-$, or $\bot$ is part of the assumptions, or $\top$ is part of the counterassumptions, or $C = \top$ for $\vdash^+$, or $C = \bot$ for $\vdash^-$. In either case, also $(\Gamma, D; \Delta) \vdash^* C$ (resp. $(\Gamma; \Delta, D) \vdash^* C$) is a conclusion of the respective zero-premise rule.
Let contraction be admissible up to derivation height $n$ and let $(\Gamma, D, D; \Delta) \vdash^* C$ (resp. $(\Gamma; \Delta, D, D) \vdash^* C$) be derivable in at most $n+1$ steps. Either the contraction formula is not principal in the last inference step or it is principal.
If $D$ is not principal in the last rule concluding the premise of contraction $(\Gamma, D, D; \Delta) \vdash^* C$, there must be one or two premises $(\Gamma', D, D; \Delta') \vdash^* C'$, $(\Gamma'', D, D; \Delta'') \vdash^* C'$ with a height of derivation $\leq n$. So by inductive hypothesis, we can derive $(\Gamma', D; \Delta') \vdash^* C'$, $(\Gamma'', D; \Delta'') \vdash^* C'$ with a height of derivation $\leq n$. Now the last rule can be applied to these premises to conclude $(\Gamma, D; \Delta) \vdash^* C$ in at most $n + 1$ steps. For the case of $(\Gamma; \Delta, D, D) \vdash^* C$ being the premise of contraction, the same argument applies respectively.
If $D$ is principal in the last rule, we have to consider four cases for each contraction rule according to the form of $D$. I will show the cases for $C^c$ this time; for $C^a$ the same arguments apply respectively.

$D = A \wedge B$. Then the last rule applied must be $\wedge L^c$ and we have as premises $(\Gamma; \Delta, A \wedge B, A) \vdash^* C$ and $(\Gamma; \Delta, A \wedge B, B) \vdash^* C$ with a derivation height $\leq n$. By the inversion lemma this means that $(\Gamma; \Delta, A, A) \vdash^* C$ and $(\Gamma; \Delta, B, B) \vdash^* C$ are also derivable with a derivation height $\leq n$. Then by inductive hypothesis, we get

$(\Gamma; \Delta, A) \vdash^* C$ and $(\Gamma; \Delta, B) \vdash^* C$ with a height of derivation $\leq n$ and by applying $\wedge L^c$ we can derive $(\Gamma; \Delta, A \wedge B) \vdash^* C$ in at most $n + 1$ steps.

$D = A \vee B$. Then the last rule applied must be $\vee L^c$ and $(\Gamma; \Delta, A \vee B, A, B) \vdash^* C$ is derivable with a height of derivation $\leq n$. By the inversion lemma, also $(\Gamma; \Delta, A, B, A, B) \vdash^* C$ is derivable with a derivation height $\leq n$. Then by inductive hypothesis (applied twice), we get $(\Gamma; \Delta, A, B) \vdash^* C$ with a height of derivation $\leq n$ and by applying $\vee L^c$ we can derive $(\Gamma; \Delta, A \vee B) \vdash^* C$ in at most $n + 1$ steps.

$D = A \rightarrow B$. Then the last rule applied must be $\rightarrow L^c$ and accordingly $(\Gamma, A; \Delta, B,$
$A \rightarrow B) \vdash^* C$ is derivable with a height of derivation $\leq n$. By the inversion lemma, then also $(\Gamma, A, A; \Delta, B, B) \vdash^* C$ is derivable with a derivation height $\leq n$. By inductive hypothesis (applied twice), we get $(\Gamma, A; \Delta, B) \vdash^* C$ with a height of derivation $\leq n$ and by applying $\rightarrow L^c$ we can derive $(\Gamma; \Delta, A \rightarrow B) \vdash^* C$ in at most $n + 1$ steps.

$D = A \prec B$. Then the last rule applied must be $\prec L^c$ and we have as premises $(\Gamma; \Delta, A \prec B, A \prec B) \vdash^- B$ and $(\Gamma; \Delta, A \prec B, A) \vdash^* C$ with a derivation height $\leq n$. The inductive hypothesis applied to the first, gives us $(\Gamma; \Delta, A \prec B) \vdash^- B$ with a derivation height $\leq n$ and the inversion lemma applied to the second, also $(\Gamma; \Delta, A, A) \vdash^* C$ and again by inductive hypothesis $(\Gamma; \Delta, A) \vdash^* C$ with a derivation height $\leq n$. By applying $\prec L^c$ we can now derive $(\Gamma; \Delta, A \prec B) \vdash^* C$ in at most $n+1$ steps. □

### 4.3.4 Admissibility of cut

Now, I will come to the main result, the proof of cut-elimination. The proof shows that cuts can be permuted upward in a derivation until they reach one of the zero-premise rules the derivation started with. When cut has reached zero-premise rules, the derivation can be transformed into one beginning with the conclusion of the cut, which can be shown by the following reasoning.

When both premises of cut are conclusions of a zero-premise rule, then the conclusion of cut is also a conclusion of one of these rules: If the left premise is $(\Gamma, \bot; \Delta) \vdash^* D$, then the conclusion also has $\bot$ in the assumptions of the antecedent. If the left premise is $(\Gamma; \Delta, \top) \vdash^* D$, then the conclusion also has $\top$ in the counterassumptions of the antecedent. If the left premise of $Cut^a$ is $(\Gamma; \Delta) \vdash^+ \top$ or the left premise of $Cut^c$ is $(\Gamma; \Delta) \vdash^- \bot$, then the right premise is $(\Gamma', \top; \Delta') \vdash^* C$ or $(\Gamma'; \Delta', \bot) \vdash^* C$ respectively. These are conclusions of zero-premise rules only in one of the following cases:

- $C$ is an atom in $\Gamma'$ for $* = +$ or $C$ is an atom in $\Delta'$ for $* =$ -

- $C = \top$ for $* = +$ or $C = \bot$ for $* =$ -
- $\bot$ is in $\Gamma'$ or $\top$ is in $\Delta'$

In each case the conclusion of cut $(\Gamma, \Gamma'; \Delta, \Delta') \vdash^* C$ is also a conclusion of the same zero-premise rule as the right premise. The last two possibilities are that the left premise is $(\Gamma, p; \Delta) \vdash^+ p$ for $Cut^a$ or $(\Gamma; \Delta, p) \vdash^- p$ for $Cut^c$ respectively. For the former case this means that the right premise is $(\Gamma', p; \Delta') \vdash^* C$. This is the conclusion of a zero-premise rule only in one of the following cases:

- For $* = +$: $C = p$, or $C$ is an atom in $\Gamma'$, or $C = \top$
- For $* =$ -: $C$ is an atom in $\Delta'$, or $C = \bot$
- $\bot$ is in $\Gamma'$, or $\top$ is in $\Delta'$

In each case the conclusion of cut $(\Gamma, p, \Gamma'; \Delta, \Delta') \vdash^* C$ is also a conclusion of the same zero-premise rule as the right premise. For the latter case this means that the right premise is $(\Gamma'; \Delta', p) \vdash^* C$. This is the conclusion of a zero-premise rule only in one of the following cases:

- For $* = +$: $C$ is an atom in $\Gamma'$, or $C = \top$
- For $* =$ -: $C = p$, or $C$ is an atom in $\Delta'$, or $C = \bot$
- $\bot$ is in $\Gamma'$, or $\top$ is in $\Delta'$

In each case the conclusion of cut $(\Gamma, \Gamma'; \Delta, p, \Delta') \vdash^* C$ is also a conclusion of the same zero-premise rule as the right premise. So, when cut has reached zero-premise rules as premises, the derivation can be transformed into one beginning with the conclusion of the cut by deleting the premises.

The proof is - as before - conducted in a manner corresponding to the proof of cut-elimination for G3ip by Negri and von Plato (2001), which means that it is by induction on the weight of the cut formula and a subinduction on the cut-height, the sum of heights of derivations of the two premises of cut.

**Definition 4.3**
*The* cut-height *of an application of one of the rules of cut in a derivation is the sum of heights of derivation of the two premises of the rule.*

In the proof permutations are given that always reduce the weight of the cut formula or the cut-height of instances of the rules. When the cut formula is not principal in at least one (or both) of the premises of cut, cut-height is reduced. In the other cases, i.e., in which the cut formula is principal in both premises, it is shown that cut-height and/or the weight of the cut formula can be reduced. This process terminates since atoms cannot be principal formulas.

The difference between the height of a derivation and cut-height needs to be emphasized here, because it is essential to understand that if there are two instances of cut, one occurring below the other in the derivation, this does not necessarily mean that the lower instance has a greater cut-height than the upper. Let us suppose the upper instance of cut occurs in the derivation of the left premise of the lower cut. The upper instance can have a cut-height which is greater than the height of either its premises because the *sum* of the premises is what matters. However, the lower instance can have as a right premise one with a much shorter derivation height than either of the premises of the upper cut, making the sum of the derivation heights of those two premises lesser than the one from the upper cut. So, what follows is that it is not enough to show that occurrences of cut can be permuted upward in a derivation in order to show that cut-height decreases, but we need to calculate exactly the cut-height of each derivation in our proof. As before, it can be assumed that in a given derivation the last instance is the one and only occurrence of cut.

**Theorem 4.3** (Cut admissibility)
*The cut rules*

$$\frac{(\Gamma;\Delta)\vdash^{+} D \qquad (\Gamma', D;\Delta')\vdash^{*} C}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^{*} C}\, Cut^{a} \qquad \textit{and} \qquad \frac{(\Gamma;\Delta)\vdash^{-} D \qquad (\Gamma';\Delta', D)\vdash^{*} C}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^{*} C}\, Cut^{c}$$

*are admissible in* ***SC2Int****.*

*Proof.* The proof is organized as follows. First, I consider the case that at least one premise in a cut is a conclusion of one of the zero-premise rules and show how cut can be eliminated in these cases. Otherwise three cases can be distinguished: 1.) The cut formula is not principal in either premise of cut, 2.) the cut formula is principal in just one premise of cut, and 3.) the cut formula is principal in both premises of cut.

## Cut with a conclusion of a zero-premise rule as premise

**Cut with a conclusion of $Rf^{+}$, $Rf^{-}$, $\bot L^{a}$, $\top L^{c}$, $\bot R^{-}$, or $\top R^{+}$ as premise:**

If at least one of the premises of cut is a conclusion of one of the zero-premise rules, we distinguish three cases for both cut rules:

### -1- Cut$^{a}$

-1.1- The left premise $(\Gamma;\Delta)\vdash^{+} D$ is a conclusion of a zero-premise-rule. There are four subcases:

(a) The cut formula $D$ is an atom in $\Gamma$. Then the conclusion $(\Gamma,\Gamma';\Delta,\Delta')\vdash^{*} C$ is derived from $(\Gamma', D;\Delta')\vdash^{*} C$ by $W^{a}$ and $W^{c}$.

(b) $\bot$ is a formula in $\Gamma$. Then $(\Gamma,\Gamma';\Delta,\Delta')\vdash^{*} C$ is also a conclusion of $\bot L^{a}$.

(c) $\top$ is a formula in $\Delta$. Then $(\Gamma, \Gamma'; \Delta, \Delta') \vdash^* C$ is also a conclusion of $\top L^c$.

(d) $\top =$ D. Then the right premise is $(\Gamma', \top; \Delta') \vdash^* C$ and $(\Gamma, \Gamma'; \Delta, \Delta') \vdash^* C$ follows by $W^{\top}_{inv}$ as well as $W^a$ and $W^c$.

-1.2- The right premise $(\Gamma', D; \Delta') \vdash^+ C$ is a conclusion of a zero-premise rule. There are six subcases:

(a) $C$ is an atom in $\Gamma'$. Then $(\Gamma, \Gamma'; \Delta, \Delta') \vdash^+ C$ is also a conclusion of $Rf^+$.

(b) $C = D$. Then the left premise is $(\Gamma; \Delta) \vdash^+ C$ and $(\Gamma, \Gamma'; \Delta, \Delta') \vdash^+ C$ follows by $W^a$ and $W^c$.

(c) $\bot$ is in $\Gamma'$. Then $(\Gamma, \Gamma'; \Delta, \Delta') \vdash^+ C$ is also a conclusion of $\bot L^a$.

(d) $\bot =$ D. Then the left premise is $(\Gamma; \Delta) \vdash^+ \bot$ and is either a conclusion of $\bot L^a$ or $\top L^c$ (in which case see 1.1 (b) or 1.1 (c)) or it has been derived by a left rule. There are eight cases according to the rule used which can be transformed into derivations with lesser cut-height. I will not show this here, since this is only a special case of the cases 3.1-3.8 below.

(e) $\top$ is in $\Delta'$. Then $(\Gamma, \Gamma'; \Delta, \Delta') \vdash^+ C$ is also a conclusion of $\top L^c$.

(f) $\top =$ C. Then $(\Gamma, \Gamma'; \Delta, \Delta') \vdash^+ C$ is also a conclusion of $\top R^+$.

-1.3- The right premise $(\Gamma', D; \Delta') \vdash^- C$ is a conclusion of a zero-premise rule. There are five subcases:

(a) $C$ is an atom in $\Delta'$. Then $(\Gamma, \Gamma'; \Delta, \Delta') \vdash^- C$ is also a conclusion of $Rf^-$.

(b) $\bot$ is in $\Gamma'$. Then $(\Gamma, \Gamma'; \Delta, \Delta') \vdash^- C$ is also a conclusion of $\bot L^a$.

(c) $\bot =$ D. Then the left premise is $(\Gamma; \Delta) \vdash^+ \bot$ and the same as mentioned in 1.2 (d) holds.

(d) $\top$ is in $\Delta'$. Then $(\Gamma, \Gamma'; \Delta, \Delta') \vdash^- C$ is also a conclusion of $\top L^c$.

(e) $\bot =$ C. Then $(\Gamma, \Gamma'; \Delta, \Delta') \vdash^- C$ is also a conclusion of $\bot R^-$.

**-2- Cut$^c$**

-2.1- The left premise $(\Gamma; \Delta) \vdash^- D$ is a conclusion of a zero-premise rule. There are four subcases:

(a) The cut formula $D$ is an atom in $\Delta$. Then the conclusion $(\Gamma, \Gamma'; \Delta, \Delta') \vdash^* C$ is derived from $(\Gamma'; \Delta', D) \vdash^* C$ by $W^a$ and $W^c$.

(b) $\bot$ is in $\Gamma$. Then $(\Gamma, \Gamma'; \Delta, \Delta') \vdash^* C$ is also a conclusion of $\bot L^a$.

(c) $\top$ is in $\Delta$. Then $(\Gamma, \Gamma'; \Delta, \Delta') \vdash^* C$ is also a conclusion of $\top L^c$.

(d) $\bot$ = D. Then the right premise is $(\Gamma'; \Delta', \bot) \vdash^* C$ and $(\Gamma, \Gamma'; \Delta, \Delta') \vdash^* C$ follows by $W^{\bot}_{inv}$ as well as $W^a$ and $W^c$.

-2.2- The right premise $(\Gamma'; \Delta', D) \vdash^+ C$ is a conclusion of a zero-premise rule. There are five subcases:

(a) $C$ is an atom in $\Gamma'$. Then $(\Gamma, \Gamma'; \Delta, \Delta') \vdash^+ C$ is also a conclusion of $Rf^+$.

(b) $\bot$ is in $\Gamma'$. Then $(\Gamma, \Gamma'; \Delta, \Delta') \vdash^+ C$ is also a conclusion of $\bot L^a$.

(c) $\top$ is in $\Delta'$. Then $(\Gamma, \Gamma'; \Delta, \Delta') \vdash^+ C$ is also a conclusion of $\top L^c$.

(d) $\top$ = D. Then the left premise is $(\Gamma; \Delta) \vdash^- \top$ and the same as mentioned in 1.2 (d) holds.

(e) $\top$ = C. Then $(\Gamma, \Gamma'; \Delta, \Delta') \vdash^+ C$ is also a conclusion of $\top R^+$.

-2.3- The right premise $(\Gamma'; \Delta', D) \vdash^- C$ is a conclusion of a zero-premise rule. There are six subcases:

(a) $C$ is an atom in $\Delta'$. Then $(\Gamma, \Gamma'; \Delta, \Delta') \vdash^- C$ is also a conclusion of $Rf^-$.

(b) $C = D$. Then the left premise is $(\Gamma; \Delta) \vdash^- C$ and $(\Gamma, \Gamma'; \Delta, \Delta') \vdash^- C$ follows by $W^a$ and $W^c$.

(c) $\bot$ is in $\Gamma'$. Then $(\Gamma, \Gamma'; \Delta, \Delta') \vdash^- C$ is also a conclusion of $\bot L^a$.

(d) $\top$ is in $\Delta'$. Then $(\Gamma, \Gamma'; \Delta, \Delta') \vdash^- C$ is also a conclusion of $\top L^c$.

(e) $\top$ = D. Then the left premise is $(\Gamma; \Delta) \vdash^- \top$ and the same as mentioned in 1.2 (d) holds.

(f) $\bot$ = C. Then $(\Gamma, \Gamma'; \Delta, \Delta') \vdash^- C$ is also a conclusion of $\bot R^-$.

## Cut with neither premise a conclusion of a zero-premise rule

We distinguish the cases that a left rule is used to derive the left premise (see -3-), a right rule is used to derive the left premise (see -5-), a right or a left rule is used to derive the right premise with the cut formula not being principal there (see -4-), and that a left rule is used to derive the right premise with the cut formula being principal (see -5-). These cases can be subsumed in a more compact form as categorized below. We assume, like Negri and von Plato (2001), that in the derivations the topsequents, from left to right, have derivation heights $n$, $m$, $k$,...

### -3- Cut not principal in the left premise

If the cut formula $D$ is not principal in the left premise, this means that this premise is derived by a left introduction rule. By permuting the order of the rules for the logical connectives with the cut rules, cut-height can be reduced in each of the following eight cases:

-3.1- $\wedge L^a$ is the last rule used to derive the left premise with $\Gamma = \Gamma'', A \wedge B$. The derivations for $Cut^a$ and $Cut^c$ with cuts of cut-height $n+1+m$ are

$$\dfrac{\dfrac{(\Gamma'', A, B; \Delta) \vdash^+ D}{(\Gamma'', A \wedge B; \Delta) \vdash^+ D}\wedge L^a \quad (\Gamma', D; \Delta') \vdash^* C}{(\Gamma'', A \wedge B, \Gamma'; \Delta, \Delta') \vdash^* C}Cut^a \qquad \dfrac{\dfrac{(\Gamma'', A, B; \Delta) \vdash^- D}{(\Gamma'', A \wedge B; \Delta) \vdash^- D}\wedge L^a \quad (\Gamma'; \Delta', D) \vdash^* C}{(\Gamma'', A \wedge B, \Gamma'; \Delta, \Delta') \vdash^* C}Cut^c$$

These can be transformed into derivations with cuts of cut-height $n+m$:

$$\dfrac{\dfrac{(\Gamma'', A, B; \Delta) \vdash^+ D \quad (\Gamma', D; \Delta') \vdash^* C}{(\Gamma'', A, B, \Gamma'; \Delta, \Delta') \vdash^* C}Cut^a}{(\Gamma'', A \wedge B, \Gamma'; \Delta, \Delta') \vdash^* C}\wedge L^a \qquad \dfrac{\dfrac{(\Gamma'', A, B; \Delta) \vdash^- D \quad (\Gamma'; \Delta', D) \vdash^* C}{(\Gamma'', A, B, \Gamma'; \Delta, \Delta') \vdash^* C}Cut^c}{(\Gamma'', A \wedge B, \Gamma'; \Delta, \Delta') \vdash^* C}\wedge L^a$$

-3.2- $\wedge L^c$ is the last rule used to derive the left premise with $\Delta = \Delta'', A \wedge B$. The derivations with cuts of cut-height $max(n,m)+1+k$ are

$$\dfrac{\dfrac{(\Gamma; \Delta'', A) \vdash^+ D \quad (\Gamma; \Delta'', B) \vdash^+ D}{(\Gamma; \Delta'', A \wedge B) \vdash^+ D}\wedge L^c \quad (\Gamma', D; \Delta') \vdash^* C}{(\Gamma, \Gamma'; \Delta'', A \wedge B, \Delta') \vdash^* C}Cut^a \qquad \dfrac{\dfrac{(\Gamma; \Delta'', A) \vdash^- D \quad (\Gamma; \Delta'', B) \vdash^- D}{(\Gamma; \Delta'', A \wedge B) \vdash^- D}\wedge L^c \quad (\Gamma'; \Delta', D) \vdash^* C}{(\Gamma, \Gamma'; \Delta'', A \wedge B, \Delta') \vdash^* C}Cut^c$$

These can be transformed into derivations each with two cuts of cut-height $n+k$ and $m+k$, respectively:

$$\dfrac{\dfrac{(\Gamma; \Delta'', A) \vdash^+ D \quad (\Gamma', D; \Delta') \vdash^* C}{(\Gamma, \Gamma'; \Delta'', A, \Delta') \vdash^* C}Cut^a \quad \dfrac{(\Gamma; \Delta'', B) \vdash^+ D \quad (\Gamma', D; \Delta') \vdash^* C}{(\Gamma, \Gamma'; \Delta'', B, \Delta') \vdash^* C}Cut^a}{(\Gamma, \Gamma'; \Delta'', A \wedge B, \Delta') \vdash^* C}\wedge L^c$$

$$\dfrac{\dfrac{(\Gamma; \Delta'', A) \vdash^- D \quad (\Gamma'; \Delta', D) \vdash^* C}{(\Gamma, \Gamma'; \Delta'', A, \Delta') \vdash^* C}Cut^c \quad \dfrac{(\Gamma; \Delta'', B) \vdash^- D \quad (\Gamma'; \Delta', D) \vdash^* C}{(\Gamma, \Gamma'; \Delta'', B, \Delta') \vdash^* C}Cut^c}{(\Gamma, \Gamma'; \Delta'', A \wedge B, \Delta') \vdash^* C}\wedge L^c$$

-3.3- $\vee L^a$ is the last rule used to derive the left premise with $\Gamma = \Gamma'', A \vee B$. The derivations with cuts of cut-height $max(n,m)+1+k$ are

$$\dfrac{\dfrac{(\Gamma'', A; \Delta) \vdash^+ D \quad (\Gamma'', B; \Delta) \vdash^+ D}{(\Gamma'', A \vee B; \Delta) \vdash^+ D}\vee L^a \quad (\Gamma', D; \Delta') \vdash^* C}{(\Gamma'', A \vee B, \Gamma'; \Delta, \Delta') \vdash^* C}Cut^a$$

$$\dfrac{\dfrac{(\Gamma'', A; \Delta) \vdash^- D \quad (\Gamma'', B; \Delta) \vdash^- D}{(\Gamma'', A \vee B; \Delta) \vdash^- D}\vee L^a \quad (\Gamma'; \Delta', D) \vdash^* C}{(\Gamma'', A \vee B, \Gamma'; \Delta, \Delta') \vdash^* C}Cut^c$$

These can be transformed into derivations each with two cuts of cut-height $n+k$ and $m+k$, respectively:

$$\dfrac{\dfrac{(\Gamma'', A; \Delta) \vdash^+ D \quad (\Gamma', D; \Delta') \vdash^* C}{(\Gamma'', A, \Gamma'; \Delta, \Delta') \vdash^* C}Cut^a \quad \dfrac{(\Gamma'', B; \Delta) \vdash^+ D \quad (\Gamma', D; \Delta') \vdash^* C}{(\Gamma'', B, \Gamma'; \Delta, \Delta') \vdash^* C}Cut^a}{(\Gamma'', A \vee B, \Gamma'; \Delta, \Delta') \vdash^* C}\vee L^a$$

$$\dfrac{\dfrac{(\Gamma'', A; \Delta) \vdash^- D \quad (\Gamma'; \Delta', D) \vdash^* C}{(\Gamma'', A, \Gamma'; \Delta, \Delta') \vdash^* C}\, Cut^c \qquad \dfrac{(\Gamma'', B; \Delta) \vdash^- D \quad (\Gamma'; \Delta', D) \vdash^* C}{(\Gamma'', B, \Gamma'; \Delta, \Delta') \vdash^* C}\, Cut^c}{(\Gamma'', A \vee B, \Gamma'; \Delta, \Delta') \vdash^* C}\, \vee L^a$$

-3.4- $\vee L^c$ is the last rule used to derive the left premise with $\Delta = \Delta'', A \vee B$. The derivations with cuts of cut-height $n + 1 + m$ are

$$\dfrac{\dfrac{(\Gamma; \Delta'', A, B) \vdash^+ D}{(\Gamma; \Delta'', A \vee B) \vdash^+ D}\, \vee L^c \quad (\Gamma', D; \Delta') \vdash^* C}{(\Gamma, \Gamma'; \Delta'', A \vee B, \Delta') \vdash^* C}\, Cut^a \qquad \dfrac{\dfrac{(\Gamma; \Delta'', A, B) \vdash^- D}{(\Gamma; \Delta'', A \vee B) \vdash^- D}\, \vee L^c \quad (\Gamma'; \Delta', D) \vdash^* C}{(\Gamma, \Gamma'; \Delta'', A \vee B, \Delta') \vdash^* C}\, Cut^c$$

These can be transformed into derivations with cuts of cut-height $n + m$:

$$\dfrac{\dfrac{(\Gamma; \Delta'', A, B) \vdash^+ D \quad (\Gamma', D; \Delta') \vdash^* C}{(\Gamma, \Gamma'; \Delta'', A, B, \Delta') \vdash^* C}\, Cut^a}{(\Gamma, \Gamma'; \Delta'', A \vee B, \Delta') \vdash^* C}\, \vee L^c \qquad \dfrac{\dfrac{(\Gamma; \Delta'', A, B) \vdash^- D \quad (\Gamma'; \Delta', D) \vdash^* C}{(\Gamma, \Gamma'; \Delta'', A, B, \Delta') \vdash^* C}\, Cut^c}{(\Gamma, \Gamma'; \Delta'', A \vee B, \Delta') \vdash^* C}\, \vee L^c$$

-3.5- $\to L^a$ is the last rule used to derive the left premise with $\Gamma = \Gamma'', A \to B$. The derivations with cuts of cut-height $max(n, m) + 1 + k$ are

$$\dfrac{\dfrac{(\Gamma'', A \to B; \Delta) \vdash^+ A \quad (\Gamma'', B; \Delta) \vdash^+ D}{(\Gamma'', A \to B; \Delta) \vdash^+ D}\, \to L^a \quad (\Gamma', D; \Delta') \vdash^* C}{(\Gamma'', A \to B, \Gamma'; \Delta, \Delta') \vdash^* C}\, Cut^a$$

$$\dfrac{\dfrac{(\Gamma'', A \to B; \Delta) \vdash^+ A \quad (\Gamma'', B; \Delta) \vdash^- D}{(\Gamma'', A \to B; \Delta) \vdash^- D}\, \to L^a \quad (\Gamma'; \Delta', D) \vdash^* C}{(\Gamma'', A \to B, \Gamma'; \Delta, \Delta') \vdash^* C}\, Cut^c$$

These can be transformed into derivations with cuts of cut-height $m + k$:

$$\dfrac{\dfrac{(\Gamma'', A \to B; \Delta) \vdash^+ A}{(\Gamma'', A \to B, \Gamma'; \Delta, \Delta') \vdash^+ A}\, W^{a/c} \quad \dfrac{(\Gamma'', B; \Delta) \vdash^+ D \quad (\Gamma', D; \Delta') \vdash^* C}{(\Gamma'', B, \Gamma'; \Delta, \Delta') \vdash^* C}\, Cut^a}{(\Gamma'', A \to B, \Gamma'; \Delta, \Delta') \vdash^* C}\, \to L^a$$

$$\dfrac{\dfrac{(\Gamma'', A \to B; \Delta) \vdash^+ A}{(\Gamma'', A \to B, \Gamma'; \Delta, \Delta') \vdash^+ A}\, W^{a/c} \quad \dfrac{(\Gamma'', B; \Delta) \vdash^- D \quad (\Gamma'; \Delta', D) \vdash^* C}{(\Gamma'', B, \Gamma'; \Delta, \Delta') \vdash^* C}\, Cut^c}{(\Gamma'', A \to B, \Gamma'; \Delta, \Delta') \vdash^* C}\, \to L^a$$

-3.6- $\to L^c$ is the last rule used to derive the left premise with $\Delta = \Delta'', A \to B$. The derivations with cuts of cut-height $n + 1 + m$ are

$$\dfrac{\dfrac{(\Gamma, A; \Delta'', B) \vdash^+ D}{(\Gamma; \Delta'', A \to B) \vdash^+ D}\, \to L^c \quad (\Gamma', D; \Delta') \vdash^* C}{(\Gamma, \Gamma'; \Delta'', A \to B, \Delta') \vdash^* C}\, Cut^a \qquad \dfrac{\dfrac{(\Gamma, A; \Delta'', B) \vdash^- D}{(\Gamma; \Delta'', A \to B) \vdash^- D}\, \to L^c \quad (\Gamma'; \Delta', D) \vdash^* C}{(\Gamma, \Gamma'; \Delta'', A \to B, \Delta') \vdash^* C}\, Cut^c$$

These can be transformed into derivations with cuts of cut-height $n + m$:

$$\dfrac{\dfrac{(\Gamma, A; \Delta'', B) \vdash^+ D \quad (\Gamma', D; \Delta') \vdash^* C}{(\Gamma, A, \Gamma'; \Delta'', B, \Delta') \vdash^* C}\, Cut^a}{(\Gamma, \Gamma'; \Delta'', A \to B, \Delta') \vdash^* C}\, \to L^c \qquad \dfrac{\dfrac{(\Gamma, A; \Delta'', B) \vdash^- D \quad (\Gamma'; \Delta', D) \vdash^* C}{(\Gamma, A, \Gamma'; \Delta'', B, \Delta') \vdash^* C}\, Cut^c}{(\Gamma, \Gamma'; \Delta'', A \to B, \Delta') \vdash^* C}\, \to L^c$$

-3.7- $\prec L^a$ is the last rule used to derive the left premise with $\Gamma = \Gamma'', A \prec B$. The derivations with cuts of cut-height $n + 1 + m$ are

$$\dfrac{\dfrac{(\Gamma'', A; \Delta, B) \vdash^+ D}{(\Gamma'', A \prec B; \Delta) \vdash^+ D}{\prec L^a} \quad (\Gamma', D; \Delta') \vdash^* C}{(\Gamma'', A \prec B, \Gamma'; \Delta, \Delta') \vdash^* C}{Cut^a} \qquad \dfrac{\dfrac{(\Gamma'', A; \Delta, B) \vdash^- D}{(\Gamma'', A \prec B; \Delta) \vdash^- D}{\prec L^a} \quad (\Gamma'; \Delta', D) \vdash^* C}{(\Gamma'', A \prec B, \Gamma'; \Delta, \Delta') \vdash^* C}{Cut^c}$$

These can be transformed into derivations with cuts of cut-height $n + m$:

$$\dfrac{\dfrac{(\Gamma'', A; \Delta, B) \vdash^+ D \quad (\Gamma', D; \Delta') \vdash^* C}{(\Gamma'', A, \Gamma'; \Delta, B, \Delta') \vdash^* C}{Cut^a}}{(\Gamma'', A \prec B, \Gamma'; \Delta, \Delta') \vdash^* C}{\prec L^a} \qquad \dfrac{\dfrac{(\Gamma'', A; \Delta, B) \vdash^- D \quad (\Gamma'; \Delta', D) \vdash^* C}{(\Gamma'', A, \Gamma'; \Delta, B, \Delta') \vdash^* C}{Cut^c}}{(\Gamma'', A \prec B, \Gamma'; \Delta, \Delta') \vdash^* C}{\prec L^a}$$

-3.8- $\prec L^c$ is the last rule used to derive the left premise with $\Delta = \Delta'', A \prec B$. The derivations with cuts of cut-height $max(n, m) + 1 + k$ are

$$\dfrac{\dfrac{(\Gamma; \Delta'', A \prec B) \vdash^- B \quad (\Gamma; \Delta'', A) \vdash^+ D}{(\Gamma; \Delta'', A \prec B) \vdash^+ D}{\prec L^c} \quad (\Gamma', D; \Delta') \vdash^* C}{(\Gamma, \Gamma'; \Delta'', A \prec B, \Delta') \vdash^* C}{Cut^a}$$

$$\dfrac{\dfrac{(\Gamma; \Delta'', A \prec B) \vdash^- B \quad (\Gamma; \Delta'', A) \vdash^- D}{(\Gamma; \Delta'', A \prec B) \vdash^- D}{\prec L^c} \quad (\Gamma'; \Delta', D) \vdash^* C}{(\Gamma, \Gamma'; \Delta'', A \prec B, \Delta') \vdash^* C}{Cut^c}$$

These can be transformed into derivations with cuts of cut-height $m + k$:

$$\dfrac{\dfrac{(\Gamma; \Delta'', A \prec B) \vdash^- B}{(\Gamma, \Gamma'; \Delta'', A \prec B, \Delta') \vdash^- B}{W^{a/c}} \quad \dfrac{(\Gamma; \Delta'', A) \vdash^+ D \quad (\Gamma', D; \Delta') \vdash^* C}{(\Gamma, \Gamma'; \Delta'', A, \Delta') \vdash^* C}{Cut^a}}{(\Gamma, \Gamma'; \Delta'', A \prec B, \Delta') \vdash^* C}{\prec L^c}$$

$$\dfrac{\dfrac{(\Gamma; \Delta'', A \prec B) \vdash^- B}{(\Gamma, \Gamma'; \Delta'', A \prec B, \Delta') \vdash^- B}{W^{a/c}} \quad \dfrac{(\Gamma; \Delta'' A) \vdash^- D \quad (\Gamma'; \Delta', D) \vdash^* C}{(\Gamma, \Gamma'; \Delta'', A, \Delta') \vdash^* C}{Cut^c}}{(\Gamma, \Gamma'; \Delta'', A \prec B, \Delta') \vdash^* C}{\prec L^c}$$

As said above, cut-height is reduced in all cases.

### -4- Cut formula *D* principal in the left premise only

The cases distinguished here concern the way the right premise is derived. We can distinguish 16 cases and show for each case that the derivation of the right premise can be transformed into one containing only occurrences of cut with a reduced cut-height.

-4.1- $\wedge L^a$ is the last rule used to derive the right premise with $\Gamma' = \Gamma'', A \wedge B$. The derivations with cuts of cut-height $n + m + 1$ are

$$\dfrac{(\Gamma; \Delta) \vdash^+ D \quad \dfrac{(\Gamma'', A, B, D; \Delta') \vdash^* C}{(\Gamma'', A \wedge B, D; \Delta') \vdash^* C}{\wedge L^a}}{(\Gamma, \Gamma'', A \wedge B; \Delta, \Delta') \vdash^* C}{Cut^a} \qquad \dfrac{(\Gamma; \Delta) \vdash^- D \quad \dfrac{(\Gamma'', A, B; \Delta', D) \vdash^* C}{(\Gamma'', A \wedge B; \Delta', D) \vdash^* C}{\wedge L^a}}{(\Gamma, \Gamma'', A \wedge B; \Delta, \Delta') \vdash^* C}{Cut^c}$$

These can be transformed into derivations with cuts of cut-height $n + m$:

$$\dfrac{\dfrac{(\Gamma;\Delta)\vdash^{+} D \quad (\Gamma'',A,B,D;\Delta')\vdash^{*} C}{(\Gamma,\Gamma'',A,B;\Delta,\Delta')\vdash^{*} C}\,Cut^{a}}{(\Gamma,\Gamma'',A\wedge B;\Delta,\Delta')\vdash^{*} C}\,\wedge L^{a} \qquad \dfrac{\dfrac{(\Gamma;\Delta)\vdash^{-} D \quad (\Gamma'',A,B;\Delta',D)\vdash^{*} C}{(\Gamma,\Gamma'',A,B;\Delta,\Delta')\vdash^{*} C}\,Cut^{c}}{(\Gamma,\Gamma'',A\wedge B;\Delta,\Delta')\vdash^{*} C}\,\wedge L^{a}$$

-4.2- $\wedge L^c$ is the last rule used to derive the right premise with $\Delta' = \Delta'', A\wedge B$. The derivations with cuts of cut-height $n + max(m,k) + 1$ are

$$\dfrac{(\Gamma;\Delta)\vdash^{+} D \quad \dfrac{(\Gamma',D;\Delta'',A)\vdash^{*} C \quad (\Gamma',D;\Delta'',B)\vdash^{*} C}{(\Gamma',D;\Delta'',A\wedge B)\vdash^{*} C}\,\wedge L^{c}}{(\Gamma,\Gamma';\Delta,\Delta'',A\wedge B)\vdash^{*} C}\,Cut^{a}$$

$$\dfrac{(\Gamma;\Delta)\vdash^{-} D \quad \dfrac{(\Gamma';\Delta'',A,D)\vdash^{*} C \quad (\Gamma';\Delta'',B,D)\vdash^{*} C}{(\Gamma';\Delta'',A\wedge B,D)\vdash^{*} C}\,\wedge L^{c}}{(\Gamma,\Gamma';\Delta,\Delta'',A\wedge B)\vdash^{*} C}\,Cut^{c}$$

These can be transformed into derivations each with two cuts of cut-height $n + m$ and $n + k$, respectively:

$$\dfrac{\dfrac{(\Gamma;\Delta)\vdash^{+} D \quad (\Gamma',D;\Delta'',A)\vdash^{*} C}{(\Gamma,\Gamma';\Delta,\Delta'',A)\vdash^{*} C}\,Cut^{a} \quad \dfrac{(\Gamma;\Delta)\vdash^{+} D \quad (\Gamma',D;\Delta'',B)\vdash^{*} C}{(\Gamma,\Gamma';\Delta,\Delta'',B)\vdash^{*} C}\,Cut^{a}}{(\Gamma,\Gamma';\Delta,\Delta'',A\wedge B)\vdash^{*} C}\,\wedge L^{c}$$

$$\dfrac{\dfrac{(\Gamma;\Delta)\vdash^{-} D \quad (\Gamma';\Delta'',A,D)\vdash^{*} C}{(\Gamma,\Gamma';\Delta,\Delta'',A)\vdash^{*} C}\,Cut^{c} \quad \dfrac{(\Gamma;\Delta)\vdash^{-} D \quad (\Gamma';\Delta'',B,D)\vdash^{*} C}{(\Gamma,\Gamma';\Delta,\Delta'',B)\vdash^{*} C}\,Cut^{c}}{(\Gamma,\Gamma';\Delta,\Delta'',A\wedge B)\vdash^{*} C}\,\wedge L^{c}$$

-4.3- $\vee L^a$ is the last rule used to derive the right premise with $\Gamma' = \Gamma'', A\vee B$. The derivations with cuts of cut-height $n + max(m,k) + 1$ are

$$\dfrac{(\Gamma;\Delta)\vdash^{+} D \quad \dfrac{(\Gamma'',A,D;\Delta')\vdash^{*} C \quad (\Gamma'',B,D;\Delta')\vdash^{*} C}{(\Gamma'',A\vee B,D;\Delta')\vdash^{*} C}\,\vee L^{a}}{(\Gamma,\Gamma'',A\vee B;\Delta,\Delta')\vdash^{*} C}\,Cut^{a}$$

$$\dfrac{(\Gamma;\Delta)\vdash^{-} D \quad \dfrac{(\Gamma'',A;\Delta',D)\vdash^{*} C \quad (\Gamma'',B;\Delta',D)\vdash^{*} C}{(\Gamma'',A\vee B;\Delta',D)\vdash^{*} C}\,\vee L^{a}}{(\Gamma,\Gamma'',A\vee B;\Delta,\Delta')\vdash^{*} C}\,Cut^{c}$$

These can be transformed into derivations each with two cuts of cut-height $n + m$ and $n + k$, respectively:

$$\dfrac{\dfrac{(\Gamma;\Delta)\vdash^{+} D \quad (\Gamma'',A,D;\Delta')\vdash^{*} C}{(\Gamma,\Gamma'',A;\Delta,\Delta')\vdash^{*} C}\,Cut^{a} \quad \dfrac{(\Gamma;\Delta)\vdash^{+} D \quad (\Gamma'',B,D;\Delta')\vdash^{*} C}{(\Gamma,\Gamma'',B;\Delta,\Delta')\vdash^{*} C}\,Cut^{a}}{(\Gamma,\Gamma'',A\vee B;\Delta,\Delta')\vdash^{*} C}\,\vee L^{a}$$

$$\dfrac{\dfrac{(\Gamma;\Delta)\vdash^{-} D \quad (\Gamma'',A;\Delta',D)\vdash^{*} C}{(\Gamma,\Gamma'',A;\Delta,\Delta')\vdash^{*} C}\,Cut^{c} \quad \dfrac{(\Gamma;\Delta)\vdash^{-} D \quad (\Gamma'',B;\Delta',D)\vdash^{*} C}{(\Gamma,\Gamma'',B;\Delta,\Delta')\vdash^{*} C}\,Cut^{c}}{(\Gamma,\Gamma'',A\vee B;\Delta,\Delta')\vdash^{*} C}\,\vee L^{a}$$

-4.4- $\vee L^c$ is the last rule used to derive the right premise with $\Delta' = \Delta'', A\vee B$. The derivations with cuts of cut-height $n + m + 1$ are

$$\dfrac{(\Gamma;\Delta)\vdash^{+} D \qquad \dfrac{(\Gamma',D;\Delta'',A,B)\vdash^{*} C}{(\Gamma',D;\Delta'',A\vee B)\vdash^{*} C}\,{\scriptstyle \vee L^{c}}}{(\Gamma,\Gamma';\Delta,\Delta'',A\vee B)\vdash^{*} C}\,{\scriptstyle Cut^{a}} \qquad \dfrac{(\Gamma;\Delta)\vdash^{-} D \qquad \dfrac{(\Gamma';\Delta'',A,B,D)\vdash^{*} C}{(\Gamma';\Delta'',A\vee B,D)\vdash^{*} C}\,{\scriptstyle \vee L^{c}}}{(\Gamma,\Gamma';\Delta,\Delta'',A\vee B)\vdash^{*} C}\,{\scriptstyle Cut^{c}}$$

These can be transformed into derivations with cuts of cut-height $n+m$:

$$\dfrac{\dfrac{(\Gamma;\Delta)\vdash^{+} D \qquad (\Gamma',D;\Delta'',A,B)\vdash^{*} C}{(\Gamma,\Gamma';\Delta,\Delta'',A,B)\vdash^{*} C}\,{\scriptstyle Cut^{a}}}{(\Gamma,\Gamma';\Delta,\Delta''A\vee B)\vdash^{*} C}\,{\scriptstyle \vee L^{c}} \qquad \dfrac{\dfrac{(\Gamma;\Delta)\vdash^{-} D \qquad (\Gamma';\Delta'',A,B,D)\vdash^{*} C}{(\Gamma,\Gamma';\Delta,\Delta'',A,B)\vdash^{*} C}\,{\scriptstyle Cut^{c}}}{(\Gamma,\Gamma';\Delta,\Delta'',A\vee B)\vdash^{*} C}\,{\scriptstyle \vee L^{c}}$$

-4.5- $\rightarrow L^{a}$ is the last rule used to derive the right premise with $\Gamma'=\Gamma'',A\rightarrow B$. The derivations with cuts of cut-height $n+max(m,k)+1$ are

$$\dfrac{(\Gamma;\Delta)\vdash^{+} D \qquad \dfrac{(\Gamma'',A\rightarrow B,D;\Delta')\vdash^{+} A \quad (\Gamma'',B,D;\Delta')\vdash^{*} C}{(\Gamma'',A\rightarrow B,D;\Delta')\vdash^{*} C}\,{\scriptstyle \rightarrow L^{a}}}{(\Gamma,\Gamma'',A\rightarrow B;\Delta,\Delta')\vdash^{*} C}\,{\scriptstyle Cut^{a}}$$

$$\dfrac{(\Gamma;\Delta)\vdash^{-} D \qquad \dfrac{(\Gamma'',A\rightarrow B;\Delta',D)\vdash^{+} A \quad (\Gamma'',B;\Delta',D)\vdash^{*} C}{(\Gamma'',A\rightarrow B;\Delta',D)\vdash^{*} C}\,{\scriptstyle \rightarrow L^{a}}}{(\Gamma,\Gamma'',A\rightarrow B;\Delta,\Delta')\vdash^{*} C}\,{\scriptstyle Cut^{c}}$$

These can be transformed into derivations each with two cuts of cut-height $n+m$ and $n+k$, respectively:

$$\dfrac{\dfrac{(\Gamma;\Delta)\vdash^{+} D \quad (\Gamma'',A\rightarrow B,D;\Delta')\vdash^{+} A}{(\Gamma,\Gamma'',A\rightarrow B;\Delta,\Delta')\vdash^{+} A}\,{\scriptstyle Cut^{a}} \qquad \dfrac{(\Gamma;\Delta)\vdash^{+} D \quad (\Gamma'',B,D;\Delta')\vdash^{*} C}{(\Gamma,\Gamma'',B;\Delta,\Delta')\vdash^{*} C}\,{\scriptstyle Cut^{a}}}{(\Gamma,\Gamma'',A\rightarrow B;\Delta,\Delta')\vdash^{*} C}\,{\scriptstyle \rightarrow L^{a}}$$

$$\dfrac{\dfrac{(\Gamma;\Delta)\vdash^{-} D \quad (\Gamma'',A\rightarrow B;\Delta',D)\vdash^{+} A}{(\Gamma,\Gamma'',A\rightarrow B;\Delta,\Delta')\vdash^{+} A}\,{\scriptstyle Cut^{c}} \qquad \dfrac{(\Gamma;\Delta)\vdash^{-} D \quad (\Gamma'',B;\Delta',D)\vdash^{*} C}{(\Gamma,\Gamma'',B;\Delta,\Delta')\vdash^{*} C}\,{\scriptstyle Cut^{c}}}{(\Gamma,\Gamma'',A\rightarrow B;\Delta,\Delta')\vdash^{*} C}\,{\scriptstyle \rightarrow L^{a}}$$

-4.6- $\rightarrow L^{c}$ is the last rule used to derive the right premise with $\Delta'=\Delta'',A\rightarrow B$. The derivations with cuts of cut-height $n+m+1$ are

$$\dfrac{(\Gamma;\Delta)\vdash^{+} D \qquad \dfrac{(\Gamma',A,D;\Delta'',B)\vdash^{*} C}{(\Gamma',D;\Delta'',A\rightarrow B)\vdash^{*} C}\,{\scriptstyle \rightarrow L^{c}}}{(\Gamma,\Gamma';\Delta,\Delta'',A\rightarrow B)\vdash^{*} C}\,{\scriptstyle Cut^{a}} \qquad \dfrac{(\Gamma;\Delta)\vdash^{-} D \qquad \dfrac{(\Gamma',A;\Delta'',B,D)\vdash^{*} C}{(\Gamma';\Delta'',A\rightarrow B,D)\vdash^{*} C}\,{\scriptstyle \rightarrow L^{c}}}{(\Gamma,\Gamma';\Delta,\Delta'',A\rightarrow B)\vdash^{*} C}\,{\scriptstyle Cut^{c}}$$

These can be transformed into derivations with cuts of cut-height $n+m$:

$$\dfrac{\dfrac{(\Gamma;\Delta)\vdash^{+} D \qquad (\Gamma',A,D;\Delta'',B)\vdash^{*} C}{(\Gamma,\Gamma',A;\Delta,\Delta'',B)\vdash^{*} C}\,{\scriptstyle Cut^{a}}}{(\Gamma,\Gamma';\Delta,\Delta'',A\rightarrow B)\vdash^{*} C}\,{\scriptstyle \rightarrow L^{c}} \qquad \dfrac{\dfrac{(\Gamma;\Delta)\vdash^{-} D \qquad (\Gamma',A;\Delta'',B,D)\vdash^{*} C}{(\Gamma,\Gamma',A;\Delta,\Delta'',B)\vdash^{*} C}\,{\scriptstyle Cut^{c}}}{(\Gamma,\Gamma';\Delta,\Delta'',A\rightarrow B)\vdash^{*} C}\,{\scriptstyle \rightarrow L^{c}}$$

-4.7- $\prec L^{a}$ is the last rule used to derive the right premise with $\Gamma'=\Gamma'',A\prec B$. The derivations with cuts of cut-height $n+m+1$ are

$$\dfrac{(\Gamma;\Delta)\vdash^{+} D \qquad \dfrac{(\Gamma'',A,D;\Delta',B)\vdash^{*} C}{(\Gamma'',A\prec B,D;\Delta')\vdash^{*} C}{\scriptstyle \prec L^a}}{(\Gamma,\Gamma'',A\prec B;\Delta,\Delta')\vdash^{*} C}{\scriptstyle Cut^a} \qquad \dfrac{(\Gamma;\Delta)\vdash^{-} D \qquad \dfrac{(\Gamma'',A;\Delta',B,D)\vdash^{*} C}{(\Gamma'',A\prec B;\Delta',D)\vdash^{*} C}{\scriptstyle \prec L^a}}{(\Gamma,\Gamma'',A\prec B;\Delta,\Delta')\vdash^{*} C}{\scriptstyle Cut^c}$$

These can be transformed into derivations with cuts of cut-height $n+m$:

$$\dfrac{\dfrac{(\Gamma;\Delta)\vdash^{+} D \qquad (\Gamma'',A,D;\Delta',B)\vdash^{*} C}{(\Gamma,\Gamma'',A;\Delta,\Delta',B)\vdash^{*} C}{\scriptstyle Cut^a}}{(\Gamma,\Gamma'',A\prec B;\Delta,\Delta')\vdash^{*} C}{\scriptstyle \prec L^a} \qquad \dfrac{\dfrac{(\Gamma;\Delta)\vdash^{-} D \qquad (\Gamma'',A;\Delta',B,D)\vdash^{*} C}{(\Gamma,\Gamma'',A;\Delta,\Delta',B)\vdash^{*} C}{\scriptstyle Cut^c}}{(\Gamma,\Gamma'',A\prec B;\Delta,\Delta')\vdash^{*} C}{\scriptstyle \prec L^a}$$

-4.8- $\prec L^c$ is the last rule used to derive the right premise with $\Delta' = \Delta'', A\prec B$. The derivations with cuts of cut-height $n + max(m,k) + 1$ are

$$\dfrac{(\Gamma;\Delta)\vdash^{+} D \qquad \dfrac{(\Gamma',D;\Delta'',A\prec B)\vdash^{-} B \quad (\Gamma',D;\Delta'',A)\vdash^{*} C}{(\Gamma',D;\Delta'',A\prec B)\vdash^{*} C}{\scriptstyle \prec L^c}}{(\Gamma,\Gamma';\Delta,\Delta'',A\prec B)\vdash^{*} C}{\scriptstyle Cut^a}$$

$$\dfrac{(\Gamma;\Delta)\vdash^{-} D \qquad \dfrac{(\Gamma';\Delta'',A\prec B,D)\vdash^{-} B \quad (\Gamma';\Delta'',A,D)\vdash^{*} C}{(\Gamma';\Delta'',A\prec B,D)\vdash^{*} C}{\scriptstyle \prec L^c}}{(\Gamma,\Gamma';\Delta,\Delta'',A\prec B)\vdash^{*} C}{\scriptstyle Cut^c}$$

These can be transformed into derivations each with two cuts of cut-height $n+m$ and $n+k$, respectively:

$$\dfrac{\dfrac{(\Gamma;\Delta)\vdash^{+} D \quad (\Gamma',D;\Delta'',A\prec B)\vdash^{-} B}{(\Gamma,\Gamma';\Delta,\Delta'',A\prec B)\vdash^{-} B}{\scriptstyle Cut^a} \qquad \dfrac{(\Gamma;\Delta)\vdash^{+} D \quad (\Gamma',D;\Delta'',A)\vdash^{*} C}{(\Gamma,\Gamma';\Delta,\Delta'',A)\vdash^{*} C}{\scriptstyle Cut^a}}{(\Gamma,\Gamma';\Delta,\Delta'',A\prec B)\vdash^{*} C}{\scriptstyle \prec L^c}$$

$$\dfrac{\dfrac{(\Gamma;\Delta)\vdash^{-} D \quad (\Gamma';\Delta'',A\prec B,D)\vdash^{-} B}{(\Gamma,\Gamma';\Delta,\Delta'',A\prec B)\vdash^{-} B}{\scriptstyle Cut^c} \qquad \dfrac{(\Gamma;\Delta)\vdash^{-} D \quad (\Gamma';\Delta'',A,D)\vdash^{*} C}{(\Gamma,\Gamma';\Delta,\Delta'',A)\vdash^{*} C}{\scriptstyle Cut^c}}{(\Gamma,\Gamma';\Delta,\Delta'',A\prec B)\vdash^{*} C}{\scriptstyle \prec L^c}$$

-4.9- $\wedge R^+$ is the last rule used to derive the right premise with $C = A\wedge B$. The derivations with cuts of cut-height $n + max(m,k) + 1$ are

$$\dfrac{(\Gamma;\Delta)\vdash^{+} D \qquad \dfrac{(\Gamma',D;\Delta')\vdash^{+} A \quad (\Gamma',D;\Delta')\vdash^{+} B}{(\Gamma',D;\Delta')\vdash^{+} A\wedge B}{\scriptstyle \wedge R^+}}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^{+} A\wedge B}{\scriptstyle Cut^a} \qquad \dfrac{(\Gamma;\Delta)\vdash^{-} D \qquad \dfrac{(\Gamma';\Delta',D)\vdash^{+} A \quad (\Gamma';\Delta',D)\vdash^{+} B}{(\Gamma';\Delta',D)\vdash^{+} A\wedge B}{\scriptstyle \wedge R^+}}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^{+} A\wedge B}{\scriptstyle Cut^c}$$

These can be transformed into derivations each with two cuts of cut-height $n+m$ and $n+k$, respectively:

$$\dfrac{\dfrac{(\Gamma;\Delta)\vdash^{+} D \quad (\Gamma',D;\Delta')\vdash^{+} A}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^{+} A}{\scriptstyle Cut^a} \qquad \dfrac{(\Gamma;\Delta)\vdash^{+} D \quad (\Gamma',D;\Delta')\vdash^{+} B}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^{+} B}{\scriptstyle Cut^a}}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^{+} A\wedge B}{\scriptstyle \wedge R^+}$$

$$\dfrac{\dfrac{(\Gamma;\Delta)\vdash^{-} D \quad (\Gamma';\Delta',D)\vdash^{+} A}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^{+} A}{\scriptstyle Cut^c} \qquad \dfrac{(\Gamma;\Delta)\vdash^{-} D \quad (\Gamma';\Delta',D)\vdash^{+} B}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^{+} B}{\scriptstyle Cut^c}}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^{+} A\wedge B}{\scriptstyle \wedge R^+}$$

-4.10.1- $\wedge R_1^-$ is the last rule used to derive the right premise with $C = A \wedge B$. The derivations with cuts of cut-height $n+m+1$ are

$$\dfrac{(\Gamma;\Delta)\vdash^+ D \qquad \dfrac{(\Gamma',D;\Delta')\vdash^- A}{(\Gamma',D;\Delta')\vdash^- A\wedge B}\,{\scriptstyle \wedge R_1^-}}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^- A\wedge B}\,{\scriptstyle Cut^a} \qquad \dfrac{(\Gamma;\Delta)\vdash^- D \qquad \dfrac{(\Gamma';\Delta',D)\vdash^- A}{(\Gamma';\Delta',D)\vdash^- A\wedge B}\,{\scriptstyle \wedge R_1^-}}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^- A\wedge B}\,{\scriptstyle Cut^c}$$

These can be transformed into derivations with cuts of cut-height $n+m$:

$$\dfrac{\dfrac{(\Gamma;\Delta)\vdash^+ D \qquad (\Gamma',D;\Delta')\vdash^- A}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^- A}\,{\scriptstyle Cut^a}}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^- A\wedge B}\,{\scriptstyle \wedge R_1^-} \qquad \dfrac{\dfrac{(\Gamma;\Delta)\vdash^- D \qquad (\Gamma';\Delta',D)\vdash^- A}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^- A}\,{\scriptstyle Cut^c}}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^- A\wedge B}\,{\scriptstyle \wedge R_1^-}$$

-4.10.2- $\wedge R_2^-$ is the last rule used to derive the right premise with $C = A \wedge B$. The derivations with cuts of cut-height $n+m+1$ are

$$\dfrac{(\Gamma;\Delta)\vdash^+ D \qquad \dfrac{(\Gamma',D;\Delta')\vdash^- B}{(\Gamma',D;\Delta')\vdash^- A\wedge B}\,{\scriptstyle \wedge R_2^-}}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^- A\wedge B}\,{\scriptstyle Cut^a} \qquad \dfrac{(\Gamma;\Delta)\vdash^- D \qquad \dfrac{(\Gamma';\Delta',D)\vdash^- B}{(\Gamma';\Delta',D)\vdash^- A\wedge B}\,{\scriptstyle \wedge R_2^-}}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^- A\wedge B}\,{\scriptstyle Cut^c}$$

These can be transformed into derivations with cuts of cut-height $n+m$:

$$\dfrac{\dfrac{(\Gamma;\Delta)\vdash^+ D \qquad (\Gamma',D;\Delta')\vdash^- B}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^- B}\,{\scriptstyle Cut^a}}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^- A\wedge B}\,{\scriptstyle \wedge R_2^-} \qquad \dfrac{\dfrac{(\Gamma;\Delta)\vdash^- D \qquad (\Gamma';\Delta',D)\vdash^- B}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^- B}\,{\scriptstyle Cut^c}}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^- A\wedge B}\,{\scriptstyle \wedge R_2^-}$$

-4.11.1- $\vee R_1^+$ is the last rule used to derive the right premise with $C = A \vee B$. The derivations with cuts of cut-height $n+m+1$ are

$$\dfrac{(\Gamma;\Delta)\vdash^+ D \qquad \dfrac{(\Gamma',D;\Delta')\vdash^+ A}{(\Gamma',D;\Delta')\vdash^+ A\vee B}\,{\scriptstyle \vee R_1^+}}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^+ A\vee B}\,{\scriptstyle Cut^a} \qquad \dfrac{(\Gamma;\Delta)\vdash^- D \qquad \dfrac{(\Gamma';\Delta',D)\vdash^+ A}{(\Gamma';\Delta',D)\vdash^+ A\vee B}\,{\scriptstyle \vee R_1^+}}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^+ A\vee B}\,{\scriptstyle Cut^c}$$

These can be transformed into derivations with cuts of cut-height $n+m$:

$$\dfrac{\dfrac{(\Gamma;\Delta)\vdash^+ D \qquad (\Gamma',D;\Delta')\vdash^+ A}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^+ A}\,{\scriptstyle Cut^a}}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^+ A\vee B}\,{\scriptstyle \vee R_1^+} \qquad \dfrac{\dfrac{(\Gamma;\Delta)\vdash^- D \qquad (\Gamma';\Delta',D)\vdash^+ A}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^+ A}\,{\scriptstyle Cut^c}}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^+ A\vee B}\,{\scriptstyle \vee R_1^+}$$

-4.11.2- $\vee R_2^+$ is the last rule used to derive the right premise with $C = A \vee B$. The derivations with cuts of cut-height $n+m+1$ are

$$\dfrac{(\Gamma;\Delta)\vdash^+ D \qquad \dfrac{(\Gamma',D;\Delta')\vdash^+ B}{(\Gamma',D;\Delta')\vdash^+ A\vee B}\,{\scriptstyle \vee R_2^+}}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^+ A\vee B}\,{\scriptstyle Cut^a} \qquad \dfrac{(\Gamma;\Delta)\vdash^- D \qquad \dfrac{(\Gamma';\Delta',D)\vdash^+ B}{(\Gamma';\Delta',D)\vdash^+ A\vee B}\,{\scriptstyle \vee R_2^+}}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^+ A\vee B}\,{\scriptstyle Cut^c}$$

These can be transformed into derivations with cuts of cut-height $n+m$:

$$\dfrac{\dfrac{(\Gamma;\Delta)\vdash^+ D \qquad (\Gamma',D;\Delta')\vdash^+ B}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^+ B}\,{\scriptstyle Cut^a}}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^+ A\vee B}\,{\scriptstyle \vee R_2^+} \qquad \dfrac{\dfrac{(\Gamma;\Delta)\vdash^- D \qquad (\Gamma';\Delta',D)\vdash^+ B}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^+ B}\,{\scriptstyle Cut^c}}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^+ A\vee B}\,{\scriptstyle \vee R_2^+}$$

-4.12- $\vee R^-$ is the last rule used to derive the right premise with $C = A \vee B$. The derivations with cuts of cut-height $n + max(m,k) + 1$ are

$$\dfrac{(\Gamma;\Delta) \vdash^+ D \quad \dfrac{(\Gamma',D;\Delta') \vdash^- A \quad (\Gamma',D;\Delta') \vdash^- B}{(\Gamma',D;\Delta') \vdash^- A \vee B}\,{\scriptstyle \vee R^-}}{(\Gamma,\Gamma';\Delta,\Delta') \vdash^- A \vee B}\,{\scriptstyle Cut^a}$$

$$\dfrac{(\Gamma;\Delta) \vdash^- D \quad \dfrac{(\Gamma';\Delta',D) \vdash^- A \quad (\Gamma';\Delta',D) \vdash^- B}{(\Gamma';\Delta',D) \vdash^- A \vee B}\,{\scriptstyle \vee R^-}}{(\Gamma,\Gamma';\Delta,\Delta') \vdash^- A \vee B}\,{\scriptstyle Cut^c}$$

These can be transformed into derivations each with two cuts of cut-height $n + m$ and $n + k$, respectively:

$$\dfrac{\dfrac{(\Gamma;\Delta) \vdash^+ D \quad (\Gamma',D;\Delta') \vdash^- A}{(\Gamma,\Gamma';\Delta,\Delta') \vdash^- A}\,{\scriptstyle Cut^a} \qquad \dfrac{(\Gamma;\Delta) \vdash^+ D \quad (\Gamma',D;\Delta') \vdash^- B}{(\Gamma,\Gamma';\Delta,\Delta') \vdash^- B}\,{\scriptstyle Cut^a}}{(\Gamma,\Gamma';\Delta,\Delta') \vdash^- A \vee B}\,{\scriptstyle \vee R^-}$$

$$\dfrac{\dfrac{(\Gamma;\Delta) \vdash^- D \quad (\Gamma';\Delta',D) \vdash^- A}{(\Gamma,\Gamma';\Delta,\Delta') \vdash^- A}\,{\scriptstyle Cut^c} \qquad \dfrac{(\Gamma;\Delta) \vdash^- D \quad (\Gamma';\Delta',D) \vdash^- B}{(\Gamma,\Gamma';\Delta,\Delta') \vdash^- B}\,{\scriptstyle Cut^c}}{(\Gamma,\Gamma';\Delta,\Delta') \vdash^- A \vee B}\,{\scriptstyle \vee R^-}$$

-4.13- $\to R^+$ is the last rule used to derive the right premise with $C = A \to B$. The derivations with cuts of cut-height $n + m + 1$ are

$$\dfrac{(\Gamma;\Delta) \vdash^+ D \quad \dfrac{(\Gamma',A,D;\Delta') \vdash^+ B}{(\Gamma',D;\Delta') \vdash^+ A \to B}\,{\scriptstyle \to R^+}}{(\Gamma,\Gamma';\Delta,\Delta') \vdash^+ A \to B}\,{\scriptstyle Cut^a} \qquad \dfrac{(\Gamma;\Delta) \vdash^- D \quad \dfrac{(\Gamma',A;\Delta',D) \vdash^+ B}{(\Gamma';\Delta',D) \vdash^+ A \to B}\,{\scriptstyle \to R^+}}{(\Gamma,\Gamma';\Delta,\Delta') \vdash^+ A \to B}\,{\scriptstyle Cut^c}$$

These can be transformed into derivations with cuts of cut-height $n + m$:

$$\dfrac{\dfrac{(\Gamma;\Delta) \vdash^+ D \quad (\Gamma',A,D;\Delta') \vdash^+ B}{(\Gamma,\Gamma',A;\Delta,\Delta') \vdash^+ B}\,{\scriptstyle Cut^a}}{(\Gamma,\Gamma';\Delta,\Delta') \vdash^+ A \to B}\,{\scriptstyle \to R^+} \qquad \dfrac{\dfrac{(\Gamma;\Delta) \vdash^- D \quad (\Gamma',A;\Delta',D) \vdash^+ B}{(\Gamma,\Gamma',A;\Delta,\Delta') \vdash^+ B}\,{\scriptstyle Cut^c}}{(\Gamma,\Gamma';\Delta,\Delta') \vdash^+ A \to B}\,{\scriptstyle \to R^+}$$

-4.14- $\to R^-$ is the last rule used to derive the right premise with $C = A \to B$. The derivations with cuts of cut-height $n + max(m,k) + 1$ are

$$\dfrac{(\Gamma;\Delta) \vdash^+ D \quad \dfrac{(\Gamma',D;\Delta') \vdash^+ A \quad (\Gamma',D;\Delta') \vdash^- B}{(\Gamma',D;\Delta') \vdash^- A \to B}\,{\scriptstyle \to R^-}}{(\Gamma,\Gamma';\Delta,\Delta') \vdash^- A \to B}\,{\scriptstyle Cut^a} \qquad \dfrac{(\Gamma;\Delta) \vdash^- D \quad \dfrac{(\Gamma';\Delta',D) \vdash^+ A \quad (\Gamma';\Delta',D) \vdash^- B}{(\Gamma';\Delta',D) \vdash^- A \to B}\,{\scriptstyle \to R^-}}{(\Gamma,\Gamma';\Delta,\Delta') \vdash^- A \to B}\,{\scriptstyle Cut^c}$$

These can be transformed into derivations each with two cuts of cut-height $n + m$ and $n + k$, respectively:

$$\dfrac{\dfrac{(\Gamma;\Delta) \vdash^+ D \quad (\Gamma',D;\Delta') \vdash^+ A}{(\Gamma,\Gamma';\Delta,\Delta') \vdash^+ A}\,{\scriptstyle Cut^a} \qquad \dfrac{(\Gamma;\Delta) \vdash^+ D \quad (\Gamma',D;\Delta') \vdash^- B}{(\Gamma,\Gamma';\Delta,\Delta') \vdash^- B}\,{\scriptstyle Cut^a}}{(\Gamma,\Gamma';\Delta,\Delta') \vdash^- A \to B}\,{\scriptstyle \to R^-}$$

$$\dfrac{\dfrac{(\Gamma;\Delta) \vdash^- D \quad (\Gamma';\Delta',D) \vdash^+ A}{(\Gamma,\Gamma';\Delta,\Delta') \vdash^+ A}\,{\scriptstyle Cut^c} \qquad \dfrac{(\Gamma;\Delta) \vdash^- D \quad (\Gamma';\Delta',D) \vdash^- B}{(\Gamma,\Gamma';\Delta,\Delta') \vdash^- B}\,{\scriptstyle Cut^c}}{(\Gamma,\Gamma';\Delta,\Delta') \vdash^- A \to B}\,{\scriptstyle \to R^-}$$

-4.15- $\prec R^+$ is the last rule used to derive the right premise with $C = A \prec B$. The derivations with cuts of cut-height $n + max(m, k) + 1$ are

$$\dfrac{(\Gamma;\Delta) \vdash^+ D \qquad \dfrac{(\Gamma', D;\Delta') \vdash^+ A \quad (\Gamma', D;\Delta') \vdash^- B}{(\Gamma', D;\Delta') \vdash^+ A \prec B}\,{\scriptstyle \prec R^+}}{(\Gamma,\Gamma';\Delta,\Delta') \vdash^+ A \prec B}\,{\scriptstyle Cut^a} \qquad \dfrac{(\Gamma;\Delta) \vdash^- D \qquad \dfrac{(\Gamma';\Delta', D) \vdash^+ A \quad (\Gamma';\Delta', D) \vdash^- B}{(\Gamma';\Delta', D) \vdash^+ A \prec B}\,{\scriptstyle \prec R^+}}{(\Gamma,\Gamma';\Delta,\Delta') \vdash^+ A \prec B}\,{\scriptstyle Cut^c}$$

These can be transformed into derivations each with two cuts of cut-height $n + m$ and $n + k$, respectively:

$$\dfrac{\dfrac{(\Gamma;\Delta) \vdash^+ D \quad (\Gamma', D;\Delta') \vdash^+ A}{(\Gamma,\Gamma';\Delta,\Delta') \vdash^+ A}\,{\scriptstyle Cut^a} \qquad \dfrac{(\Gamma;\Delta) \vdash^+ D \quad (\Gamma', D;\Delta') \vdash^- B}{(\Gamma,\Gamma';\Delta,\Delta') \vdash^- B}\,{\scriptstyle Cut^a}}{(\Gamma,\Gamma';\Delta,\Delta') \vdash^+ A \prec B}\,{\scriptstyle \prec R^+}$$

$$\dfrac{\dfrac{(\Gamma;\Delta) \vdash^- D \quad (\Gamma';\Delta', D) \vdash^+ A}{(\Gamma,\Gamma';\Delta,\Delta') \vdash^+ A}\,{\scriptstyle Cut^c} \qquad \dfrac{(\Gamma;\Delta) \vdash^- D \quad (\Gamma';\Delta', D) \vdash^- B}{(\Gamma,\Gamma';\Delta,\Delta') \vdash^- B}\,{\scriptstyle Cut^c}}{(\Gamma,\Gamma';\Delta,\Delta') \vdash^+ A \prec B}\,{\scriptstyle \prec R^+}$$

-4.16- $\prec R^-$ is the last rule used to derive the right premise with $C = A \prec B$. The derivations with cuts of cut-height $n + m + 1$ are

$$\dfrac{(\Gamma;\Delta) \vdash^+ D \qquad \dfrac{(\Gamma', D;\Delta', B) \vdash^- A}{(\Gamma', D;\Delta') \vdash^- A \prec B}\,{\scriptstyle \prec R^-}}{(\Gamma,\Gamma';\Delta,\Delta') \vdash^- A \prec B}\,{\scriptstyle Cut^a} \qquad \dfrac{(\Gamma;\Delta) \vdash^- D \qquad \dfrac{(\Gamma';\Delta', B, D) \vdash^- A}{(\Gamma';\Delta', D) \vdash^- A \prec B}\,{\scriptstyle \prec R^-}}{(\Gamma,\Gamma';\Delta,\Delta') \vdash^- A \prec B}\,{\scriptstyle Cut^c}$$

These can be transformed into derivations with cuts of cut-height $n + m$:

$$\dfrac{\dfrac{(\Gamma;\Delta) \vdash^+ D \qquad (\Gamma', D;\Delta', B) \vdash^- A}{(\Gamma,\Gamma';\Delta,\Delta', B) \vdash^- A}\,{\scriptstyle Cut^a}}{(\Gamma,\Gamma';\Delta,\Delta') \vdash^- A \prec B}\,{\scriptstyle \prec R^-} \qquad \dfrac{\dfrac{(\Gamma;\Delta) \vdash^- D \qquad (\Gamma';\Delta', B, D) \vdash^- A}{(\Gamma,\Gamma';\Delta,\Delta', B) \vdash^- A}\,{\scriptstyle Cut^c}}{(\Gamma,\Gamma';\Delta,\Delta') \vdash^- A \prec B}\,{\scriptstyle \prec R^-}$$

It is shown that cut-height is reduced in all cases.

### -5- Cut formula *D* principal in both premises

For each cut rule four cases can be distinguished. Here, it can be shown for each case that the derivations can be transformed into ones in which the occurrences of cut have a reduced cut-height or the cut formula has a lower weight (or both).

-5.1- $D = A \wedge B$. The derivation for $Cut^a$ with a cut of cut-height $max(n, m) + 1 + k + 1$ is

$$\dfrac{\dfrac{(\Gamma;\Delta) \vdash^+ A \quad (\Gamma;\Delta) \vdash^+ B}{(\Gamma;\Delta) \vdash^+ A \wedge B}\,{\scriptstyle \wedge R^+} \qquad \dfrac{(\Gamma', A, B;\Delta') \vdash^* C}{(\Gamma', A \wedge B;\Delta') \vdash^* C}\,{\scriptstyle \wedge L^a}}{(\Gamma,\Gamma';\Delta,\Delta') \vdash^* C}\,{\scriptstyle Cut^a}$$

and can be transformed into a derivation with two cuts of cut-height (from top to bottom) $n + k$ and $m + max(n, k) + 1$:

$$\dfrac{(\Gamma;\Delta)\vdash^{+} B \qquad \dfrac{(\Gamma;\Delta)\vdash^{+} A \qquad (\Gamma',A,B;\Delta')\vdash^{*} C}{(\Gamma,\Gamma',B;\Delta,\Delta')\vdash^{*} C}\,Cut^a}{\dfrac{(\Gamma,\Gamma,\Gamma';\Delta,\Delta,\Delta')\vdash^{*} C}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^{*} C}\,C_{a/c}}\,Cut^a$$

Note that in both cases the weight of the cut formula is reduced. The upper cut is also reduced in height, while with the lower cut we have a case where cut-height is not necessarily reduced.

The possible derivations for $Cut^c$ with a cut of cut-height $n+1+max(m,k)+1$ are

$$\dfrac{\dfrac{(\Gamma;\Delta)\vdash^{-} A}{(\Gamma;\Delta)\vdash^{-} A\wedge B}\,\wedge R_1^{-} \qquad \dfrac{(\Gamma';\Delta',A)\vdash^{*} C \quad (\Gamma';\Delta',B)\vdash^{*} C}{(\Gamma';\Delta',A\wedge B)\vdash^{*} C}\,\wedge L^c}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^{*} C}\,Cut^c$$

or

$$\dfrac{\dfrac{(\Gamma;\Delta)\vdash^{-} B}{(\Gamma;\Delta)\vdash^{-} A\wedge B}\,\wedge R_2^{-} \qquad \dfrac{(\Gamma';\Delta',A)\vdash^{*} C \quad (\Gamma';\Delta',B)\vdash^{*} C}{(\Gamma';\Delta',A\wedge B)\vdash^{*} C}\,\wedge L^c}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^{*} C}\,Cut^c$$

and those can be transformed into derivations with cuts of cut-height $n+m$ or $n+k$, respectively:

$$\dfrac{(\Gamma;\Delta)\vdash^{-} A \qquad (\Gamma';\Delta',A)\vdash^{*} C}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^{*} C}\,Cut^c \qquad \dfrac{(\Gamma;\Delta)\vdash^{-} B \qquad (\Gamma';\Delta',B)\vdash^{*} C}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^{*} C}\,Cut^c$$

Here, both cut-height and weight of the cut formulas are reduced.

-5.2- $D = A\vee B$. The possible derivations for $Cut^a$ with a cut of cut-height $n+1+max(m,k)+1$ are

$$\dfrac{\dfrac{(\Gamma;\Delta)\vdash^{+} A}{(\Gamma;\Delta)\vdash^{+} A\vee B}\,\vee R_1^{+} \qquad \dfrac{(\Gamma',A;\Delta')\vdash^{*} C \quad (\Gamma',B;\Delta')\vdash^{*} C}{(\Gamma',A\vee B;\Delta')\vdash^{*} C}\,\vee L^a}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^{*} C}\,Cut^a$$

or

$$\dfrac{\dfrac{(\Gamma;\Delta)\vdash^{+} B}{(\Gamma;\Delta)\vdash^{+} A\vee B}\,\vee R_2^{+} \qquad \dfrac{(\Gamma',A;\Delta')\vdash^{*} C \quad (\Gamma',B;\Delta')\vdash^{*} C}{(\Gamma',A\vee B;\Delta')\vdash^{*} C}\,\vee L^a}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^{*} C}\,Cut^a$$

and those can be transformed into derivations with cuts of cut-height $n+m$ and $n+k$, respectively:

$$\dfrac{(\Gamma;\Delta)\vdash^{+} A \qquad (\Gamma',A;\Delta')\vdash^{*} C}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^{*} C}\,Cut^a \qquad \dfrac{(\Gamma;\Delta)\vdash^{+} B \qquad (\Gamma',B;\Delta')\vdash^{*} C}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^{*} C}\,Cut^a$$

Again, both cut-height and weight of the cut formulas are reduced.

The derivation for $Cut^c$ with a cut of cut-height $max(n,m)+1+k+1$ is

$$\dfrac{\dfrac{(\Gamma;\Delta)\vdash^- A \quad (\Gamma;\Delta)\vdash^- B}{(\Gamma;\Delta)\vdash^- A\vee B}\vee R^- \qquad \dfrac{(\Gamma';\Delta',A,B)\vdash^* C}{(\Gamma';\Delta',A\vee B)\vdash^* C}\vee L^c}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^* C}\,Cut^c$$

and can be transformed into a derivation with two cuts of cut-height $n+k$ and $m+max(n,k)+1$:

$$\dfrac{\dfrac{(\Gamma;\Delta)\vdash^- B \qquad \dfrac{(\Gamma;\Delta)\vdash^- A \quad (\Gamma';\Delta',A,B)\vdash^* C}{(\Gamma,\Gamma';\Delta,\Delta',B)\vdash^* C}\,Cut^c}{(\Gamma,\Gamma,\Gamma';\Delta,\Delta,\Delta')\vdash^* C}\,Cut^c}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^* C}\,C^{a/c}$$

Note that again, in the case of the lower cut, although the cut-height might increase, the weight of the cut formula is reduced. For the upper cut both cut-height and weight of the cut formula is reduced.

-5.3- $D = A \to B$. The derivation for $Cut^a$ with a cut of cut-height $n+1+max(m,k)+1$ is

$$\dfrac{\dfrac{(\Gamma,A;\Delta)\vdash^+ B}{(\Gamma;\Delta)\vdash^+ A\to B}\to R^+ \qquad \dfrac{(\Gamma',A\to B;\Delta')\vdash^+ A \quad (\Gamma',B;\Delta')\vdash^* C}{(\Gamma',A\to B;\Delta')\vdash^* C}\to L^a}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^* C}\,Cut^a$$

and this can be transformed into a derivation with three cuts of cut-height (from left to right and from top to bottom) $n+1+m$, $n+k$, and $max(n+1,m)+1+max(n,k)+1$ respectively:

$$\dfrac{\dfrac{\dfrac{\dfrac{(\Gamma,A;\Delta)\vdash^+ B}{(\Gamma;\Delta)\vdash^+ A\to B}\to R^+ \quad (\Gamma',A\to B;\Delta')\vdash^+ A}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^+ A}\,Cut^a \qquad \dfrac{(\Gamma,A;\Delta)\vdash^+ B \quad (\Gamma',B;\Delta')\vdash^* C}{(\Gamma,A,\Gamma';\Delta,\Delta')\vdash^* C}\,Cut^a}{(\Gamma,\Gamma,\Gamma',\Gamma';\Delta,\Delta,\Delta',\Delta')\vdash^* C}\,Cut^a}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^* C}\,C^{a/c}$$

In the first case cut-height is reduced, in the second case cut-height and weight of the cut formula is reduced and in the third case weight of the cut formula is reduced.

The derivation for $Cut^c$ with a cut of cut-height $max(n,m)+1+k+1$ is

$$\dfrac{\dfrac{(\Gamma;\Delta)\vdash^+ A \quad (\Gamma;\Delta)\vdash^- B}{(\Gamma;\Delta)\vdash^- A\to B}\to R^- \qquad \dfrac{(\Gamma',A;\Delta',B)\vdash^* C}{(\Gamma';\Delta',A\to B)\vdash^* C}\to L^c}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^* C}\,Cut^c$$

This can be transformed into a derivation with two cuts of cut-height $n+k$ and $m+max(n,k)+1$:

$$\dfrac{\dfrac{(\Gamma;\Delta)\vdash^{-} B \qquad \dfrac{(\Gamma;\Delta)\vdash^{+} A \qquad (\Gamma', A;\Delta', B)\vdash^{*} C}{(\Gamma,\Gamma';\Delta,\Delta', B)\vdash^{*} C}\,Cut^{a}}{(\Gamma,\Gamma,\Gamma';\Delta,\Delta,\Delta')\vdash^{*} C}\,Cut^{c}}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^{*} C}\,C^{a/c}$$

In the first case cut-height and weight of the cut formula is reduced, while in the second case the weight of the cut formula is reduced. Here we can observe a result specific for this calculus due to the mixture of derivability relations $\vdash^{+}$ and $\vdash^{-}$ in $\rightarrow R^{-}$ and the position of the active formulas in the assumptions *and* in the counterassumptions in $\rightarrow L^{c}$: Derivations containing instances of $Cut^{c}$ are not necessarily transformed into derivations with a lesser cut-height or a reduced weight of the cut formula of another instance of $Cut^{c}$ but it can also happen that $Cut^{c}$ is replaced by $Cut^{a}$.

-5.4- $D = A \prec B$. The derivation for $Cut^{a}$ with a cut of cut-height $max(n,m)+1+k+1$ is

$$\dfrac{\dfrac{(\Gamma;\Delta)\vdash^{+} A \quad (\Gamma;\Delta)\vdash^{-} B}{(\Gamma;\Delta)\vdash^{+} A\prec B}\,{\prec}R^{+} \qquad \dfrac{(\Gamma', A;\Delta', B)\vdash^{*} C}{(\Gamma', A\prec B;\Delta')\vdash^{*} C}\,{\prec}L^{a}}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^{*} C}\,Cut^{a}$$

This can be transformed into a derivation with two cuts of cut-height $n+k$ and $m+max(n,k)+1$:

$$\dfrac{\dfrac{(\Gamma;\Delta)\vdash^{-} B \qquad \dfrac{(\Gamma;\Delta)\vdash^{+} A \qquad (\Gamma', A;\Delta', B)\vdash^{*} C}{(\Gamma,\Gamma';\Delta,\Delta', B)\vdash^{*} C}\,Cut^{a}}{(\Gamma,\Gamma,\Gamma';\Delta,\Delta,\Delta')\vdash^{*} C}\,Cut^{c}}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^{*} C}\,C^{a/c}$$

Again, due to the mixture of derivability relations $\vdash^{+}$ and $\vdash^{-}$ in ${\prec}R^{+}$ and the presence of the active formulas both in assumptions and counterassumptions in ${\prec}L^{a}$, in this case $Cut^{a}$ can be replaced by instances of $Cut^{c}$ with a reduced weight of the cut formula. In the upper cut we have a reduction of both cut-height and weight of the cut formula.

The derivation for $Cut^{c}$ with a cut of cut-height $n+1+max(m,k)+1$ is

$$\dfrac{\dfrac{(\Gamma;\Delta, B)\vdash^{-} A}{(\Gamma;\Delta)\vdash^{-} A\prec B}\,{\prec}R^{-} \qquad \dfrac{(\Gamma';\Delta', A\prec B)\vdash^{-} B \quad (\Gamma';\Delta', A)\vdash^{*} C}{(\Gamma';\Delta', A\prec B)\vdash^{*} C}\,{\prec}L^{c}}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^{*} C}\,Cut^{c}$$

and this can be transformed into a derivation with three cuts of cut-height (from left to right and from top to bottom) $n+1+m$, $n+k$, and $max(n+1,m)+1+max(n,k)+1$ respectively:

$$\dfrac{\dfrac{\dfrac{\dfrac{(\Gamma;\Delta,B)\vdash^{-} A}{(\Gamma;\Delta)\vdash^{-} A \prec B}{\scriptstyle \prec R^{-}} \quad (\Gamma';\Delta',A\prec B)\vdash^{-} B}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^{-} B}{\scriptstyle Cut^{c}} \qquad \dfrac{(\Gamma;\Delta,B)\vdash^{-} A \quad (\Gamma';\Delta',A)\vdash^{*} C}{(\Gamma,\Gamma';\Delta,B,\Delta')\vdash^{*} C}{\scriptstyle Cut^{c}}}{(\Gamma,\Gamma,\Gamma',\Gamma';\Delta,\Delta,\Delta',\Delta')\vdash^{*} C}{\scriptstyle Cut^{c}}}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^{*} C}{\scriptstyle C^{a/c}}$$

In the first case cut-height is reduced, in the second case cut-height and weight of the cut formula and in the third case weight of the cut formula.

□

## 4.4 Conclusion

By applying the proof methods that Negri and von Plato (2001) use for their calculus `G3ip`, we were able to show that `SC2Int` is a cut-free sequent calculus for the bi-intuitionistic logic `2Int`. A proof can be given for the admissibility of the structural rules of weakening, contraction and cut in the system.

# 5 Uniqueness of logical connectives in a bilateralist setting

## 5.1 Introduction

The question of uniqueness is the question whether a connective is characterized by the rules governing its use in a way that there is *at most* one connective playing its specific inferential role. The usual way to test this is to create a 'copy-cat' connective governed by the same rules and show that formulas containing these connectives are interderivable. Important work has been conducted showing the problematic features of certain logics leading to the failure of uniqueness for some connectives in these systems, along with refinements of the requirements for uniqueness (see Section 5.3.2). In this paper I will deal with bilateralist proof systems, more specifically with proof systems for the logic `2Int`, which are bilateral in that they display two consequence relations: one for provability and one for dual provability (see Section 5.2.2). In such a setting, according to the common understanding of uniqueness, the question could be raised, whether this bilateralist proof-theoretic semantics (PTS) framework does not lead to different meanings depending on whether we prove or refute. Making this problem and my solution fully understandable requires laying some groundwork on bilateralism (see Section 5.2.1) and uniqueness (see Section 5.3.1) first. My aim is to show that the problems occurring in a bilateralist setting extend the problematic settings and solutions to ensure uniqueness that have been detected so far. Based on this analysis, I will propose a modification of our characterization of uniqueness that enables us to deal with uniqueness in bilateralism (see Section 5.3.3). Finally, I will point out what these considerations may imply when it comes to evaluating different (related) proof systems (see Section 5.3.4).

## 5.2 Bilateralism

### 5.2.1 Bilateralism and proof-theoretic semantics

The topic of bilateralism has received more and more attention in different areas within the past years including the area of PTS. In a nutshell, bilateralism is the view that dual concepts like truth and falsity, assertion and denial, or, in our context, proof and refutation should each be considered equally important, and not, like it is traditionally done, to concentrate solely on the former concepts. The debate started out in the context of considerations regarding an approach to the meaning of logical connectives, called "proof-theoretic semantics".[67] In PTS, situated in the broader context of *inferentialism*, the meaning of logical connectives is determined

[67] See (Schroeder-Heister, 2022) for an extensive overview of this area, as well as (Francez, 2015), which also covers the relation to bilateralism.

by the rules of inference that govern their use in proofs. Bilateralism is, then, an approach to meaning which questions the established view, famously held by Frege (1993[1918/1919]) and especially endorsed by Dummett (e.g., in 1976; 1981; 1991), that denying a proposition $A$ is equal to asserting the negation of $A$.[68] This has been opposed by several authors claiming that denial is a concept prior to negation and hence, should not be analyzed in terms of it (e.g., Martin-Löf, 1996; Restall, 2005). Thus, bilateralism demands an equal consideration of these dual concepts in that they should *both* be taken as primitive concepts, i.e., not reducible to each other.

Applying this to the proof-theoretic context, this amounts to demanding a proof system not only to characterize the proof (or verification) conditions of connectives but also their refutation (or falsification) conditions. Traditionally, in proof systems like natural deduction systems, the focus is only on the former, whereas, if we consider these notions to be on a par, we need to extend these systems with rules that capture falsification conditions. This is what Rumfitt (2000) proposes in his seminal paper connecting bilateralism and PTS, in which he introduces a natural deduction system with signed formulas for assertion and denial. Wansing (2017) goes one step further and argues that considering the speech acts of assertion and denial as well as their internally corresponding attitudes of judgment and dual judgment on a par, gives rise to also considering a consequence relation dual to our usual consequence relation. He claims that, in order to take bilateralism seriously in the context of proof theory, we need to embed this principle of duality on a level deeper than that of formulas: Next to our usual consequence relation ($\vdash^+$), which captures the notion of verification from premises to conclusion, we also need to consider a *dual* consequence relation ($\vdash^-$) capturing the dual notion of falsification from premises to conclusion.[69]

### 5.2.2 Bilateralist calculi: `N2Int` and `SC2Int`

To realize this, Wansing (2017) devises a natural deduction system for the bi-intuitionistic logic `2Int`, which comprises not only proofs (indicated by using single lines) but also *dual proofs* (indicated by using double lines). Also, a distinction is drawn in the premises between *assumptions* (taken to be true) and *counterassumptions* (taken to be false). This is indicated by an ordered pair $(\Gamma; \Delta)$ (with $\Gamma$ and $\Delta$ being finite, possibly empty multisets) of assumptions ($\Gamma$) and counterassumptions ($\Delta$). Single square brackets denote a possible discharge of assumptions, while double square brackets denote a possible discharge of counterassumptions. The language

[68] For an analysis of the established view as well as different ways to tackle it, see also (Ripley, 2011).

[69] In the spirit of Hacking's (1979, p. 292) conception of the sequent calculus as a metatheory, I use "$\vdash^+$" and "$\vdash^-$" both when talking about consequence relations in the metalanguage as well as for the sequent signs in the sequent calculus system which I will introduce below.

$\mathscr{L}_{2Int}$ of `2Int`, as given by Wansing, is defined in Backus-Naur form as follows:
$A ::= p \mid \bot \mid \top \mid (A \wedge A) \mid (A \vee A) \mid (A \rightarrow A) \mid (A \prec A)$.
I will in general use $p, q, r, ...$ for atomic formulas, $A, B, C, ...$ for arbitrary formulas, and $\Gamma, \Delta, \Gamma', ...$ for multisets of formulas. In a rule, the formula containing the respective connective of that rule is called the *principal formula*, while its components mentioned explicitly in the premises are called the *active formulas*.

As can be seen, we have a non-standard connective in this language, namely the operator of co-implication $\prec$,[70] which acts as a dual to implication, just like conjunction and disjunction can be seen as dual connectives. With that we are in the realms of so-called *bi-intuitionistic logic*, which is a conservative extension of intuitionistic logic by co-implication. Note that there is also a use of "bi-intuitionistic logic" in the literature to refer to a specific system, namely `BiInt`, also called "Heyting-Brouwer logic". Co-implication is there to be understood to internalize the preservation of non-truth from the conclusion to the premises in a valid inference. The system `2Int`, which is treated here, uses the same language as `BiInt`, but the meaning of co-implication differs in that it internalizes the preservation of falsity from the premises to the conclusion in a dually valid inference (Wansing, 2016a, 2016c, 2017, p. 30ff.).

From the viewpoint of bilateralism, i.e., considering falsificationism being on a par with verificationism, it is quite natural to extend our language by a connective for co-implication. The reason for this is that co-implication plays the same role in falsificationism as implication in verificationism: Both can be understood to express a concept of entailment in the object language. If we expect $\vdash^+$ to capture verification from the premises to the conclusion in a valid inference and $\vdash^-$ to capture falsification from the premises to the conclusion in a dually valid inference, then, just like implication internalizes provability in that we have in our system $(A; \emptyset) \vdash^+ B$ iff $(\emptyset; \emptyset) \vdash^+ A \rightarrow B$, likewise co-implication internalizes dual provability in that we have $(\emptyset; A) \vdash^- B$ iff $(\emptyset; \emptyset) \vdash^- B \prec A$.

With the two implication connectives also two negation connectives are defined: intuitionistic negation with $\neg A := A \rightarrow \bot$ and co-negation with $-A := \top \prec A$. Concerning switching between proofs and dual proofs, there is a division of labor between those negations in that we can move from proofs to dual proofs with intuitionistic negation and from dual proofs to proofs with co-negation: $(\Gamma; \Delta) \vdash^+ A$ iff $(\Gamma; \Delta) \vdash^- \neg A$ and $(\Gamma; \Delta) \vdash^- A$ iff $(\Gamma; \Delta) \vdash^+ -A$.[71]

Besides the usual introduction and elimination rules (henceforth: the *proof rules*) for intuitionistic logic, the natural deduction system `N2Int`, which is presented below, also contains rules that allow us to introduce and eliminate our connectives

[70] Sometimes also called "pseudo-difference", e.g., in (Rauszer, 1974), or "subtraction", e.g., in (Restall, 1997), and used with different symbols.

[71] I will not consider negation further in this paper, since I am concerned with connectives which are defined by their rules. See (Wansing, 2016a, 2016c, 2017) for a more detailed discussion, though.

into and from dual proofs. These so-called *dual proof rules* are obtained by a dualization of the proof rules (for the description and the rules of the calculus, see Wansing, 2017, p. 32-34) and having these two independent sets of rules is exactly what reflects the bilateralism of the proof system.

`N2Int`

$$\dfrac{\begin{array}{c}(\Gamma;\Delta)\\ \vdots\\ \overline{\bot}\end{array}}{A}\ \bot E \qquad \begin{array}{c}(\Gamma;\Delta)\\ \vdots\\ \overline{\overline{\top}}\\ \hline\hline A\end{array}\ \top E^d$$

$$\dfrac{\begin{array}{c}(\Gamma;\Delta)\\ \vdots\\ \overline{A}\end{array} \quad \begin{array}{c}(\Gamma';\Delta')\\ \vdots\\ \overline{B}\end{array}}{A \wedge B}\ \wedge I \qquad \dfrac{\begin{array}{c}(\Gamma;\Delta)\\ \vdots\\ \overline{A \wedge B}\end{array}}{A}\ \wedge E_1 \qquad \dfrac{\begin{array}{c}(\Gamma;\Delta)\\ \vdots\\ \overline{A \wedge B}\end{array}}{B}\ \wedge E_2$$

$$\begin{array}{c}(\Gamma;\Delta)\\ \vdots\\ \overline{\overline{A}}\\ \hline\hline A \wedge B\end{array}\ \wedge I^d_1 \qquad \begin{array}{c}(\Gamma;\Delta)\\ \vdots\\ \overline{\overline{B}}\\ \hline\hline A \wedge B\end{array}\ \wedge I^d_2 \qquad \begin{array}{ccc}(\Gamma;\Delta) & (\Gamma';\Delta',[\![A]\!]) & (\Gamma'';\Delta'',[\![B]\!])\\ \vdots & \vdots & \vdots\\ \overline{\overline{A \wedge B}} & \overline{\overline{C}} & \overline{\overline{C}}\\ \hline\hline & C & \end{array}\ \wedge E^d$$

$$\dfrac{\begin{array}{c}(\Gamma;\Delta)\\ \vdots\\ \overline{A}\end{array}}{A \vee B}\ \vee I_1 \qquad \dfrac{\begin{array}{c}(\Gamma;\Delta)\\ \vdots\\ \overline{B}\end{array}}{A \vee B}\ \vee I_2 \qquad \dfrac{\begin{array}{c}(\Gamma;\Delta)\\ \vdots\\ \overline{A \vee B}\end{array} \quad \begin{array}{c}([A],\Gamma';\Delta')\\ \vdots\\ \overline{C}\end{array} \quad \begin{array}{c}([B],\Gamma'';\Delta'')\\ \vdots\\ \overline{C}\end{array}}{C}\ \vee E$$

$$\begin{array}{cc}(\Gamma;\Delta) & (\Gamma';\Delta')\\ \vdots & \vdots\\ \overline{\overline{A}} & \overline{\overline{B}}\\ \hline\hline \multicolumn{2}{c}{A \vee B}\end{array}\ \vee I^d \qquad \begin{array}{c}(\Gamma;\Delta)\\ \vdots\\ \overline{\overline{A \vee B}}\\ \hline\hline A\end{array}\ \vee E^d_1 \qquad \begin{array}{c}(\Gamma;\Delta)\\ \vdots\\ \overline{\overline{A \vee B}}\\ \hline\hline B\end{array}\ \vee E^d_2$$

$$\dfrac{\begin{array}{c}([A],\Gamma;\Delta)\\ \vdots\\ \overline{B}\end{array}}{A \to B}\ {\to} I \qquad \dfrac{\begin{array}{c}(\Gamma;\Delta)\\ \vdots\\ \overline{A}\end{array} \quad \begin{array}{c}(\Gamma';\Delta')\\ \vdots\\ \overline{A \to B}\end{array}}{B}\ {\to} E$$

$$\begin{array}{cc}(\Gamma;\Delta) & (\Gamma';\Delta')\\ \vdots & \vdots\\ \overline{A} & \overline{\overline{B}}\\ \hline\hline \multicolumn{2}{c}{A \to B}\end{array}\ {\to} I^d \qquad \begin{array}{c}(\Gamma;\Delta)\\ \vdots\\ \overline{\overline{A \to B}}\\ \hline\hline A\end{array}\ {\to} E^d_1 \qquad \begin{array}{c}(\Gamma;\Delta)\\ \vdots\\ \overline{\overline{A \to B}}\\ \hline\hline B\end{array}\ {\to} E^d_2$$

$$\dfrac{\begin{array}{c}(\Gamma;\Delta)\\ \vdots\\ \overline{A}\end{array} \quad \begin{array}{c}(\Gamma';\Delta')\\ \vdots\\ \overline{\overline{B}}\end{array}}{A \prec B}\ {\prec} I \qquad \dfrac{\begin{array}{c}(\Gamma;\Delta)\\ \vdots\\ \overline{A \prec B}\end{array}}{A}\ {\prec} E_1 \qquad \begin{array}{c}(\Gamma;\Delta)\\ \vdots\\ \overline{A \prec B}\\ \hline\hline B\end{array}\ {\prec} E_2$$

$$\dfrac{\dfrac{\begin{matrix}(\Gamma;\Delta,\llbracket A\rrbracket)\\ \vdots\end{matrix}}{B}}{B\prec A}\prec I^d \qquad \dfrac{\dfrac{\begin{matrix}(\Gamma;\Delta)\\ \vdots\end{matrix}}{B\prec A}\quad \dfrac{\begin{matrix}(\Gamma';\Delta')\\ \vdots\end{matrix}}{A}}{B}\prec E^d$$

What I will present here additionally, is a sequent calculus, which I will call `SC2Int`. `SC2Int` corresponds to `N2Int` in that we have a proof in `N2Int` of $A$ from the pair $(\Gamma;\Delta)$, iff the sequent $(\Gamma;\Delta)\vdash^+ A$ is derivable in `SC2Int` and we have a dual proof of $A$ from the pair $(\Gamma;\Delta)$, iff the sequent $(\Gamma;\Delta)\vdash^- A$ is derivable in `SC2Int`. While Wansing (2017) proves a normal form theorem for `N2Int`, for `SC2Int` also a cut-elimination theorem can be proven (Ayhan, 2020). Since this means that our system enjoys the subformula property, this ensures the conservativeness of our system.[72] Sequents are of the form $(\Gamma;\Delta)\vdash^* C$ (with $\Gamma$ and $\Delta$ being finite, possibly empty multisets and $*\in\{+,-\}$). Within the right introduction rules we need to distinguish whether the derivability relation expresses verification or falsification by using the superscripts + and -. Within the left rules this is not necessary, but what is needed here instead is distinguishing an introduction of the principal formula into the *assumptions* (indexed by superscript $a$) from an introduction into the *counterassumptions* (indexed by superscript $c$). Thus, the set of *proof rules* in `SC2Int` consists of the rules marked with + or with $a$, while the set of *dual proof rules* consists of the rules marked with - or with $c$. When a rule contains multiple occurrences of $*$, application of this rule requires that all such occurrences are instantiated in the same way, i.e., either as + or as -.

`SC2Int`

For $*\in\{+,-\}$ :

$$\dfrac{}{(\Gamma,p;\Delta)\vdash^+ p}\,Rf^+ \qquad \dfrac{}{(\Gamma;\Delta,p)\vdash^- p}\,Rf^-$$

$$\dfrac{}{(\Gamma;\Delta)\vdash^- \bot}\,\bot R^- \qquad \dfrac{}{(\Gamma,\bot;\Delta)\vdash^* C}\,\bot L^a \qquad \dfrac{}{(\Gamma;\Delta)\vdash^+ \top}\,\top R^+ \qquad \dfrac{}{(\Gamma;\Delta,\top)\vdash^* C}\,\top L^c$$

$$\dfrac{(\Gamma;\Delta)\vdash^+ A \qquad (\Gamma;\Delta)\vdash^+ B}{(\Gamma;\Delta)\vdash^+ A\wedge B}\,\wedge R^+ \qquad \dfrac{(\Gamma,A,B;\Delta)\vdash^* C}{(\Gamma,A\wedge B;\Delta)\vdash^* C}\,\wedge L^a$$

$$\dfrac{(\Gamma;\Delta)\vdash^- A}{(\Gamma;\Delta)\vdash^- A\wedge B}\,\wedge R^-_1 \qquad \dfrac{(\Gamma;\Delta)\vdash^- B}{(\Gamma;\Delta)\vdash^- A\wedge B}\,\wedge R^-_2 \qquad \dfrac{(\Gamma;\Delta,A)\vdash^* C \quad (\Gamma;\Delta,B)\vdash^* C}{(\Gamma;\Delta,A\wedge B)\vdash^* C}\,\wedge L^c$$

[72] The exact relation between conservativeness and cut-elimination is debatable and, more specifically, depends on the system that is used (Hacking, 1979; Kremer, 1988) but given that we can also prove admissibility of the other structural rules, this should be a safe assumption for our system.

$$\frac{(\Gamma;\Delta)\vdash^{+} A}{(\Gamma;\Delta)\vdash^{+} A\vee B}\,{\vee}R^{+}_{1}\qquad \frac{(\Gamma;\Delta)\vdash^{+} B}{(\Gamma;\Delta)\vdash^{+} A\vee B}\,{\vee}R^{+}_{2}\qquad \frac{(\Gamma,A;\Delta)\vdash^{*} C\quad (\Gamma,B;\Delta)\vdash^{*} C}{(\Gamma,A\vee B;\Delta)\vdash^{*} C}\,{\vee}L^{a}$$

$$\frac{(\Gamma;\Delta)\vdash^{-} A\qquad (\Gamma;\Delta)\vdash^{-} B}{(\Gamma;\Delta)\vdash^{-} A\vee B}\,{\vee}R^{-}\qquad \frac{(\Gamma;\Delta,A,B)\vdash^{*} C}{(\Gamma;\Delta,A\vee B)\vdash^{*} C}\,{\vee}L^{c}$$

$$\frac{(\Gamma,A;\Delta)\vdash^{+} B}{(\Gamma;\Delta)\vdash^{+} A\to B}\,{\to}R^{+}\qquad \frac{(\Gamma,A\to B;\Delta)\vdash^{+} A\qquad (\Gamma,B;\Delta)\vdash^{*} C}{(\Gamma,A\to B;\Delta)\vdash^{*} C}\,{\to}L^{a}$$

$$\frac{(\Gamma;\Delta)\vdash^{+} A\qquad (\Gamma;\Delta)\vdash^{-} B}{(\Gamma;\Delta)\vdash^{-} A\to B}\,{\to}R^{-}\qquad \frac{(\Gamma,A;\Delta,B)\vdash^{*} C}{(\Gamma;\Delta,A\to B)\vdash^{*} C}\,{\to}L^{c}$$

$$\frac{(\Gamma;\Delta)\vdash^{+} A\qquad (\Gamma;\Delta)\vdash^{-} B}{(\Gamma;\Delta)\vdash^{+} A\prec B}\,{\prec}R^{+}\qquad \frac{(\Gamma,A;\Delta,B)\vdash^{*} C}{(\Gamma,A\prec B;\Delta)\vdash^{*} C}\,{\prec}L^{a}$$

$$\frac{(\Gamma;\Delta,B)\vdash^{-} A}{(\Gamma;\Delta)\vdash^{-} A\prec B}\,{\prec}R^{-}\qquad \frac{(\Gamma;\Delta,A\prec B)\vdash^{-} B\qquad (\Gamma;\Delta,A)\vdash^{*} C}{(\Gamma;\Delta,A\prec B)\vdash^{*} C}\,{\prec}L^{c}$$

The following structural rules of weakening, contraction, and cut can be shown to be admissible in `SC2Int`:

$$\frac{(\Gamma;\Delta)\vdash^{*} C}{(\Gamma,A;\Delta)\vdash^{*} C}\,W^{a}\qquad \frac{(\Gamma;\Delta)\vdash^{*} C}{(\Gamma;\Delta,A)\vdash^{*} C}\,W^{c}$$

$$\frac{(\Gamma,A,A;\Delta)\vdash^{*} C}{(\Gamma,A;\Delta)\vdash^{*} C}\,C^{a}\qquad \frac{(\Gamma;\Delta,A,A)\vdash^{*} C}{(\Gamma;\Delta,A)\vdash^{*} C}\,C^{c}$$

$$\frac{(\Gamma;\Delta)\vdash^{+} D\qquad (\Gamma',D;\Delta')\vdash^{*} C}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^{*} C}\,Cut^{a}\qquad \frac{(\Gamma;\Delta)\vdash^{-} D\qquad (\Gamma';\Delta',D)\vdash^{*} C}{(\Gamma,\Gamma';\Delta,\Delta')\vdash^{*} C}\,Cut^{c}$$

## 5.3 Uniqueness

### 5.3.1 The notion of uniqueness

The issue of uniqueness has not received much attention in the literature. It was introduced more or less *en passant* in Belnap's (1962) famous response to the `tonk`-attack by Prior (1960) against an inferentialist view on the meaning of connectives.[73] Prior's intention in using `tonk` is to show that it leads the idea of PTS[74] *ad absurdum.* He argues that if the rules of inference governing the use of a connective would indeed

[73]Belnap refers to a lecture by Hiż as being the actual origin of this idea.

[74]The term "proof-theoretic semantics" emerged much later of course but I use it whenever the idea fits to whatever terminology may be used in other places.

be all there is to the meaning of it, then nothing would prevent the inclusion of a seemingly non-sensical connective, which ultimately trivializes our system, since it allows anything to be derived from everything. Belnap's proposal to solve this so-called *existence* issue of connectives was to demand extensions of a given system to be "conservative". In addition to that, he claims, one could wonder about the *uniqueness* issue of connectives. Once we have settled that it is allowed to extend our system with a certain connective, we can ask whether the rules of inference governing the connective characterize this connective *uniquely*.

Uniqueness as a requirement for a connective means that characterizing its inference rules amounts to exactly specifying its role in inference. There can be *at most* one connective playing this role; duplication of that connective with the same characterizing rules does not change its behavior, neither in the premises nor in the conclusion. However, since Belnap's first requirement of conservativeness of the system was seen (by the responding literature and also by himself) to be far more important, the uniqueness requirement was more or less forgotten until it resurfaced in (Došen & Schroeder-Heister, 1985, 1988), which cover quite technical treatments of the issue as well as of connections to other proof-theoretic features. After that, the topic is absent from the debate for a long time again. A recent resuming of it can be found in (Naibo & Petrolo, 2015), which targets the question whether the uniqueness condition for connectives is the same as Hacking's "deducibility of identicals"-criterion[75]. Humberstone (2011, 2019, 2020b) is one of the few scholars who treats the topic quite extensively, dedicating one chapter of his monumental work on connectives to the question of uniqueness. His observations on the connections between (failure of) uniqueness of connectives, proof systems, and features of the consequence relation are of particular importance for the present purpose.

On the usual account of uniqueness two connectives # and #', which are defined by exactly the same set of inference rules and $\vdash$ being the consequence relation generated by the combined set of the rules, play exactly the same inferential role iff it can be shown for all $A$ and $B$ that $A \# B \dashv\vdash A \#' B$. Let us assume, for the moment, a common intuitionistic calculus and the example of conjunction. It can easily be shown that $\wedge$ is uniquely characterized by its usual natural deduction (resp. sequent calculus) rules (i.e., in our systems above: by its *proof rules*) governing it, since we can derive $A \wedge B$ from $A \wedge' B$ and vice versa, taking $\wedge'$ to be a connective governed by exactly the same rules as $\wedge$:

$$\dfrac{\dfrac{A \wedge' B}{A}\,{\scriptstyle \wedge' E} \qquad \dfrac{A \wedge' B}{B}\,{\scriptstyle \wedge' E}}{A \wedge B}\,{\scriptstyle \wedge I} \qquad\qquad \dfrac{\dfrac{A \wedge B}{A}\,{\scriptstyle \wedge E} \qquad \dfrac{A \wedge B}{B}\,{\scriptstyle \wedge E}}{A \wedge' B}\,{\scriptstyle \wedge I'}$$

[75] The condition that the structural rule of reflexivity for arbitrary formulas is provably admissible for every connective, i.e., each derivation using an application of it with a complex formula can be replaced by a derivation using applications of the rule with only atomic formulas (Hacking, 1979).

Thus, the interderivability requirement makes clear why, as I mentioned above, it is important to consider the underlying consequence relation when asking about the uniqueness of connectives.

Belnap's (1962, p. 133) original counterexample for satisfying the uniqueness condition is the connective `plonk`. We define `plonk` by the following rule: $A$ `plonk` $B$ can be derived from $B$. Since an extension with `plonk` (in the system Belnap is presupposing) is conservative, it can be stated that there is such a connective. However, it is not unique, since there can be another connective, which he calls `plink` defined by exactly the same rule, i.e., $A$ `plink` $B$ can be derived from $B$, which can otherwise play a different inferential role. The uniqueness requirement, as Belnap puts it, demands that another connective specified by exactly the same rules ought to play exactly the same role in inference, both as premise and as conclusion. In his system with reflexivity, weakening, permutation, contraction, and transitivity as structural rules, this amounts to showing that $A$ `plonk` $B$ and $A$ `plink` $B$ are interderivable. This, however, is not possible given that there is only this one rule governing the connectives and hence, he concludes, `plonk` is not uniquely determined by its definition.

### 5.3.2 Problematic settings

There are several examples of connectives which are not uniquely characterized. This can be shown not only for 'ad hoc' connectives, in the sense that they are only thought of for this purpose, but also for connectives existing in calculi actually used, as, e.g., $\neg$ in `FDE` or $\Box$ in system `K`.[76] Failure of uniqueness can - among other reasons - occur due to the specific formulation of the proof system, non-congruentiality of the logic, or impurity of the rules. Humberstone (2011, p. 595f.) emphasizes that what does or does not uniquely characterize a given connective is the *set of rules* governing the connective, while sets of rules can be seen as a set of *conditions on consequence relations*.

The usual system Humberstone refers to when showing the non-uniqueness (e.g., of the examples mentioned in the last paragraph) is what he calls "sequent-to-sequent rules in the framework SET-FMLA", i.e., sequent rules with a set of formulas on the left side of the sequent operator and exactly one formula on the right. He also gives examples, however, where we have uniqueness in one particular formulation of the rules but not in another. Negation in Minimal Logic, for example, cannot be uniquely characterized by any collection of SET-FMLA-rules, but can be by others, which allow *at most* one formula on the right side of the sequent operator (Humberstone, 2020a, p. 186). Another example would be that disjunction

[76] Or for that matter $\Box$ in every normal modal logic except for the Post-complete ones (Humberstone, 2011, p. 601-605). Examples of failure of uniqueness are given in (Humberstone, 2011, 2019, 2020a; Naibo & Petrolo, 2015).

is not uniquely characterized by its classical (or intuitionistic) rules when those are formulated in a zero-premise SET-FMLA system (Humberstone, 2011, p. 600).

Another important issue concerning uniqueness is the question of congruentiality, which can be a property of connectives, consequence relations, or logics (depending on the specific understanding of those concepts). A logic is congruential, if for all formulas $A$, $B$, $C$, whenever $A$ and $B$ are equivalent insofar as they are interderivable according to a defined consequence relation of the logic, equivalence also holds when we replace $A$ and $B$ in a more complex formula $C$ (Wójcicki, 1979).[77] This is closely connected to the notion of synonymy between formulas, since synonymy means that they are not only equivalent but also that replacing one by the other in any complex formula results in equivalent formulas. In view of (non-)congruentiality Humberstone (2011, p. 579f.) refines what I described as 'the usual account' (which he calls *uniqueness to within equivalence*) in that he claims that # is uniquely characterized by its set of rules iff every compound formed by that connective is *synonymous* to every compound (with the same components) formed by #' governed by exactly the same rules as #, which he calls *uniqueness to within synonymy.* This distinction coincides in the congruential case, but when the consequence relation is non-congruential, it can make a difference whether we demand the stronger or the weaker notion (2020a, p. 183, 187).

Another terminological refinement is needed when we have systems with connectives governed by impure rules, i.e., rules which govern more than one connective. In this case, Humberstone (2011, p. 580f.) argues, we need to speak of the connective in question being uniquely characterized *in terms of* whichever connective also appears in its rules. An example would be another non-congruential logic, namely Nelson's constructive logic with strong negation, `N4`. The rules for `N4` are impure because the rules for conjunction and implication also display the strong negation connective.[78]

### 5.3.3 Problems in a bilateralist system

The problem that occurs when asking about uniqueness in a bilateralist setting is connected to the points addressed in the last section. What causes trouble in the bilateralist proof systems laid out above - if we assume the common characterization of uniqueness (to within equivalence or synonymy) - is that we have two sets of rules for each connective and two consequence relations. It would make sense, then, to think of the proof rules as generating the consequence relation for provability and the dual proof rules as generating the dual consequence relation for dual provability.

[77] Wójcicki actually uses the term "self-extensional" instead of "congruential". The latter is used by (Humberstone, 2011, p. 175) for the case of connectives and consequence relations.

[78] At least this is the case for the traditional (unilateral) calculi given for `N4` (e.g., Prawitz, 1965, p. 97). In Section 5.3.4 I elaborate on this and compare such a proof system with a bilateral sequent calculus given in (Kamide & Wansing, 2012).

The specific consequence relation is of course important, since we usually test for uniqueness via interderivability, and in `2Int` it can be shown for both relations *individually* that our connectives are uniquely characterized by only a part of the whole set of rules. Consider the case of conjunction, for example: We can show that $\wedge$ is uniquely characterized by its proof rules, since we can show (see derivations in Section 5.3.1) that both $(A \wedge B; \emptyset) \vdash^+ A \wedge' B$ and $(A \wedge' B; \emptyset) \vdash^+ A \wedge B$ are derivable. Likewise, taking $\wedge''$ to be a connective governed by exactly the same $I^d$- and $E^d$-rules from `N2Int` as $\wedge$ (resp. $R^-$ and $L^c$-rules from `SC2Int`), we can show that it is also uniquely characterized by its dual proof rules, since $(\emptyset; A \wedge B) \vdash^- A \wedge'' B$ and $(\emptyset; A \wedge'' B) \vdash^- A \wedge B$ are derivable. To show it for `SC2Int`:

$$\dfrac{\dfrac{\dfrac{}{(\emptyset; A) \vdash^- A}\,{\scriptstyle Rf^-}}{(\emptyset; A) \vdash^- A \wedge'' B}\,{\scriptstyle \wedge'' R_1^-} \quad \dfrac{\dfrac{}{(\emptyset; B) \vdash^- B}\,{\scriptstyle Rf^-}}{(\emptyset; B) \vdash^- A \wedge'' B}\,{\scriptstyle \wedge'' R_2^-}}{(\emptyset; A \wedge B) \vdash^- A \wedge'' B}\,{\scriptstyle \wedge L^c} \qquad \dfrac{\dfrac{\dfrac{}{(\emptyset; A) \vdash^- A}\,{\scriptstyle Rf^-}}{(\emptyset; A) \vdash^- A \wedge B}\,{\scriptstyle \wedge R_1^-} \quad \dfrac{\dfrac{}{(\emptyset; B) \vdash^- B}\,{\scriptstyle Rf^-}}{(\emptyset; B) \vdash^- A \wedge B}\,{\scriptstyle \wedge R_2^-}}{(\emptyset; A \wedge'' B) \vdash^- A \wedge B}\,{\scriptstyle \wedge'' L^c}$$

However, there is no possibility to determine by this characterization that there is *only one* connective $\wedge$ because it is not possible to derive the following sequents:

$$(A \wedge B; \emptyset) \vdash^+ A \wedge'' B \qquad (\emptyset; A \wedge B) \vdash^- A \wedge' B$$
$$(A \wedge'' B; \emptyset) \vdash^+ A \wedge B \qquad (\emptyset; A \wedge' B) \vdash^- A \wedge B$$

The difference to `plonk` and `plink` is that in this case the one rule governing those connectives was ‘not enough’ to uniquely characterize a role in inference, while here a partial duplication of the rules (with proof rules only or dual proof rules only) is already enough for a unique characterization. So, in a way, we could say, the bilateral sets of rules overdetermine our connectives. However, since on the one hand both the proof rules as well as the dual proof rules uniquely characterize a connective, but on the other hand, there is no interderivability ‘across’ the consequence relations possible, how can we know that there is one conjunction with a unique meaning? Wouldn’t that mean that we would be forced to say that there are actually two conjunctions, $\wedge^+$ and $\wedge^-$, one for the context of provability and one for dual provability? Thus, we could not confidently claim that our conjunction is uniquely characterized and has only one meaning in a system like `N2Int` or `SC2Int`, which would certainly have to be considered problematic.

However, let us take a look at our rules again, especially at the ones for implication and co-implication: What we can see here is that the different consequence relations are intertwined in characterizing these connectives. In `N2Int` this is observable by a mixture of single and double lines in the dual proof rules of implication, $\rightarrow I^d$ and $\rightarrow E_1^d$, and in the proof rules of co-implication, $\prec I$ and $\prec E_2$. In `SC2Int` this is indicated in the dual proof rules of implication, $\rightarrow R^-$ and $\rightarrow L^c$, as well as in the proof rules of co-implication, $\prec R^+$ and $\prec L^a$, by a mixture of $\vdash^+$ and $\vdash^-$ in the right introduction rules and for the left introduction rules by the fact that active formulas are part of the assumptions *as well as* of the counterassumptions. Thus, the rules for implication as well as for co-implication need *both* consequence relations in one and the same rule application. This indicates that it would not be correct to think of the proof rules as generating the consequence relation and the dual proof rules as generating the dual consequence relation. Instead, both relations are generated by rules of both sets.[79] And this fact would support the point that we are not allowed to use different duplications of a connective when trying to show its uniqueness. Thus, when duplicating a connective, we need to use *the same duplication* for both proof rules *and* dual proof rules. By doing so, it is guaranteed that we are not talking about different connectives in different proof contexts.

So, my proposal is to modify our characterization of uniqueness in a way that it also fits the context of bilateralism: In a bilateralist setting, instead of taking interderivability as a sufficient criterion for uniqueness, we also have to consider dual interderivability.

**Definition of uniqueness for bilaterally defined connectives:**

In a bilateralist setting with consequence relations for verification as well as falsification, two n-place connectives # and #', which are defined by exactly the same set of inference rules, play exactly the same inferential role, i.e., are unique, iff for all $A_1, \ldots, A_n$ the formulas $\#(A_1, \ldots, A_n)$ and $\#'(A_1, \ldots, A_n)$ are interderivable as well as dually interderivable. To express this formally for the case of `2Int`:

(i) $(A \# B; \emptyset) \vdash^+ A \#' B$ and $(A \#' B; \emptyset) \vdash^+ A \# B$

(ii) $(\emptyset; A \# B) \vdash^- A \#' B$ and $(\emptyset; A \#' B) \vdash^- A \# B$.

With this definition of uniqueness we can state that all connectives of `2Int` are uniquely characterized by their rules with respect to `N2Int` and `SC2Int`.[80]

A last question to consider, having Humberstone's distinction in mind, would be if this holds for *uniqueness to within equivalence* only or also for *uniqueness to within synonymy*. The question needs to be asked since `2Int` is in fact also a non-congruential logic. The non-congruentiality in `2Int` stems from the fact that not all formulas that are equivalent with respect to $\vdash^+$ are also equivalent with respect to $\vdash^-$. While for example $-(A \to B)$ and $A \wedge -B$ are interderivable with respect to $\vdash^+$, this does not hold for $\vdash^-$. Fortunately, the answer is that with the definition above we indeed get uniqueness to within synonymy because the following holds in `2Int`: If we have equivalence, i.e., interderivability, of formulas both with respect to $\vdash^+$ as well as to $\vdash^-$, then it is guaranteed that these formulas are also replaceable in any more complex formula, i.e., then it is guaranteed that they are synonymous (for

[79] `SC2Int` shows this feature of 'mixedness' even nicer than `N2Int`, since in the former we have a $\vdash^*$ in *all* left rules, meaning that the rule holds for both verification and falsification.

[80] This also holds for the constants $\top$ and $\bot$, since in the case of n=0, $\#(A_1, \ldots, A_n) = \#$.

the proof see Wansing, 2016a). So the upshot of this definition is that we do not only get uniqueness to within equivalence but even uniqueness to within synonymy, without the need to consider compound formulas.

### 5.3.4 Implications for the choosing the 'right' proof system

I think that the problems for uniqueness that have been shown by Humberstone and in this paper for the case of bilateralism as well as the solution I outlined can be taken to argue in a more general way that there are better and worse proof systems. In particular, I want to compare the unilateral ND system of Nelson's constructive logic with strong negation, N4,[81] that Prawitz (e.g., 1965, p. 97) proposed, to a proof system for N4 in which bilateralism is incorporated by having two derivability relations and thereby I show that the latter exhibit more desirable features from a PTS point of view.

The bilateral system is a sequent calculus system, called Sn4, and was proposed in (Kamide & Wansing, 2012). The rules in Prawitz' system are impure as you can see, e.g., with the rules for implication because they also contain the connective of strong negation:

$$\frac{A \qquad \sim B}{\sim (A \to B)} \sim\!\to\! I \qquad\qquad \frac{\sim (A \to B)}{A} \sim\!\to\! E_1 \qquad\qquad \frac{\sim (A \to B)}{\sim B} \sim\!\to\! E_2$$

The corresponding rules in Sn4 are the following pure rules:

$$\frac{\Gamma : \Delta \vdash B : \emptyset \quad \Gamma' : \Delta' \vdash \emptyset : A}{\Gamma, \Gamma' : \Delta, \Delta' \vdash A \to B : \emptyset} \to\! R- \qquad\qquad \frac{\Gamma, B : \Delta, A \vdash C}{\Gamma, A \to B : \Delta \vdash C} \to\! L-$$

Note that in the notation of Sn4 the positions are the other way around as opposed to the notation of SC2Int, i.e., the left side marks the 'negative' side and the right one the 'positive'. Also, as can be seen, instead of having the turnstile marked with $^+$ and $^-$, here we have two positions not only on the left side of the turnstile but also on the right side. Thus, a sequent that would be written $(\Gamma; \Delta) \vdash^- A$ in SC2Int would become $(\Delta : \Gamma) \vdash A : \emptyset$ in Sn4. At this example one can see how purity of rules can be achieved by internalizing different consequence relations within the proof system.[82]

What does considering uniqueness imply now for a choice between these two proof systems? As laid out above, if we would ask for the uniqueness of $\to$ in N4

[81] I choose this example because N4 and 2Int are related in that strong negation $\sim$ in Nelson's logic can be read as a direct toggle between proofs and dual proofs, if it were added to 2Int, i.e., we would have $\vdash^+ A$ iff $\vdash^- \sim A$ and $\vdash^- A$ iff $\vdash^+ \sim A$.

[82] Likewise, Drobyshevich (2019) introduces the notion of a signed consequence relation between a set of signed formulas and a single signed formula as a bilateral variant of the notion of a Tarskian consequence relation and gives a bilateral natural deduction system for N4, which also contains pure rules only.

with impure rules, the question would always have to be "Is $\to$ uniquely characterized by its rules in terms of $\sim$?". What is more, in N4 strong negation leads to the system's non-congruentiality, since for two formulas to be equivalently replaceable *in all contexts* it is not sufficient for the formulas to be provably equivalent, but additionally, we also need equivalence between the *negated* formulas.[83] For uniqueness this would mean that firstly, due to the congruentiality issue we would have to demand uniqueness to within synonymy. Thus, this would tie uniqueness in this system to strong negation, since we would have to demand not only the interderivability of all formulas containing the connective in question with the formula containing the 'copy-cat' connective, but also the same interderivability with the strongly negated formulas. However, given that the connectives can only be uniquely characterized in terms of strong negation, which by itself *cannot* be uniquely characterized by its rules,[84] this does not seem like a desirable system or a good solution to recover uniqueness. It seems circular, or at least not well-formed, to have to characterize the uniqueness of connectives in terms of one specific connective, which at the same time is not uniquely defined by its rules and plays the establishing role for the definition of uniqueness for the system in general. In a presentation of the rules in the way of Sn4, though, we do not encounter this problem since the rules are there presented in pure form.

## 5.4 Conclusion

It has been made clear in other works that there are several features in logical systems which may cause problems for the claim that the connectives are uniquely characterized by the rules of that system. In this paper I examined the specific problem that occurs in a bilateralist setting in which we have two consequence relations, one for provability and one for dual provability. The refinements that are needed in such a setting differ from the ones that have been detected so far. In our specified case we also need to require that the interderivability of the formulas containing the connective is satisfied *for both consequence relations*. In other bilateral systems the specific formulation of what we require for uniqueness may differ, but in one way or another we will always need a requirement which holds not only for the context of verification (or assertion, or provability), but also for the context of falsification (or denial, or dual provability).

[83] A counterexample to congruentiality of N4 is that equivalence holds between $\sim (A \to B)$ and $(A \wedge \sim B)$ but not between $\sim\sim (A \to B)$ and $\sim (A \wedge \sim B)$ (Wansing, 2016a, p. 445).

[84] Given that the only rules governing $\sim$ are only thos two: $A \vdash \sim\sim A$ and $\sim\sim A \vdash A$.

# 6 Meaning and identity of proofs in a bilateralist setting

## A two-sorted typed $\lambda$-calculus for proofs and refutations

## 6.1 Introduction

In this paper I will develop a type theory for a bi-intuitionistic logic and discuss its implications for the notions of sense and denotation of derivations in a bilateralist setting. Thus, I will use the Curry-Howard correspondence, which has been well-established between the simply typed $\lambda$-calculus and natural deduction systems for intuitionistic logic, and apply it to a bilateralist proof system displaying two derivability relations, one for proving and one for refuting. The basis will be the natural deduction system of Wansing's bi-intuitionistic logic `2Int` (2016a; 2017), which I will turn into a term-annotated form. Therefore, we need a type theory that extends to a two-sorted typed $\lambda$-calculus similar to what Wansing (2016b) presents for the bi-connexive logic `2C`. He uses a type theory à la Church, though, while I will introduce a Curry-style type theory. I will argue that this gives us interesting insights into questions about sense and denotation as well as synonymy and identity of proofs from a bilateralist point of view.

## 6.2 A type theory for `2Int`: $\lambda^{2Int}$

### 6.2.1 Term-annotated `N2Int` and some results for the system

Let `Prop` be a countably infinite set of atomic formulas. Elements from `Prop` will be denoted $\rho$, $\sigma$, $\tau$, $\rho_1$, $\rho_2$ ... etc. Formulas generated from `Prop` will be denoted $A, B, C, A_1, A_2$, ... etc. We use $\Gamma$, $\Delta$,... for multisets of formulas. The concatenation $\Gamma$, $A$ stands for $\Gamma \cup \{A\}$.

The language $\mathscr{L}_{2Int}$ of `2Int`, as given by Wansing, is defined in Backus-Naur form as follows:

$A ::= \rho \mid \bot \mid \top \mid (A \wedge A) \mid (A \vee A) \mid (A \rightarrow A) \mid (A \prec A)$.

As can be seen, we have a non-standard connective in this language, namely the operator of co-implication $\prec$,[85] which acts as a dual to implication, just like conjunction and disjunction can be seen as dual connectives of each other. With that we are in the realms of so-called *bi-intuitionistic logic*, which is a conservative extension of intuitionistic logic by co-implication. Note that there is also a use of "bi-intuitionistic logic" in the literature to refer to a specific system, namely `BiInt`,

[85] Sometimes also called "pseudo-difference", e.g., in Rauszer (1974), or "subtraction", e.g., in Restall (1997), and used with different symbols.

also called "Heyting-Brouwer logic". Co-implication is there to be understood to internalize the preservation of non-truth from the conclusion to the premises in a valid inference. The system `2Int`, which is treated here, uses the same language as `BiInt`, but the meaning of co-implication differs in that it internalizes the preservation of falsity from the premises to the conclusion in a dually valid inference (Wansing, 2016a, 2016c, 2017, p. 30ff.).

In (Wansing, 2017) a natural deduction system for `2Int` is given and a normal form theorem is proven for it. Besides the usual introduction and elimination rules for intuitionistic logic (henceforth: the *proof rules*, indicated by using single lines), the system `N2Int` also comprises rules that allow us to introduce and eliminate our connectives into and from dual proofs.[86] These so-called *dual proof rules* (indicated by using double lines) are obtained by a dualization of the proof rules (for the description and the rules of the calculus, see Wansing, 2017, p. 32-34) and having these two independent sets of rules is exactly what reflects the bilateralism of the proof system.[87] Also, a distinction is drawn in the premises between *assumptions* (taken to be true) and *counterassumptions* (taken to be false). This is indicated by an ordered pair $(\Gamma; \Delta)$ (with $\Gamma$ and $\Delta$ being finite, possibly empty multisets) of assumptions ($\Gamma$) and counterassumptions ($\Delta$), together called the *basis* of a derivation. Single square brackets denote a possible discharge of assumptions, while double square brackets denote a possible discharge of counterassumptions. If there is a derivation of $A$ from a (possibly empty) basis $(\Gamma; \Delta)$ whose last inference step is constituted by a proof rule, this will be indicated by $(\Gamma; \Delta) \vdash^{+} A$. If there is a derivation of $A$ from a (possibly empty) basis $(\Gamma; \Delta)$ whose last inference step is constituted by a dual proof rule, this will be indicated by $(\Gamma; \Delta) \vdash^{-} A$.

Whenever the superscript $^*$ is used with a symbol, this is to indicate that the superscript can be either $+$ or $-$ (called *polarities*). When $^*$ is used multiple times within a symbol, this is meant to always denote the same polarity. In contrast, when $^\dagger$ is used next to $^*$ in a symbol, this means that it can - but does not have to - be of another polarity (yet again multiple $^\dagger$ denote the same polarity, i.e., for example $\texttt{case}\ r^*\{x^*.t^\dagger|y^*.s^\dagger\}^\dagger$ could either stand for $\texttt{case}\ r^+\{x^+.t^+|y^+.s^+\}^+$, $\texttt{case}\ r^-\{x^-.t^-|y^-.s^-\}^-$, $\texttt{case}\ r^+\{x^+.t^-|y^+.s^-\}^-$, or $\texttt{case}\ r^-\{x^-.t^+|y^-.s^+\}^+$ but not for, e.g., $\texttt{case}\ r^+\{x^+.t^+|y^+.s^-\}^-$. Furthermore, we use '$\equiv$' to denote *syntactic identity* between terms, types, or derivations.

[86] Especially in the later sections when I will discuss more philosophical issues, I will often use "refutations" instead of "dual proofs". The latter is the terminologically stricter expression, which is appropriate when we speak about the proof system, but it expresses essentially the same concept as the former.

[87] Apart from the term annotations, our presentation of `N2Int` given below differs in two minor aspects from the presentation in (Wansing, 2017) adopted above in 5.2.2: Firstly, we include explicit introduction rules for $\bot$ and $\top$, and secondly, we use dashed lines in four of the rules indicating that the conclusion can be obtained either by a proof or by a dual proof. These versions of the rules are derivable in Wansing's original `N2Int`, though.

**Definition 6.1**
*The set of* type symbols *(or just* types*) is the set of all formulas of* $\mathscr{L}_{2Int}$. *Let* $\mathtt{Var}_{2Int}$ *be a countably infinite set of two-sorted term variables. Elements from* $\mathtt{Var}_{2Int}$ *will be denoted* $x^*$, $y^*$, $z^*$, $x_1^*$, $x_2^*$ ... *etc. The two-sorted terms generated from* $\mathtt{Var}_{2Int}$ *will be denoted* $t^*, r^*, s^*, t_1^*, t_2^*$, ... *etc. The set* $\mathtt{Term}_{2Int}$ *can be defined in Backus-Naur form as follows:*
$t ::= x^* \mid abort(t^*)^\dagger \mid \langle t^*, t^*\rangle^* \mid fst(t^*)^* \mid snd(t^*)^* \mid inl(t^*)^* \mid inr(t^*)^*$
$\mid \mathtt{case}\ t^*\{x^*.t^\dagger | x^*.t^\dagger\}^\dagger \mid (\lambda x^*.t^*)^* \mid App(t^*, t^*)^* \mid \{t^+, t^-\}^* \mid \pi_1(t^*)^\dagger \mid \pi_2(t^*)^\dagger$.

**Definition 6.2**
*A* (type assignment) statement *is of the form* $t : A$ *with term* $t$ *being the* subject *and type* $A$ *the* predicate *of the statement. It is read "term* $t$ *is of type* $A$*" or, in the 'proof-reading', "*$t$ *is a proof of formula* $A$*".*

We are thus using a type-system à la Curry, in which the terms are not typed, in the sense that the types are part of the term's structure, but are *assigned* types. Substitution is expressed by $t[s/x]$, meaning that in term $t$ every free occurrence of $x$ is substituted with $s$. The usual capture-avoiding requirements for variable substitution are to be observed.

**Definition 6.3**
*A statement* $t : A$ *is* derivable *in term-annotated* **N2Int** *from a (possibly empty) basis* $(\Gamma; \Delta)$, *i.e., there is a* derivation $(\Gamma; \Delta) \vdash^*_{N2Int} t : A$, *if* $t : A$ *can be produced as the conclusion from the premises* $(\Gamma; \Delta)$ *according to the following rules:*[88]

`Term-annotated N2Int`

$$\frac{}{t^- : \bot}\ \bot I \qquad \frac{\begin{array}{c}(\Gamma;\Delta)\\ \vdots \\ \hline t^+ : \bot\end{array}}{abort(t^+)^* : A}\ \bot E \qquad \frac{}{t^+ : \top}\ \top I \qquad \frac{\begin{array}{c}(\Gamma;\Delta)\\ \vdots \\ \hline\hline t^- : \top\end{array}}{abort(t^-)^* : A}\ \top E^d$$

$$\frac{\begin{array}{c}(\Gamma;\Delta)\\ \vdots \\ \hline s^+ : A\end{array} \quad \begin{array}{c}(\Gamma';\Delta')\\ \vdots \\ \hline t^+ : B\end{array}}{\langle s^+, t^+\rangle^+ : A \wedge B}\ \wedge I \qquad \frac{\begin{array}{c}(\Gamma;\Delta)\\ \vdots \\ \hline t^+ : A \wedge B\end{array}}{fst(t^+)^+ : A}\ \wedge E_1 \qquad \frac{\begin{array}{c}(\Gamma;\Delta)\\ \vdots \\ \hline t^+ : A \wedge B\end{array}}{snd(t^+)^+ : B}\ \wedge E_2$$

$$\frac{\begin{array}{c}(\Gamma;\Delta)\\ \vdots \\ \hline\hline t^- : A\end{array}}{inl(t^-)^- : A \wedge B}\ \wedge I_1^d \qquad \frac{\begin{array}{c}(\Gamma;\Delta)\\ \vdots \\ \hline\hline t^- : B\end{array}}{inr(t^-)^- : A \wedge B}\ \wedge I_2^d$$

[88] The subscript of the turnstile will be omitted henceforth unless there is a possibility for confusion.

$$\frac{\begin{array}{c}(\Gamma;\Delta)\\ \vdots\\ \hline\hline r^- : A\wedge B\end{array} \quad \begin{array}{c}(\Gamma';\Delta',[\![x^- : A]\!])\\ \vdots\\ \hline s^* : C\end{array} \quad \begin{array}{c}(\Gamma'';\Delta'',[\![y^- : B]\!])\\ \vdots\\ \hline t^* : C\end{array}}{\mathsf{case}\ r^-\{x^-.s^*|y^-.t^*\}^* : C}\ \wedge E^d$$

$$\frac{\begin{array}{c}(\Gamma;\Delta)\\ \vdots\\ \hline t^+ : A\end{array}}{inl(t^+)^+ : A\vee B}\ \vee I_1 \qquad \frac{\begin{array}{c}(\Gamma;\Delta)\\ \vdots\\ \hline t^+ : B\end{array}}{inr(t^+)^+ : A\vee B}\ \vee I_2$$

$$\frac{\begin{array}{c}(\Gamma;\Delta)\\ \vdots\\ \hline r^+ : A\vee B\end{array} \quad \begin{array}{c}([x^+ : A],\Gamma';\Delta')\\ \vdots\\ \hline s^* : C\end{array} \quad \begin{array}{c}([y^+ : B],\Gamma'';\Delta'')\\ \vdots\\ \hline t^* : C\end{array}}{\mathsf{case}\ r^+\{x^+.s^*|y^+.t^*\}^* : C}\ \vee E$$

$$\frac{\begin{array}{c}(\Gamma;\Delta)\\ \vdots\\ \hline\hline s^- : A\end{array} \quad \begin{array}{c}(\Gamma';\Delta')\\ \vdots\\ \hline\hline t^- : B\end{array}}{\langle s^-,t^-\rangle^- : A\vee B}\ \vee I^d \qquad \frac{\begin{array}{c}(\Gamma;\Delta)\\ \vdots\\ \hline\hline t^- : A\vee B\end{array}}{fst(t^-)^- : A}\ \vee E^d_1 \qquad \frac{\begin{array}{c}(\Gamma;\Delta)\\ \vdots\\ \hline\hline t^- : A\vee B\end{array}}{snd(t^-)^- : B}\ \vee E^d_2$$

$$\frac{\begin{array}{c}([x^+ : A],\Gamma;\Delta)\\ \vdots\\ \hline t^+ : B\end{array}}{(\lambda x^+.t^+)^+ : A\to B}\ \to I \qquad \frac{\begin{array}{c}(\Gamma;\Delta)\\ \vdots\\ \hline s^+ : A\to B\end{array} \quad \begin{array}{c}(\Gamma';\Delta')\\ \vdots\\ \hline t^+ : A\end{array}}{App(s^+,t^+)^+ : B}\ \to E$$

$$\frac{\begin{array}{c}(\Gamma;\Delta)\\ \vdots\\ \hline s^+ : A\end{array} \quad \begin{array}{c}(\Gamma';\Delta')\\ \vdots\\ \hline\hline t^- : B\end{array}}{\{s^+,t^-\}^- : A\to B}\ \to I^d \qquad \frac{\begin{array}{c}(\Gamma;\Delta)\\ \vdots\\ \hline\hline t^- : A\to B\end{array}}{\pi_1(t^-)^+ : A}\ \to E^d_1 \qquad \frac{\begin{array}{c}(\Gamma;\Delta)\\ \vdots\\ \hline\hline t^- : A\to B\end{array}}{\pi_2(t^-)^- : B}\ \to E^d_2$$

$$\frac{\begin{array}{c}(\Gamma;\Delta)\\ \vdots\\ \hline t^+ : B\end{array} \quad \begin{array}{c}(\Gamma';\Delta')\\ \vdots\\ \hline\hline s^- : A\end{array}}{\{t^+,s^-\}^+ : B\prec A}\ \prec I \qquad \frac{\begin{array}{c}(\Gamma;\Delta)\\ \vdots\\ \hline t^+ : B\prec A\end{array}}{\pi_1(t^+)^+ : B}\ \prec E_1 \qquad \frac{\begin{array}{c}(\Gamma;\Delta)\\ \vdots\\ \hline t^+ : B\prec A\end{array}}{\pi_2(t^+)^- : A}\ \prec E_2$$

$$\frac{\begin{array}{c}(\Gamma;\Delta,[\![x^- : A]\!])\\ \vdots\\ \hline\hline t^- : B\end{array}}{(\lambda x^-.t^-)^- : B\prec A}\ \prec I^d \qquad \frac{\begin{array}{c}(\Gamma;\Delta)\\ \vdots\\ \hline\hline s^- : B\prec A\end{array} \quad \begin{array}{c}(\Gamma';\Delta')\\ \vdots\\ \hline\hline t^- : A\end{array}}{App(s^-,t^-)^- : B}\ \prec E^d$$

**Definition 6.4**
*The height of a derivation is the greatest number of successive applications of rules in it, where assumptions have height 0.*

The following lemmata show how terms of a certain form are typed and we need them to prove the Subject Reduction Theorem as well as our Dualization Theorem. The terminology and presentation of the lemmata and proofs are to a great extent in the style of (Barendregt, 1992) and (Sørensen & Urzyczyn, 2006).

**Lemma 6.1** (Generation Lemma)

*1. Assumptions and zero-premise rules*

*1.1 For every x,* $(\Gamma;\Delta) \vdash^{+} x^{+} : A \Rightarrow (x^{+} : A) \in \Gamma$ *or* $A \equiv \top$ *and* $\Gamma = \Delta = \emptyset$

*1.2 For every x,* $(\Gamma;\Delta) \vdash^{-} x^{-} : A \Rightarrow (x^{-} : A) \in \Delta$ *or* $A \equiv \bot$ *and* $\Gamma = \Delta = \emptyset$

*2.* $\top/\bot$*-E-rules*

*2.1* $(\Gamma;\Delta) \vdash^{*} abort(t^{+})^{*} : A \Rightarrow (\Gamma;\Delta) \vdash^{+} t^{+} : \bot$

*2.2* $(\Gamma;\Delta) \vdash^{*} abort(t^{-})^{*} : A \Rightarrow (\Gamma;\Delta) \vdash^{-} t^{-} : \top$

*3.* $\rightarrow$*-rules*

*3.1* $(\Gamma;\Delta) \vdash^{+} (\lambda x^{+}.t^{+})^{+} : C \Rightarrow \exists A, B[(\Gamma, x^{+} : A;\Delta) \vdash^{+} t^{+} : B \,\&\, C \equiv A \rightarrow B]$

*3.2* $(\Gamma,\Gamma';\Delta,\Delta') \vdash^{+} App(s^{+},t^{+})^{+} : B \Rightarrow \exists A[(\Gamma;\Delta) \vdash^{+} s^{+} : A \rightarrow B \,\&\, (\Gamma';\Delta') \vdash^{+} t^{+} : A]$

*3.3* $(\Gamma,\Gamma';\Delta,\Delta') \vdash^{-} \{s^{+},t^{-}\}^{-} : C \Rightarrow \exists A, B[(\Gamma;\Delta) \vdash^{+} s^{+} : A \,\&\, (\Gamma';\Delta') \vdash^{-} t^{-} : B \,\&\, C \equiv A \rightarrow B]$

*3.4* $(\Gamma;\Delta) \vdash^{+} \pi_1(t^{-})^{+} : A \Rightarrow \exists B[(\Gamma;\Delta) \vdash^{-} t^{-} : A \rightarrow B]$

*3.5* $(\Gamma;\Delta) \vdash^{-} \pi_2(t^{-})^{-} : B \Rightarrow \exists A[(\Gamma;\Delta) \vdash^{-} t^{-} : A \rightarrow B]$

*4.* $\prec$*-rules*

*4.1* $(\Gamma,\Gamma';\Delta,\Delta') \vdash^{+} \{s^{+},t^{-}\}^{+} : C \Rightarrow \exists A, B[(\Gamma;\Delta) \vdash^{+} s^{+} : B \,\&\, (\Gamma';\Delta') \vdash^{-} t^{-} : A \,\&\, C \equiv B \prec A]$

*4.2* $(\Gamma;\Delta) \vdash^{+} \pi_1(t^{+})^{+} : B \Rightarrow \exists A[(\Gamma;\Delta) \vdash^{+} t^{+} : B \prec A]$

*4.3* $(\Gamma;\Delta) \vdash^{-} \pi_2(t^{+})^{-} : A \Rightarrow \exists B[(\Gamma;\Delta) \vdash^{+} t^{+} : B \prec A]$

*4.4* $(\Gamma;\Delta) \vdash^{-} (\lambda x^{-}.t^{-})^{-} : C \Rightarrow \exists A, B[(\Gamma;\Delta, x^{-} : A) \vdash^{-} t^{-} : B \,\&\, C \equiv B \prec A]$

*4.5* $(\Gamma,\Gamma';\Delta,\Delta') \vdash^{-} App(s^{-},t^{-})^{-} : B \Rightarrow \exists A[(\Gamma;\Delta) \vdash^{-} s^{-} : B \prec A \,\&\, (\Gamma';\Delta') \vdash^{-} t^{-} : A]$

*5.* $\wedge$*-rules*

*5.1* $(\Gamma,\Gamma';\Delta,\Delta') \vdash^{+} \langle s^{+},t^{+}\rangle^{+} : C \Rightarrow \exists A, B[(\Gamma;\Delta) \vdash^{+} s^{+} : A \,\&\, (\Gamma';\Delta') \vdash^{+} t^{+} : B \,\&\, C \equiv A \wedge B]$

*5.2* $(\Gamma;\Delta) \vdash^{+} fst(t^{+})^{+} : A \Rightarrow \exists B[(\Gamma;\Delta) \vdash^{+} t^{+} : A \wedge B]$

*5.3* $(\Gamma;\Delta) \vdash^{+} snd(t^{+})^{+} : B \Rightarrow \exists A[(\Gamma;\Delta) \vdash^{+} t^{+} : A \wedge B]$

*5.4* $(\Gamma;\Delta) \vdash^- inl(t^-)^- : C \Rightarrow \exists A, B[(\Gamma;\Delta) \vdash^- t^- : A \ \& \ C \equiv A \wedge B]$

*5.5* $(\Gamma;\Delta) \vdash^- inr(t^-)^- : C \Rightarrow \exists A, B[(\Gamma;\Delta) \vdash^- t^- : B \ \& \ C \equiv A \wedge B]$

*5.6* $(\Gamma, \Gamma', \Gamma''; \Delta, \Delta', \Delta'') \vdash^* \texttt{case}\ r^-\{x^-.s^*|y^-.t^*\}^* : C \Rightarrow \exists A, B[(\Gamma;\Delta) \vdash^- r^- : A \wedge B \ \& \ (\Gamma';\Delta', x^- : A) \vdash^* s^* : C \ \& \ (\Gamma'';\Delta'', y^- : B) \vdash^* t^* : C]$

*6.* $\vee$*-rules*

*6.1* $(\Gamma;\Delta) \vdash^+ inl(t^+)^+ : C \Rightarrow \exists A, B[(\Gamma;\Delta) \vdash^+ t^+ : A \ \& \ C \equiv A \vee B]$

*6.2* $(\Gamma;\Delta) \vdash^+ inr(t^+)^+ : C \Rightarrow \exists A, B[(\Gamma;\Delta) \vdash^+ t^+ : B \ \& \ C \equiv A \vee B]$

*6.3* $(\Gamma, \Gamma', \Gamma''; \Delta, \Delta', \Delta'') \vdash^* \texttt{case}\ r^+\{x^+.s^*|y^+.t^*\}^* : C \Rightarrow \exists A, B[(\Gamma;\Delta) \vdash^+ r^+ : A \vee B \ \& \ (\Gamma', x^+ : A;\Delta') \vdash^* s^* : C \ \& \ (\Gamma'', y^+ : B;\Delta'') \vdash^* t^* : C]$

*6.4* $(\Gamma, \Gamma';\Delta, \Delta') \vdash^- \langle s^-, t^- \rangle^- : C \Rightarrow \exists A, B[(\Gamma;\Delta) \vdash^- s^- : A \ \& \ (\Gamma';\Delta') \vdash^- t^- : B \ \& \ C \equiv A \vee B]$

*6.5* $(\Gamma;\Delta) \vdash^- fst(t^-)^- : A \Rightarrow \exists B[(\Gamma;\Delta) \vdash^- t^- : A \vee B]$

*6.6* $(\Gamma;\Delta) \vdash^- snd(t^-)^- : B \Rightarrow \exists A[(\Gamma;\Delta) \vdash^- t^- : A \vee B]$

*Proof.* By induction on the height $n$ of the derivation. If $n = 0$, then $(\Gamma;\Delta) \vdash^* t^* : A$ must consist of either an arbitrary single assumption, in which case $t \equiv x$ and (by definition of the basis $(\Gamma;\Delta)$) either $(x^+ : A) \in \Gamma$ or $(x^- : A) \in \Delta$, or of a proof of $\top$ or a refutation of $\bot$, which always hold.

Assume now that the clauses of the Generation Lemma hold for all derivations of height $n$. Then for all clauses 2-6 the following holds: If there is a derivation of height $n + 1$ of the form given on the left side of $\Rightarrow$, then by the rules given above for `Term-annotated N2Int` there must be a derivation of height $n$ of the form given on the right side of $\Rightarrow$.

□

**Lemma 6.2** (Substitution Lemma)

*1. If* $(\Gamma;\Delta) \vdash^* t^* : A$*, then* $(\Gamma[B/C];\Delta[D/E]) \vdash^* t^* : A[B/C; D/E]$.

*2. If* $(\Gamma, x^+ : A;\Delta) \vdash^* t^* : B$ *and* $(\Gamma';\Delta') \vdash^+ s^+ : A$*, then* $(\Gamma, \Gamma';\Delta, \Delta') \vdash^* t[s/x]^* : B$.

*3. If* $(\Gamma;\Delta, x^- : A) \vdash^* t^* : B$ *and* $(\Gamma';\Delta') \vdash^- s^- : A$*, then* $(\Gamma, \Gamma';\Delta, \Delta') \vdash^* t[s/x]^* : B$.

*Proof.* 1. By induction on the derivation of $t^* : A$ using the Generation Lemma.
2 & 3. By induction on the generation of $(\Gamma, x^+ : A;\Delta) \vdash^* t^* : B$, respectively $(\Gamma;\Delta, x^- : A) \vdash^* t^* : B$.
For the base cases (we'll leave the even more trivial cases, where we have a proof of $\top$ and a refutation of $\bot$, out), we would have
$(\Gamma, x^+ : A;\Delta) \vdash^+ x^+ : A$ and $(\Gamma';\Delta') \vdash^+ s^+ : A$, respectively
$(\Gamma;\Delta, x^- : A) \vdash^- x^- : A$ and $(\Gamma';\Delta') \vdash^- s^- : A$.

Thus, trivially (by the usual conception of derivation)
$(\Gamma, \Gamma'; \Delta, \Delta') \vdash^{+} s^{+} : A$, respectively
$(\Gamma, \Gamma'; \Delta, \Delta') \vdash^{-} s^{-} : A$.
We will consider two exemplary cases (choosing ones with a mixture of polarities to make it more interesting) to show that the proof is straightforward. For clause 2, let us consider the case that the last rule applied is $\prec E_2$. Then we have
$(\Gamma, x^{+} : A; \Delta) \vdash^{-} \pi_2(r^{+})^{-} : B$ and $(\Gamma'; \Delta') \vdash^{+} s^{+} : A$.
Thus, by Generation Lemma 4.3, for some $C$
$(\Gamma, x^{+} : A; \Delta) \vdash^{+} r^{+} : C \prec B$.
By our inductive hypothesis
$(\Gamma, \Gamma'; \Delta, \Delta') \vdash^{+} r[s/x]^{+} : C \prec B$.
Thus, by $\prec E_2$
$(\Gamma, \Gamma'; \Delta, \Delta') \vdash^{-} \pi_2(r[s/x]^{+})^{-} : B$.
For clause 3, let us consider the case that the last rule applied is $\wedge I$. Then we have
$(\Gamma; \Delta, x^{-} : A) \vdash^{+} \langle r^{+}, u^{+}\rangle^{+} : C \wedge D$ and $(\Gamma'; \Delta') \vdash^{-} s^{-} : A$.
Thus, by Generation Lemma 5.1 with $\Gamma = \Gamma'' \cup \Gamma'''$ and $\Delta = \Delta'' \cup \Delta'''$ either
$(\Gamma''; \Delta'', x^{-} : A) \vdash^{+} r^{+} : C$ and $(\Gamma'''; \Delta''') \vdash^{+} u^{+} : D$ or
$(\Gamma''; \Delta'') \vdash^{+} r^{+} : C$ and $(\Gamma'''; \Delta''', x^{-} : A) \vdash^{+} u^{+} : D$.
By our inductive hypothesis then either
$(\Gamma', \Gamma''; \Delta', \Delta'') \vdash^{+} r[s/x]^{+} : C$ or
$(\Gamma', \Gamma'''; \Delta', \Delta''') \vdash^{+} u[s/x]^{+} : D$.
Thus, by $\wedge I$ either
$(\Gamma, \Gamma'; \Delta, \Delta') \vdash^{+} \langle r[s/x]^{+}, u^{+}\rangle^{+} : C \wedge D$ or
$(\Gamma, \Gamma'; \Delta, \Delta') \vdash^{+} \langle r^{+}, u[s/x]^{+}\rangle^{+} : C \wedge D$. □

The inductive definition of a *compatible* relation will be outsourced to the appendix because for $\lambda^{2Int}$ we need a lot of clauses (see Definition 6.7). Suffice it to say that a "compatible relation 'respects' the syntactic constructions" (Sørensen & Urzyczyn, 2006, p. 12) of the terms, i.e., let $\mathscr{R}$ be a compatible relation on $\texttt{Term}_{2Int}$, then for all $t, r, s \in \texttt{Term}_{2Int}$: if $t\mathscr{R}r$, then $(\lambda x^{*}.t^{*})^{*}\mathscr{R}(\lambda x^{*}.r^{*})^{*}$, $App(t^{*}, s^{*})^{*}\mathscr{R}App(r^{*}, s^{*})^{*}$, $App(s^{*}, t^{*})^{*}\mathscr{R}App(s^{*}, r^{*})^{*}$, etc.

**Definition 6.5** (Reductions)
*1. The least compatible relation* $\rightsquigarrow_{1\beta}$ *on* $\texttt{Term}_{2Int}$ *satisfying the following clauses is called* $\beta$-reduction*:*

$$App((\lambda x^{*}.t^{*})^{*}, s^{*})^{*} \rightsquigarrow_{1\beta} t[s^{*}/x^{*}]^{*}$$
$$\pi_1(\{s^{+}, t^{-}\}^{*})^{+} \rightsquigarrow_{1\beta} s^{+} \qquad \pi_2(\{s^{+}, t^{-}\}^{*})^{-} \rightsquigarrow_{1\beta} t^{-}$$
$$fst(\langle s^{*}, t^{*}\rangle^{*})^{*} \rightsquigarrow_{1\beta} s^{*} \qquad snd(\langle s^{*}, t^{*}\rangle^{*})^{*} \rightsquigarrow_{1\beta} t^{*}$$
$$\texttt{case}\ inl(r^{*})^{*}\{x^{*}.s^{\dagger}|y^{*}.t^{\dagger}\}^{\dagger} \rightsquigarrow_{1\beta} s[r^{*}/x^{*}]^{\dagger}$$
$$\texttt{case}\ inr(r^{*})^{*}\{x^{*}.s^{\dagger}|y^{*}.t^{\dagger}\}^{\dagger} \rightsquigarrow_{1\beta} t[r^{*}/y^{*}]^{\dagger}$$

*2. For all clauses the term on the left of* $\rightsquigarrow_{1\beta}$ *is called* $\beta$-redex, *while the term on* *the right is its* contractum.

*3. A term $t$ is said to be in $\beta$-*normal form *iff $t$ does not contain a $\beta$-redex.*

*4. The relation $\rightsquigarrow$ (multi-step $\beta$-reduction) is the transitive and reflexive closure of $\rightsquigarrow_{1\beta}$.*

**Theorem 6.1** (Subject Reduction Theorem for $\lambda^{2Int}$)
*If $(\Gamma; \Delta) \vdash^{*} t^{*} : C$ and $t \rightsquigarrow_{\beta} t'$, then $(\Gamma'; \Delta') \vdash^{*} t'^{*} : C$ for $\Gamma' \subseteq \Gamma$ and $\Delta' \subseteq \Delta$.*

*Proof.* By induction on the generation of $\rightsquigarrow_{\beta}$ using the Generation and Substitution Lemmata.
We will spell out one of the reductions for each connective. For the connectives where we have two reductions it will be straightforward that the same reasoning can be applied.

Suppose $t \equiv App((\lambda x^{*}.r^{*})^{*}, s^{*})^{*}$, $t' \equiv r[s^{*}/x^{*}]^{*}$, $\Gamma = \Gamma' \cup \Gamma''$ and $\Delta = \Delta' \cup \Delta''$. If
$(\Gamma; \Delta) \vdash^{*} App((\lambda x^{*}.r^{*})^{*}, s^{*})^{*} : C$,
then by Generation Lemma 3.2 and 4.5 there must be some $A$ such that either
$(\Gamma'; \Delta') \vdash^{+} (\lambda x^{+}.r^{+})^{+} : A \to C$ and $(\Gamma''; \Delta'') \vdash^{+} s^{+} : A$, or
$(\Gamma'; \Delta') \vdash^{-} (\lambda x^{-}.r^{-})^{-} : C \prec A$ and $(\Gamma''; \Delta'') \vdash^{-} s^{-} : A$.
Thus, again by Generation Lemma 3.1 and 4.4 it follows that either
$(\Gamma', x^{+} : A; \Delta') \vdash^{+} r^{+} : C$ and $(\Gamma''; \Delta'') \vdash^{+} s^{+} : A$ or
$(\Gamma'; \Delta', x^{-} : A) \vdash^{-} r^{-} : C$ and $(\Gamma''; \Delta'') \vdash^{-} s^{-} : A$.
Therefore, by the Substitution Lemma either
$(\Gamma; \Delta) \vdash^{+} r[s^{+}/x^{+}]^{+} : C$ or
$(\Gamma; \Delta) \vdash^{-} r[s^{-}/x^{-}]^{-} : C$.

Suppose $t \equiv \pi_1(\{s^{+}, r^{-}\}^{*})^{+}$, $t' \equiv s^{+}$, $\Gamma' \subseteq \Gamma$ and $\Delta' \subseteq \Delta$. If
$(\Gamma; \Delta) \vdash^{+} \pi_1(\{s^{+}, r^{-}\}^{*})^{+} : C$,
then by Generation Lemma 3.4 and 4.2 there must be some $A$ such that either
$(\Gamma; \Delta) \vdash^{-} \{s^{+}, r^{-}\}^{-} : C \to A$ or
$(\Gamma; \Delta) \vdash^{+} \{s^{+}, r^{-}\}^{+} : C \prec A$.
Thus, again by Generation Lemma 3.3 and 4.1 it follows that in both cases
$(\Gamma'; \Delta') \vdash^{+} s^{+} : C$.

Suppose $t \equiv fst(\langle s^{*}, r^{*}\rangle^{*})^{*}$, $t' \equiv s^{*}$, $\Gamma' \subseteq \Gamma$ and $\Delta' \subseteq \Delta$. If
$(\Gamma; \Delta) \vdash^{*} fst(\langle s^{*}, r^{*}\rangle^{*})^{*} : C$,
then by Generation Lemma 5.2 and 6.5 there must be some $A$ such that either
$(\Gamma; \Delta) \vdash^{+} \langle s^{+}, r^{+}\rangle^{+} : C \wedge A$ or
$(\Gamma; \Delta) \vdash^{-} \langle s^{-}, r^{-}\rangle^{-} : C \vee A$.
Thus, again by Generation Lemma 5.1 and 6.4 it follows that in both cases
$(\Gamma'; \Delta') \vdash^{*} s^{*} : C$.

Suppose $t \equiv \texttt{case}\ inl(r^{*})^{*}\{x^{*}.s^{\dagger}|y^{*}.u^{\dagger}\}^{\dagger}$, $t' \equiv s[r^{*}/x^{*}]^{\dagger}$, $\Gamma = \Gamma' \cup \Gamma'' \cup \Gamma'''$ and $\Delta = \Delta' \cup \Delta'' \cup \Delta'''$. If

$(\Gamma; \Delta) \vdash^{\dagger} \texttt{case}\ inl(r^*)^*\{x^*.s^{\dagger}|y^*.u^{\dagger}\}^{\dagger} : C$,
then by Generation Lemma 5.6 and 6.3 there must be some $A$ and $B$ such that either
$(\Gamma'; \Delta') \vdash^{+} inl(r^{+})^{+} : A \vee B$ and $(\Gamma'', x^{+} : A; \Delta'') \vdash^{\dagger} s^{\dagger} : C$ and $(\Gamma''', y^{+} : B; \Delta''') \vdash^{\dagger} u^{\dagger} : C$ or
$(\Gamma'; \Delta') \vdash^{-} inl(r^{-})^{-} : A \wedge B$ and $(\Gamma''; \Delta'', x^{-} : A) \vdash^{\dagger} s^{\dagger} : C$ and $(\Gamma'''; \Delta''', y^{-} : B) \vdash^{\dagger} u^{\dagger} : C$
Thus, again by Generation Lemma 5.4 and 6.1 it follows that either
$(\Gamma'; \Delta') \vdash^{+} r^{+} : A$ or
$(\Gamma'; \Delta') \vdash^{-} r^{-} : A$.
Therefore, by the Substitution Lemma either
$(\Gamma', \Gamma''; \Delta', \Delta'') \vdash^{\dagger} s[r^{+}/x^{+}]^{\dagger} : C$ or
$(\Gamma', \Gamma''; \Delta', \Delta'') \vdash^{\dagger} s[r^{-}/x^{-}]^{\dagger} : C$. □

The philosophical importance of having established subject reduction for this calculus will be made explicit below (see Section 6.3). From now on we will omit the superscripts of subterms in the cases where the superscript of the whole term clearly determines the other polarities, i.e., instead of, e.g., $(\lambda x^{+}.t^{+})^{+}$ writing $(\lambda x.t)^{+}$ suffices.

### 6.2.2 Duality in $\lambda^{2Int}$

Now we want to examine a bit closer how the polarities in $\lambda^{2Int}$ relate to each other, and thereby, more generally speaking, the relation between proofs and refutations in this system. Therefore, we will define dualities in $\lambda^{2Int}$ and then prove our Dualization Theorem, which will be the key feature for the philosophical implications discussed in the next section.

**Definition 6.6**
*We will define a duality function d mapping types to their dual types, terms to their dual terms and contexts to their dual contexts as follows:*[89]

1. $d(\rho) = \rho$
2. $d(\top) = \bot$
3. $d(\bot) = \top$
4. $d(A \wedge B) = d(A) \vee d(B)$
5. $d(A \vee B) = d(A) \wedge d(B)$
6. $d(A \to B) = d(B) \prec d(A)$
7. $d(A \prec B) = d(B) \to d(A)$
8. $d(x^*) = x^d$
9. $d(abort(t^*)^{\dagger}) = abort(d(t^*))^d$

[89] The superscript $^d$ is used to indicate the dual polarity of whatever polarity $^*$ stands for in its respective dual version.

10. $d(\langle t^*, s^* \rangle^*) = \langle d(t^*), d(s^*) \rangle^d$
11. $d(inl(t^*)^*) = inl(d(t^*))^d$
12. $d(inr(t^*)^*) = inr(d(t^*))^d$
13. $d((\lambda x^*.t^*)^*) = (\lambda d(x^*).d(t^*))^d$
14. $d(\{t^+, s^-\}^*) = \{d(s^+), d(t^-)\}^d$
15. $d(fst(t^*)^*) = fst(d(t^*))^d$
16. $d(snd(t^*)^*) = snd(d(t^*))^d$
17. $d(\texttt{case}\ r^*\{x^*.s^\dagger | y^*.t^\dagger\}^\dagger) = \texttt{case}\ d(r^*)\{d(x^*).d(s^\dagger) | d(y^*).d(t^\dagger)\}^d$
18. $d(App(s^*, t^*)^*) = App(d(s^*), d(t^*))^d$
19. $d(\pi_1(t^*)^\dagger) = \pi_2(d(t^*))^d$
20. $d(\pi_2(t^*)^\dagger) = \pi_1(d(t^*))^d$
21. $d((\Gamma; \Delta)) = (d(\Delta); d(\Gamma))$, *with* $d(\Delta) = \{d(t^*) \mid t^* \in \Delta\}$, *resp. for* $d(\Gamma)$

**Theorem 6.2** (Dualization)
*If* $(\Gamma; \Delta) \vdash^* t^* : A$ *with a height of derivation at most* $n$*, then* $(d(\Delta); d(\Gamma)) \vdash^d d(t^*) : d(A)$ *(called its* dual derivation*) with a height of derivation at most* $n$*. This means that whenever we have a proof (refutation) of a formula, we can construct a refutation (proof) with the same height of its dual formula in our system.*

*Proof.* By induction on the height of derivation $n$ using the Generation Lemma.

If $n = 0$, then one of the four cases holds:

1. $(x^+ : A; \emptyset) \vdash^+ x^+ : A$
2. $(\emptyset; x^- : A) \vdash^- x^- : A$
3. $(\emptyset; \emptyset) \vdash^+ t^+ : \top$
4. $(\emptyset; \emptyset) \vdash^- t^- : \bot$

In case 1 the dual derivation is $(\emptyset; x^- : d(A)) \vdash^- x^- : d(A)$.
In case 2 the dual derivation is $(x^+ : d(A); \emptyset) \vdash^+ x^+ : d(A)$.
In case 3 the dual derivation is $(\emptyset; \emptyset) \vdash^- t^- : \bot$.
In case 4 the dual derivation is $(\emptyset; \emptyset) \vdash^+ t^+ : \top$.
All dual derivations can be trivially constructed with a height of $n = 0$.

Assume height-preserving dualization up to derivations of height at most $n$.

If $(\Gamma; \Delta) \vdash^* abort(t^+)^* : A$ is of height $n + 1$, then (by Generation Lemma 2.1) we have $(\Gamma; \Delta) \vdash^+ t^+ : \bot$ with height at most $n$. If $(\Gamma; \Delta) \vdash^* abort(t^-))^* : A$ is of height $n+1$, then (by Generation Lemma 2.2) we have $(\Gamma; \Delta) \vdash^- t^- : \top$ with height at most $n$.

Then by inductive hypothesis $(d(\Delta); d(\Gamma)) \vdash^- d(t^+) : d(\bot)$, resp. $(d(\Delta); d(\Gamma)) \vdash^+ d(t^-) : d(\top)$ are of height at most $n$ as well.

By application of $\top E^d$, resp. $\bot E$, we can construct a derivation of height $n+1$ s.t. $(d(\Delta); d(\Gamma)) \vdash^* abort(d(t^+))^* : d(A)$, resp. $(d(\Delta); d(\Gamma)) \vdash^* abort(d(t^-))^* : d(A)$ with $*$ being the dual polarity of $*$ in the original derivations. By our definition of dual terms $d(abort(t^+)^*) = abort(d(t^+))^d$ and $d(abort(t^-)^*) = abort(d(t^-))^d$.

If $(\Gamma; \Delta) \vdash^+ \langle s^+, t^+ \rangle^+ : A \wedge B$, resp. $(\Gamma; \Delta) \vdash^- \langle s^-, t^- \rangle^- : A \vee B$ is of height $n+1$, then (by Generation Lemma 5.1, resp. 6.4) we have $(\Gamma; \Delta) \vdash^+ s^+ : A$ and $(\Gamma'; \Delta') \vdash^+ t^+ : B$, resp. $(\Gamma; \Delta) \vdash^- s^- : A$ and $(\Gamma'; \Delta') \vdash^- t^- : B$ with height at most $n$.

Then by inductive hypothesis $(d(\Delta); d(\Gamma)) \vdash^- d(s^+) : d(A)$ and $(d(\Delta'); d(\Gamma')) \vdash^- d(t^+) : d(B)$, resp. $(d(\Delta); d(\Gamma)) \vdash^+ d(s^-) : d(A)$ and $(d(\Delta'); d(\Gamma')) \vdash^+ d(t^-) : d(B)$ are of height at most $n$ as well.

By application of $\vee I^d$, resp. $\wedge I$, we can construct a derivation of height $n+1$ s.t. $(d(\Delta); d(\Gamma)) \vdash^- \langle d(s^+), d(t^+) \rangle^- : d(A) \vee d(B)$, resp. $(d(\Delta); d(\Gamma)) \vdash^+ \langle d(s^-), d(t^-) \rangle^+ : d(A) \wedge d(B)$. By our definition of dual terms $d(\langle s^*, t^* \rangle^*) = \langle d(s^*), d(t^*) \rangle^d$.

For the further cases, see Appendix. □

Let us have a look at an example here considering the following derivation:

$$\dfrac{\dfrac{\dfrac{}{t^+ : \top}\,{\scriptstyle \top I} \qquad \dfrac{\dfrac{[x^+ : A \prec B]}{\pi_1(x^+)^+ : A}\,{\scriptstyle \prec E_1} \qquad \dfrac{[x^+ : A \prec B]}{\pi_2(x^+)^- : B}\,{\scriptstyle \prec E_2}}{\{\pi_1(x^+)^+, \pi_2(x^+)^-\}^- : A \to B}\,{\scriptstyle \to I^d}}{\{t^+, \{\pi_1(x^+)^+, \pi_2(x^+)^-\}^-\}^+ : \top \prec (A \to B)}\,{\scriptstyle \prec I}}{(\lambda x^+.\{t^+, \{\pi_1(x^+)^+, \pi_2(x^+)^-\}^-\}^+)^+ : (A \prec B) \to (\top \prec (A \to B))}\,{\scriptstyle \to I}$$

Now we dualize the term and the formula by our duality function $d$ yielding the following:

$d((\lambda x^+.\{t^+, \{\pi_1(x^+)^+, \pi_2(x^+)^-\}^-\}^+)^+) =$
$(\lambda x^-.\{\{\pi_1(x^-)^+, \pi_2(x^-)^-\}^+, t^-\}^-)^-$

$d((A \prec B) \to (\top \prec (A \to B))) = ((B \prec A) \to \bot) \prec (B \to A)$

We can now build a derivation with the dualized term $(\lambda x^-.\{\{\pi_1(x^-)^+, \pi_2(x^-)^-\}^+, t^-\}^-)^-$ as end-term being of the type of the dualized formula $((B \prec A) \to \bot) \prec (B \to A)$:

$$\dfrac{\dfrac{\dfrac{\dfrac{\llbracket x^- : B \to A \rrbracket}{\pi_1(x^-)^+ : B}\,{\scriptstyle \to E_1^d} \qquad \dfrac{\llbracket x^- : B \to A \rrbracket}{\pi_2(x^-)^- : A}\,{\scriptstyle \to E_2^d}}{\{\pi_1(x^-)^+, \pi_2(x^-)^-\}^+ : B \prec A}\,{\scriptstyle \prec I} \qquad \dfrac{}{t^- : \bot}\,{\scriptstyle \bot I^d}}{\{\{\pi_1(x^-)^+, \pi_2(x^-)^-\}^+, t^-\}^- : (B \prec A) \to \bot}\,{\scriptstyle \to I^d}}{(\lambda x^-.\{\{\pi_1(x^-)^+, \pi_2(x^-)^-\}^+, t^-\}^-)^- : ((B \prec A) \to \bot) \prec (B \to A)}\,{\scriptstyle \prec I^d}$$

The duality between those derivations is literally 'visible' in that they look like the mirrored version of each other with respect to the construction of the proof tree and the use of single and double lines. At each step we have the dual terms with the dual types applied according to the respective dual rules. So the case can be made - and this is what I want to argue for in the next section - that these derivations represent essentially the *same* underlying construction, although in one case it is delivered as a proof and in the other as a refutation.

## 6.3 Meaning and identity of proofs in `2Int`

I will lay out a conception of a Fregean distinction between sense and denotation of proofs based on (Tranchini, 2016) and (Ayhan, 2021b). The background of this conception is located in the area of what has been called *general proof theory*, which purports the idea that proofs are interesting objects of study in their own right, and *proof-theoretic semantics* (PTS), which can be seen as a more narrow understanding of this.[90] PTS opposes the traditional conception of model-theoretic semantics in that it takes the meaning of logical connectives not to be given in terms of truth tables, first-order models, etc. but by the rules that govern their use in inferences. The very general thought, then, underlying a Fregean distinction of sense and denotation of proofs, is simply that there are different ways to deliver a derivation of the same proof. A standard example for this would be two derivations, one being in non-normal form and the other being in its respective normal form. We will distinguish (as it is also done in the literature, e.g., in Kreisel (1971); Martin-Löf (1975); Prawitz (1971)) for these purposes between a *proof* as the underlying object (conceived of as a mental entity in line with the intuitionistic tradition) and a *derivation* as its respective linguistic representation. Since the derivation in normal form is the most direct way of representing the proof, it can be argued that this captures the denotation best. Thus, in the case with a derivation in normal form and one in non-normal form, though reducible to the former, the denotation would be the same since they share the same normal form. Their sense would differ, however, because the way of representing the denotation is essentially different in these cases. Such a conception for proofs can be found, e.g., in Girard (1989), but has not drawn considerable attention in the standard literature on this topic. Tranchini (2016) spelled out in more detail, then, how such a distinction could be usefully applied in the context of PTS. He argues that a derivation can only have sense if all the rules applied in it have reductions available (as opposed to, e.g., rules for `tonk`), since the reductions are what transfers a derivation into its normal form, i.e., its denotation.[91] Thus, the reductions are the way to get to the denotation of proofs which seems to

[90] For background literature on this, see, e.g., (Kreisel, 1971; Prawitz, 1971, 1973; Schroeder-Heister, 2022).

[91] Tranchini uses this framework to distinguish between derivations which have sense *and* denotation ('normal', well-behaved proofs), derivations which have sense, yet lack denotation (paradoxical derivations, since reductions can be applied to them but they cannot be brought into normal form by this) and derivations having neither sense nor denotation (which would be, e.g., ones containing a connective like `tonk`, for which there are no reductions available at all).

fit a Fregean conception of sense nicely.

(Ayhan, 2021b) builds upon Tranchini's idea in that the criterion for a derivation to have sense is adopted but also further developed by firstly, transferring it to a setting with $\lambda$-term-annotated proof systems and secondly, by giving a concrete account of what constitutes the sense of a derivation. While the denotation of a derivation in such systems is the normal form of the $\lambda$-term annotating the conclusion of the derivation, the sense of the derivation is taken to be the set of all $\lambda$-terms occurring within the derivation since these reflect the operations used in the derivation. Thus, they can be seen as encoding a procedure that takes us to the denotation, since the procedure finally yields the end-term.[92] The benefit of this is that we can thereby not only speak of sameness when it comes to denotation, i.e., identity of proofs, but also about sameness when it comes to sense, which would be the question of *synonymy* of proofs. Also, by using the $\lambda$-terms we can compare sense and denotation across different kinds of proofs systems, e.g., between natural deduction and sequent calculus systems.

So, what can we say on this basis about sense and denotation of proofs and dual proofs in $\lambda^{\texttt{2Int}}$? Having established that we have well-behaved reductions (by the Subject Reduction Theorem), we can safely assume that our derivations in $\lambda^{\texttt{2Int}}$ do have sense and (although we have not proven a normalization theorem for the terms here) since a normal form theorem is proven for the non-annotated natural deduction system `N2Int`, we can assume that they have denotation as well. What I want to argue for is that in this system it seems reasonable to extend our criterion for identity of proofs along the following lines. What the dualization theorem states is that for every proof, resp. dual proof, of a formula, we can express it as one or the other. This gives us a - from a bilateralist viewpoint - perfect balance in our system: There is no priority for proofs! Thus, proofs and dual proofs should be viewed as two sides of one coin. Taking this image seriously, what this amounts to is taking them as different representations of the *same* object, i.e., proofs and refutations of the respective dual formulas are *essentially the same*. So, my claim is - in Fregean terminology - that those derivations have the same denotation, because the underlying construction is *one* object, but they differ in sense because the way it is represented is essentially different.

Is it intuitive, though, to identify proofs and refutations, given that they are seemingly rather quite the opposite of each other? In the traditional literature on falsification, e.g., Nelson (1949) and López-Escobar (1972), such a thought is indeed

[92]For such an interpretation of Fregean sense, see, e.g., Dummett (1973, pp. 232, 323, 636) speaking of a "procedure" to determine the denotation, (1973, p. 96) "names with different senses but the same referent correspond to different routes leading to the same destination", Girard (1989, p. 2) "a sequence of *instructions*", or Horty (2007, pp. 66-69) "senses as procedures". Girard even mentions this in the context of relating this to "proofs as programs", i.e., a Curry-Howard conception.

expressed, namely that one and the same entity can act as verifying one formula while falsifying another. And I also think there are cases from mathematical reasoning or from our empirical way of 'proving' something, e.g., in court, where it makes a lot of sense to do so, i.e., where we have one and the same construction/evidence/etc. yielding a proof of a proposition while simultaneously refuting the dual proposition. If we think of what a proof of the statement "11 is a prime number" would look like, it would probably be something along the following lines: A constructive way of showing that 11 is a natural number greater than 1 and is not a product of two smaller natural numbers. So, basically we could have a program running through all the natural numbers up to 11 and checking whether they could form 11 as a product. If this is not the case, then we have our proof. The same program, however, could be used just as well to refute the statement "11 is a composite number". Or to take a 'real-life' example, let us suppose we are in court and a video is shown recording person X shooting person Z, while it is person Y being in the dock. Given that the video tape has been checked by experts for authenticity, it is clear in lighting, etc., this video would probably be taken as a refutation of the court's charges that Y is the murderer of Z, or, to put it in a slightly odd way in natural language, it would be taken as proof that Y is a non-murderer of Z.[93]

Similar suggestions, drawing on Nelson and López-Escobar, have been made in the more recent literature as well, e.g., by Wansing (2016b) and Ferguson (2020) for Nelson's constructive logic with strong negation, `N4`, namely that a construction $c$ can be taken as a proof of $A$ iff $c$ is a disproof of $\sim A$ (and vice versa). Ferguson (2020, p. 1507) argues here that there can be a coextensionality between a verifier of one formula and a falsifier of another, which, however, would not entail their identity because their sense differs. While I would agree with the latter, I think that we must distinguish here very carefully what we mean by 'identity'. Whereas Ferguson seems to understand it as 'being the same on all levels', I would understand it in the Fregean way, for which Frege uses '='. Identity between '$a$' and '$b$' means that they have the same denotation but not necessarily that they have the same sense. The intuition that verifiers and falsifiers may, in certain settings, be coextensional while differing in sense, can be well captured with the system proposed here, though. However, from a bilateralist point of view, it is in my opinion preferable not to have strong negation as a primitive connective in the language since, in a way, it stands against the bilateralist idea that refutation (or denial, or rejection, etc.) is a concept *prior* to negation. To briefly explain my concerns here: In `2Int` you can have a negation, even two, namely the intuitionistic negation, defined by $A \to \bot$ and the dual intuitionistic negation, which we call *co-negation*, defined by $\top \prec A$. Since they are defined

[93] Of course, in our natural language we do not have a strict definition of a "dual proposition" but one can come up with intuitive examples as I tried to do here. "(Dual) proposition" and "(dual) statement" are used synonymously in this context.

via implication and co-implication, which are manifestations of the two derivability relations in the object language, this seems to me in accordance with refutation (and here also proof) being the more primitive concept upon which negation is defined. However, incorporating strong negation would mean to have a *primitive* connective that is basically expressing exactly what is expressed by our derivability relations. Firstly, I simply do not see the need for that.[94] What is more, this would give strong negation a special place from the bilateralist PTS point of view as opposed to the other connectives, which do not express this relation between proofs and refutations. I do not deem this desirable if we have the bilateralist meaning-giving component already incorporated in the proof system via the derivability relations. Finally, strong negation would be non-congruential in our system, which leads to problems from a PTS point of view when it comes to the question of uniqueness.[95]

So, I think we should read the '=' used in Definition 5 for the mapping of terms to their dual terms in the Fregean way telling us that we should identify those terms in the sense that they refer to the same object. Thus, if we take the denotation of derivations to be referred to by their end-term in normal form, in a bilateralist setting derivational constructions[96] are not only identified by them being encoded by the same $\beta$-normal $\lambda$-term[97] but also by them being connected via the duality function $d$. The sense, however, must clearly be more fine-grained and thus, should not be identified over the duality function, just as it is not identified over its $\beta$-normal form: Remember, two derivations ending on different, though $\beta$-normal-form-equal, end-terms always differ in sense. The reason for this is that the way the proof object is presented is taken to be essentially different. In the context of comparing proofs and

[94]Since we do have negations, an objection coming from a "Frege-Geach-point" angle (see, e.g., Horwich, 2005), that we need a negation in our language to express it in subclauses of sentences (where an interpretation as refutation would not suffice), does not seem to be a concern here. This question came up in a discussion about an earlier draft of this paper with Dave Ripley, whom I want to thank for helping to clarify my thoughts on this.

[95]See (Humberstone, 2011, p. 579f.) and (2020a, p. 183, 187) on this, or on the issue of uniqueness specifically in bilateralist systems (Ayhan, 2021a). An example to show that strong negation would be non-congruential are the formulas $\sim (A \to B)$ and $A \wedge \sim B$, which are interderivable only w.r.t. $\vdash^+$ and $\sim\sim (A \to B)$ and $\sim (A \wedge \sim B)$, which are interderivable only w.r.t. $\vdash^-$. Note how this is different from the non-congruentiality described in Chapter 5.3.3 caused by the same example with co-negation: Since strong negation would be a primitive connective, for this connective uniqueness could not be retained by the definition for bilateralist settings given in that chapter.

[96]I use this somewhat clumsy expression instead of "proof objects" in order to avoid sounding like giving preference to proofs over dual proofs again.

[97]We will pass over the question here whether terms in $\beta$- or $\beta$-$\eta$-normal form should be identified. The literature is divided with respect to that question. Martin-Löf (1975, p. 100), for example, does not agree that $\eta$-conversions are identity-preserving. Prawitz (1971, p. 257), on the other hand, seems to lean towards it when he claims that it would seem "unlikely that any interesting property of proofs is sensitive to differences created by an expansion". He does not make a clear decision on that, though. Widebäck (2001), relating to results in the literature on the typed $\lambda$-calculus like (Friedman, 1975) and (Statman, 1983), argues for $\beta$-$\eta$-equality to give the right account of identity of proofs and Girard (1989, p. 16) does the same, although he also mentions that $\eta$-equations "have never been given adequate status" compared to the $\beta$-equations. For our purposes here it suffices to go with the 'safe' option of considering $\beta$-normal form only.

refutations it is the same: Although one can argue that the underlying derivational construction is the same, the way it is constructed is essentially different; in the one case by proving something and in the other by refuting something.

Let us use the following exemplary derivations to illustrate this point:

$$\frac{[x^+ : p]}{(\lambda x.x)^+ : p \to p} \to_I \qquad \frac{[x^+ : q]}{(\lambda x.x)^+ : q \to q} \to_I \qquad \frac{[y^+ : q]}{(\lambda y.y)^+ : q \to q} \to_I$$

$$\frac{[\![x^- : p]\!]}{(\lambda x.x)^- : p \prec p} \prec_I^d \qquad \frac{[\![x^- : q]\!]}{(\lambda x.x)^- : q \prec q} \prec_I^d \qquad \frac{[\![y^- : q]\!]}{(\lambda y.y)^- : q \prec q} \prec_I^d$$

In these cases the respective derivations on the vertical as well as on the diagonal axes are different in sense but not in denotation since their end-terms can be obtained from each other by our duality function. For this it does not matter that different formulas are derived because what we are interested in is not the denotation of the formulas but of the derivation, i.e., the structure of the construction is decisive here. The same holds for derivations with not only the same denotation but also the *same sense*, which we have on the horizontal axes. Although the signs, which are used, differ from each other, this difference is negligible because when it comes to the meaning of *derivations* (not formulas or propositions etc.), it should not make a difference which atomic formulas are chosen as long as the derived formula is structurally the same. In terms of type theory we can say that it makes no difference as long as the *principal type* of the term, i.e., the most general type that can be assigned to a term, is the same.

This is also why I prefer to use a Curry-style typing over a Church-style typing. In the latter system each term is usually uniquely typed, i.e., we would get a collapse of signs and sense: Since the sense is constituted by the terms occurring in a derivation, a differently typed term would automatically lead to a different sense. What leads to a difference in sense in our system is a difference in the principal types of the terms or a difference in the polarities. That the polarity makes a difference in sense can also be motivated by looking at substitution, which was one of Frege's motivations to make this distinction: In intensional contexts we cannot substitute expressions with a different sense *salva veritate*. Here, we cannot substitute same terms of different polarities for one another. In the first derivation on the left above, e.g., we could not just substitute $x^+$ with $x^-$ and leave the rest unchanged. The application of $\to_I$ would not be feasible. Thus, derivations can be claimed to constitute an intensional context. This seems at least not inappropriate for the general idea of PTS, which is, as Schroeder-Heister (2022) puts it, "intensional in spirit, as it is interested in proofs and not just provability".[98]

I want to make two final remarks about possible concerns that might be raised. First of all, it should be emphasized that identifying proofs and refutations does

[98]See also (Schroeder-Heister, 2016) and (Tranchini, 2021) about an intensional notion of harmony.

*not* mean to ultimately retreat to unilateralism again just because we have only one underlying object. To begin with, unilateralism means more than relying on one concept. It rather means to favor a certain concept, more specifically a 'positive' one, over the other 'negative' one. This does not happen here, though: both proofs and refutations are on a par, they are just conceived of being different ways to do the same thing. Furthermore, the identity between proofs and refutations is only stated for the denotation, not for the sense, though. Thus, we still do have means to distinguish between proofs and refutations, i.e., there is not a complete collapse between these two concepts. Secondly, it has been remarked and questioned whether we can truly speak of *one* underlying account to distinguish sense and denotation of derivations given that there is apparently a fundamental difference between this account in a unilateral vs. a bilateral setting. It is true that there is an asymmetry here because in the bilateral setting we always have two senses of the same denotation which are on equal standing, unlike in the unilateral case, where differences in sense often go along with non-normality vs. normality. But first of all, that is not necessarily so: we could have two non-normal derivations (in a unilateral setting) reducing to the same normal form and thus, we would have two different senses without one being the 'prior' one. Moreover, I do not see why this is a worry about this still being the same account: unilateralism and bilateralism *are* fundamentally different, so the fact that the application of my account to these settings also delivers different outcomes does not seem surprising to me.

## 6.4 Conclusion

In this paper I established an extension of the $\lambda$-calculus with which a natural deduction system for the logic `2Int`, containing proofs and refutations, can be annotated. For this system, called $\lambda^{2Int}$, I proved certain properties, which are typically considered important for $\lambda$-calculi, such as subject reduction. Furthermore, using a duality function for terms and types, I established and proved a duality theorem, which states that for every proof (resp. for every refutation) of a formula, a refutation (resp. a proof) of the of the dual formula can be given in this system. On this basis, I argued that proofs and refutations should be identified when they are connected by our duality function, since the underlying construction of the derivations is fundamentally the same. In a Fregean manner of distinguishing sense and denotation, then, proofs and refutations can be seen to have the same denotation, while, being presented in a different way, having a different sense. Thus, we have a, from a bilateralist point of view, very desirable equality between proofs and refutations. Neither is reduced to the other but rather both are considered to be on equal footing, since they are simply different ways to present the same object.

## 6.5 Appendix

### 6.5.1 Definition of compatibility

**Definition 6.7**
*A binary relation* $\mathscr{R}$ *on* $\mathtt{Term}_{2Int}$ *is* compatible *iff it satisfies the following clauses for all* $t, r, s, u \in \mathtt{Term}_{2Int}$*:*

1. *If* $t\mathscr{R}r$ *then* $abort(t^*)^\dagger \mathscr{R} abort(r^*)^\dagger$.
2. *If* $t\mathscr{R}r$ *then* $\langle t^*, s^*\rangle^* \mathscr{R} \langle r^*, s^*\rangle^*$.
3. *If* $t\mathscr{R}r$ *then* $\langle s^*, t^*\rangle^* \mathscr{R} \langle s^*, r^*\rangle^*$.
4. *If* $t\mathscr{R}r$ *then* $inl(t^*)^* \mathscr{R} inl(r^*)^*$.
5. *If* $t\mathscr{R}r$ *then* $inr(t^*)^* \mathscr{R} inr(r^*)^*$.
6. *If* $t\mathscr{R}r$ *then* $(\lambda x^*.t^*)^* \mathscr{R} (\lambda x^*.r^*)^*$*, for all variables* $x$.
7. *If* $t\mathscr{R}r$ *then* $\{t^+, s^-\}^* \mathscr{R} \{r^+, s^-\}^*$.
8. *If* $t\mathscr{R}r$ *then* $\{s^+, t^-\}^* \mathscr{R} \{s^+, r^-\}^*$.
9. *If* $t\mathscr{R}r$ *then* $fst(t^*)^* \mathscr{R} fst(r^*)^*$.
10. *If* $t\mathscr{R}r$ *then* $snd(t^*)^* \mathscr{R} snd(r^*)^*$.
11. *If* $t\mathscr{R}r$ *then* $\mathtt{case}\ t^*\{x^*.s^\dagger | y^*.u^\dagger\}^\dagger \mathscr{R}\, \mathtt{case}\ r^*\{x^*.s^\dagger | y^*.u^\dagger\}^\dagger$*, for all variables* $x, y$.
12. *If* $t\mathscr{R}r$ *then* $\mathtt{case}\ s^*\{x^*.t^\dagger | y^*.u^\dagger\}^\dagger \mathscr{R}\, \mathtt{case}\ s^*\{x^*.r^\dagger | y^*.u^\dagger\}^\dagger$*, for all variables* $x, y$.
13. *If* $t\mathscr{R}r$ *then* $\mathtt{case}\ s^*\{x^*.u^\dagger | y^*.t^\dagger\}^\dagger \mathscr{R}\, \mathtt{case}\ s^*\{x^*.u^\dagger | y^*.r^\dagger\}^\dagger$*, for all variables* $x, y$.
14. *If* $t\mathscr{R}r$ *then* $App(t^*, s^*)^* \mathscr{R} App(r^*, s^*)^*$.
15. *If* $t\mathscr{R}r$ *then* $App(s^*, t^*)^* \mathscr{R} App(s^*, r^*)^*$.
16. *If* $t\mathscr{R}r$ *then* $\pi_1(t^*)^\dagger \mathscr{R} \pi_1(r^*)^\dagger$.
17. *If* $t\mathscr{R}r$ *then* $\pi_2(t^*)^\dagger \mathscr{R} \pi_2(r^*)^\dagger$.

### 6.5.2 Proof of Dualization Theorem

*Proof of Dualization Theorem cont.* If $(\Gamma; \Delta) \vdash^+ inl(t^+)^+ : A \vee B$, resp. $(\Gamma; \Delta) \vdash^- inl(t^-)^- : A \wedge B$, is of height $n+1$, then (by Generation Lemma 6.1, resp. 5.4) we have $(\Gamma; \Delta) \vdash^+ t^+ : A$, resp. $(\Gamma; \Delta) \vdash^- t^- : A$ with height at most $n$.

Then by inductive hypothesis $(d(\Delta); d(\Gamma)) \vdash^- d(t^+) : d(A)$, resp. $(d(\Delta); d(\Gamma)) \vdash^+ d(t^-) : d(A)$ are of height at most $n$ as well.

By application of $\wedge I_1^d$, resp. $\vee I_1$, we can construct a derivation of height $n+1$ s.t. $(d(\Delta); d(\Gamma)) \vdash^- inl(d(t^+)^- : d(A) \wedge d(B)$, resp. $(d(\Delta); d(\Gamma)) \vdash^+ inl(d(t^-))^+ : d(A) \vee d(B)$. By our definition of dual terms $d(inl(t^*)^*) = inl(d(t^*))^d$. The same holds for the two cases of $inr(t^*)^*$.

If $(\Gamma; \Delta) \vdash^+ (\lambda x^+.t^+)^+ : A \to B$ is of height $n+1$, then (by Generation Lemma 3.1) we have $(\Gamma, x^+ : A; \Delta) \vdash^+ t^+ : B$ with height at most $n$.

Then by inductive hypothesis $(d(\Delta); d(\Gamma), x^- : d(A)) \vdash^- d(t^+) : d(B)$ is of height at most $n$ as well.

By application of $\prec I^d$ we can construct a derivation of height $n+1$ s.t. $(d(\Delta); d(\Gamma)) \vdash^- (\lambda x^-.d(t^+))^- : d(B) \prec d(A)$. By our definition of dual terms $d((\lambda x^+.t^+)^+) = (\lambda x^-.d(t^+))^-$.

If $(\Gamma; \Delta) \vdash^- (\lambda x^-.t^-)^- : A \prec B$ is of height $n+1$, then (by Generation Lemma 4.4) we have $(\Gamma; \Delta, x^- : B) \vdash^- t^- : A$ with height at most $n$.

Then by inductive hypothesis $(d(\Delta), x^+ : d(B); d(\Gamma)) \vdash^+ d(t^-) : d(A)$ is of height at most $n$ as well.

By application of $\to I$ we can construct a derivation of height $n+1$ s.t. $(d(\Delta); d(\Gamma)) \vdash^+ (\lambda x^+.d(t^-))^+ : d(B) \to d(A)$. By our definition of dual terms $d((\lambda x^-.t^-)^-) = (\lambda x^+.d(t^-))^+$.

If $(\Gamma; \Delta) \vdash^- \{s^+, t^-\}^- : A \to B$, resp. $(\Gamma; \Delta) \vdash^+ \{s^+, t^-\}^+ : A \prec B$ is of height $n+1$, then (by Generation Lemma 3.3, resp. 4.1) we have $(\Gamma; \Delta) \vdash^+ s^+ : A$ and $(\Gamma'; \Delta') \vdash^- t^- : B$ with height at most $n$.

Then by inductive hypothesis $(d(\Delta); d(\Gamma)) \vdash^- d(s^+) : d(A)$ and $(d(\Delta'); d(\Gamma')) \vdash^+ d(t^-) : d(B)$ are of height at most $n$ as well.

By application of $\prec I$, resp. $\to I^d$, we can construct a derivation of height $n+1$ s.t. $(d(\Delta); d(\Gamma)) \vdash^+ \{d(t^-), d(s^+)\}^+ : d(B) \prec d(A)$, resp. $(d(\Delta); d(\Gamma)) \vdash^- \{d(t^-), d(s^+)\}^- : d(B) \to d(A)$. By our definition of dual terms $d(\{s^+, t^-\}^*) = \{d(t^-), d(s^+)\}^d$.

If $(\Gamma; \Delta) \vdash^+ fst(t^+)^+ : A$ is of height $n+1$, then (by Generation Lemma 5.2) we have $(\Gamma; \Delta) \vdash^+ t^+ : A \wedge B$ with height at most $n$.

Then by inductive hypothesis $(d(\Delta); d(\Gamma)) \vdash^- d(t^+) : d(A) \vee d(B)$ are of height at most $n$ as well.

By application of $\vee E_1^d$ we can construct a derivation of height $n+1$ s.t. $(d(\Delta); d(\Gamma)) \vdash^- fst(d(t^+))^- : d(A)$ . By our definition of dual terms $d(fst(t^+)^+) = fst(d(t^+))^-$.

If $(\Gamma; \Delta) \vdash^+ snd(t^+)^+ : B$ is of height $n+1$, then (by Generation Lemma 5.3) we have $(\Gamma; \Delta) \vdash^+ t^+ : A \wedge B$ with height at most $n$.

Then by inductive hypothesis $(d(\Delta); d(\Gamma)) \vdash^- d(t^+) : d(A) \vee d(B)$ is of height at most $n$ as well.

By application of $\vee E_2^d$ we can construct a derivation of height $n+1$ s.t. $(d(\Delta); d(\Gamma)) \vdash^- snd(d(t^+))^- : d(B)$. By our definition of dual terms $d(snd(t^+)^+) = snd(d(t^+))^-$.

If $(\Gamma; \Delta) \vdash^- fst(t^-)^- : A$ is of height $n+1$, then (by Generation Lemma 6.5) we have $(\Gamma; \Delta) \vdash^- t^- : A \vee B$ with height at most $n$.

Then by inductive hypothesis $(d(\Delta); d(\Gamma)) \vdash^+ d(t^-) : d(A) \wedge d(B)$ is of height at most $n$ as well.

By application of $\wedge E_1$ we can construct a derivation of height $n+1$ s.t. $(d(\Delta); d(\Gamma)) \vdash^+ fst(d(t^-))^+ : d(A)$. By our definition of dual terms $d(fst(t^-)^-) = fst(d(t^-))^+$.

If $(\Gamma; \Delta) \vdash^- snd(t^-)^- : B$ is of height $n+1$, then (by Generation Lemma 6.6) we have $(\Gamma; \Delta) \vdash^- t^- : A \vee B$ with height at most $n$.

Then by inductive hypothesis $(d(\Delta); d(\Gamma)) \vdash^+ d(t^-) : d(A) \wedge d(B)$ are of height at most $n$ as well.

By application of $\wedge E_2$ we can construct a derivation of height $n+1$ s.t. $(d(\Delta); d(\Gamma)) \vdash^+ snd(d(t^-))^+ : d(B)$. By our definition of dual terms $d(snd(t^-)^-) = snd(d(t^-))^+$.

If $(\Gamma; \Delta) \vdash^* \texttt{case}\ r^+\{x^+.s^*|y^+.t^*\}^* : C$, resp. $(\Gamma; \Delta) \vdash^* \texttt{case}\ r^-\{x^-.s^*|y^-.t^*\}^* : C$ is of height $n+1$, then (by Generation Lemma 6.3, resp. 5.6) we have $(\Gamma; \Delta) \vdash^+ r^+ : A \vee B$, $(\Gamma', x^+ : A; \Delta') \vdash^* s^* : C$ and $(\Gamma'', y^+ : B; \Delta'') \vdash^* t^* : C$, resp. $(\Gamma; \Delta) \vdash^- r^- : A \wedge B$, $(\Gamma'; \Delta', x^- : A) \vdash^* s^* : C$ and $(\Gamma''; \Delta'', y^- : B) \vdash^* t^* : C$ with height at most $n$.

Then by inductive hypothesis $(d(\Delta); d(\Gamma)) \vdash^- d(r^+) : d(A) \wedge d(B)$, $(d(\Delta'); d(\Gamma'), x^- : d(A)) \vdash^d d(s^*) : d(C)$ and $(d(\Delta''); d(\Gamma''), y^- : d(B)) \vdash^d d(t^*) : d(C)$, resp. $(d(\Delta); d(\Gamma)) \vdash^+ d(r^-) : d(A) \vee d(B)$, $(d(\Delta'), x^+ : d(A); d(\Gamma')) \vdash^d d(s^*) : d(C)$ and $(d(\Delta''), y^+ : d(B); d(\Gamma'')) \vdash^d d(t^*) : d(C)$ are of height at most $n$ as well.
By application of $\wedge E^d$, resp. $\vee E$, we can construct a derivation of height $n+1$ s.t. $(d(\Delta); d(\Gamma)) \vdash^d \texttt{case}\ d(r^+)\{x^-.d(s^*)|y^-.d(t^*)\}^d : d(C)$, resp. $(d(\Delta); d(\Gamma)) \vdash^d \texttt{case}\ d(r^-)\{x^+.d(s^*)|y^+.d(t^*)\}^d : d(C)$.
By our definition of dual terms
$d(\texttt{case}\ r^+\{x^+.s^*|y^+.t^*\}^*) = \texttt{case}\ d(r^+)\{x^-.d(s^*)|y^-.d(t^*)\}^d$ and
$d(\texttt{case}\ r^-\{x^-.s^*|y^-.t^*\}^*) = \texttt{case}\ d(r^-)\{x^+.d(s^*)|y^+.d(t^*)\}^d$.

If $(\Gamma; \Delta) \vdash^+ App(s^+, t^+)^+ : B$ is of height $n+1$, then (by Generation Lemma 3.2) we have $(\Gamma; \Delta) \vdash^+ s^+ : A \to B$ and $(\Gamma'; \Delta') \vdash^+ t^+ : A$ with height at most $n$.

Then by inductive hypothesis $(d(\Delta); d(\Gamma)) \vdash^- d(s^+) : d(B) \prec d(A)$ and $(d(\Delta'); d(\Gamma')) \vdash^- d(t^+) : d(A)$ is of height at most $n$ as well.

By application of $\prec E^d$ we can construct a derivation of height $n+1$ s.t. $(d(\Delta); d(\Gamma)) \vdash^- App(d(s^+), d(t^+))^- : d(B)$.

By our definition of dual terms $d(App(s^+, t^+)^+) = App(d(s^+), d(t^+))^-$.

If $(\Gamma; \Delta) \vdash^- App(s^-, t^-)^- : B$ is of height $n+1$, then (by Generation Lemma 4.5) we have $(\Gamma; \Delta) \vdash^- s^- : B \prec A$ and $(\Gamma'; \Delta') \vdash^- t^- : A$ with height at most $n$.

Then by inductive hypothesis $(d(\Delta); d(\Gamma)) \vdash^+ d(s^-) : d(A) \to d(B)$ and $(d(\Delta'); d(\Gamma')) \vdash^+ d(t^-) : d(A)$ is of height at most $n$ as well.

By application of $\to E$ we can construct a derivation of height $n+1$ s.t. $(d(\Delta); d(\Gamma)) \vdash^+ App(d(s^-), d(t^-))^+ : d(B)$. By our definition of dual terms $d(App(s^-, t^-)^-) = App(d(s^-), d(t^-))^+$.

If $(\Gamma; \Delta) \vdash^+ \pi_1(t^-)^+ : A$ is of height $n+1$, then (by Generation Lemma 3.4) we have $(\Gamma; \Delta) \vdash^- t^- : A \to B$ with height at most $n$.

Then by inductive hypothesis $(d(\Delta); d(\Gamma)) \vdash^+ d(t^-) : d(B) \prec d(A)$ is of height at most $n$ as well.

By application of $\prec E_2$ we can construct a derivation of height $n+1$ s.t. $(d(\Delta); d(\Gamma)) \vdash^- \pi_2(d(t^-))^- : d(A)$. By our definition of dual terms $d(\pi_1(t^-)^+) = \pi_2(d(t^-))^-$.

If $(\Gamma; \Delta) \vdash^- \pi_2(t^+)^- : A$ is of height $n+1$, then (by Generation Lemma 4.3) we have $(\Gamma; \Delta) \vdash^+ t^+ : B \prec A$ with height at most $n$.

Then by inductive hypothesis $(d(\Delta); d(\Gamma)) \vdash^- d(t^+) : d(A) \to d(B)$ is of height at most $n$ as well.

By application of $\to E_1^d$ we can construct a derivation of height $n+1$ s.t. $(d(\Delta); d(\Gamma)) \vdash^+ \pi_1(d(t^+))^+ : d(A)$. By our definition of dual terms $d(\pi_2(t^+)^-) = \pi_1(d(t^+))^+$. □

# 7 Conclusion

This thesis comprises five individual papers dealing with various specific problems and topics in the realm of proof-theoretic semantics. However, the questions are not at all unrelated. In the first two papers (Chapters 2 and 3) I establish and elaborate criteria and conceptions for a theory about meaning and identity of proofs. In the second two papers (Chapters 4 and 5) I am concerned with a specific bilateralist logic, its proof systems and the related problems and virtues. Finally, in the last paper constituting this thesis (Chapter 6) these two sides are brought together by the question of how to deal with meaning and identity of derivations in a bilateralist setting.

In Chapter 2 the aim was to give a distinction between the meaning and denotation of proofs, following a Fregean conception of these concepts. While the question about the denotation of proofs has been treated in the literature of general proof theory, this is not the case for the question about meaning of proofs. My approach was to use a $\lambda$-term-annotated proof system, in which - following the standard conception in the literature - the end-term of a derivation is conceived of as referring to the *denotation* of the derivation. The *sense* of a derivation, on the other hand, then, consists in the set of all $\lambda$-terms occurring in the derivation since these reflect the operations of the derivation, i.e., they show the way how the end-term, and thus the denotation, is achieved. I argued that this approach has two advantages. Firstly, we get a more fine-grained criterion to distinguish and/or identify proofs since we can distinguish sameness of denotations, i.e., *identity* of proofs, but also sameness of sense, i.e., *synonymy* of proofs. Secondly, we can compare sense and denotation over different kinds of proof systems, which we showed for ND and SC. An important background assumption concerning the sense of derivations, though, is that the rules of the connectives appearing in the derivation have *reductions* available. I followed Tranchini's (2016) argumentation on this issue, who connects the sense of derivations with reductions being available because the reductions are needed to convert a derivation into its normal form, i.e., its denotation. Thus, they can be seen as a sequence of instructions leading us our way to the denotation, which is how Fregean sense has been characterized in the literature.

Since reductions are, therefore, not only relevant from a technical point of view in proof theory but also for philosophical considerations, it seems important to have a precise conception about what can count as a proper reduction for a system. This question has not been considered systematically, as is also mentioned in the papers that *are* concerned with this question (Schroeder-Heister & Tranchini, 2017, 2018). Therefore, this was the aim of Chapter 3. In this chapter I addressed the question, which reductions of derivations should be accepted as admissible and which ones should not. The background is that it has been argued that we should not only consider the 'common' reductions, which are directly connected to the rules for the respective connectives, but that anything that in any way cuts out a redundant detour of a derivation can be regarded as an (admissible) reduction. An example of such a reduction is given in (Ekman, 1994, 1998), on which basis it is argued that the proof-theoretic conception of paradoxes given in (Tennant, 1982) had to be abandoned. Schroeder-Heister and Tranchini's (2017) answer to this is that it is the underlying Ekman-reduction that is problematic and not the conception of paradoxes. Assuming a reduction-based conception of proof identity, they motivate their claim by showing that accepting Ekman-reduction in a system would result in a trivialization of this conception: we would have to identify all proofs of the same formula. While I agree very much with their reasoning, what I aimed for in Chapter 3 is to generalize their approach to make it more widely applicable. For this purpose, I showed, on the one hand, why Ekman-reduction cannot be considered acceptable even if one rejects a reduction-based conception of identity between derivations. Building on this, I then gave a general criterion, which can be fairly easily applied to test whether or not it holds for a given reduction. I call this criterion "weak subject reduction" in dependence of the well-known property of "subject reduction" stating that the type of a redex should be preserved under reduction, i.e., its contractum should be of the same type, which holds for reductions in most 'well-behaved' $\lambda$-calculi. Thus, as in Chapter 2, we are considering proof systems with $\lambda$-term annotations, for which the criterion can then be said to hold (or not). What weak subject reduction states is that for a given reduction either 'full' subject reduction must hold, or if not, then what cannot be the case is that redex and contractum are of arbitrarily different types. Arbitrariness is then defined in relation to what extent the rules of type assignment do (not) determine the type reconstruction of the redex. The philosophical motivation that supports my argumentation is that apart from committing to a certain kind of theory of proof identity, having reductions in our system that do not adhere to weak subject reduction would cause such an arbitrariness that the proof system could not be considered meaning-giving anymore.

In Chapter 4 I introduced `SC2Int`, a sequent calculus for the bi-intuitionistic logic `2Int,` displaying two derivability relations through its sequent signs. Thus, this calculus can be interpreted as bilateralist, capturing a relation of *proof* as well as one of *refutation*. By applying the proof methods that are used in (Negri & von Plato, 2001) for the calculus `G3ip`, I showed that `SC2Int` is a cut-free bilateralist sequent calculus and gave proofs for the admissibility of the structural rules of weakening, contraction and cut in the system.

In Chapter 5, then, I examined a specific proof-theoretic property for the logic `2Int`, namely uniqueness. This property, establishing that the rules defining our connectives actually determine *at most one* connective playing the inferential role that is given by the rules, is relevant from a PTS point of view since it seems like a reasonable demand to ask if we want the rules to be meaning-giving. There are several features in logical systems which may cause problems for uniqueness, like non-congruentiality, the specific representation of the proof system, etc. Chapter 5's contribution to the debate is that I examined the specific problems that occur in a bilateralist setting. In such a system the question could arise, according to the common understanding of uniqueness, whether a bilateral PTS does not lead to different meanings of the connectives depending on whether we prove or refute. I showed how such a conception can be avoided and proposed a characterization of uniqueness which can be applied to a bilateralist proof system in which we have two derivability relations, one for provability and one for dual provability. This comes down to requiring that interderivability of the formulas containing the connective is satisfied for both derivability relations. Finally, I compared this to other (related) proof systems and argue that having the bilateral aspect integrated via the derivability relations bears certain advantages as opposed to other proof-theoretic representations.

As mentioned above, Chapter 6 unites the threads of the previous chapters in a sense, since, on the one hand, it is about a bilateral system, and on the other hand, about sense and denotation in this system, for which, of course, the reductions are also of importance again. Using the natural deduction system of `2Int` as a basis again, in this chapter I designed an extension of the $\lambda$-calculus, which, having two-sorted terms, is suitable for annotating proofs as well as refutations in this system. Then I went on to prove some properties of that calculus, called $\lambda^{2Int}$, which are important for our purposes. Furthermore, as one of the main results, I established and proved a duality theorem, which states that for every proof (resp. for every refutation) of a formula $A$ that can be given in the system, a refutation (resp. a proof) of the dual formula of $A$ can be given. Based on this theorem, I subsequently argued that all derivations and their respective dual derivations should be identified. Thus, for every proof (resp. for every refutation) we can say that there is a refutation (resp. a proof), such that they have the same underlying denotation. This is motivated by the observation that the construction, which is the basis of the proof/refutation, is essentially the same. However, since the ways of representing the denotation are fundamentally different in these two cases, one by proving something and the other by refuting something, the sense must be said to differ here. Thus, we still have legitimate means to distinguish proofs and refutations, while at the same time having a strong basis to argue in a bilateralist spirit for the point that proofs and refutations are on a par. Since they are ultimately seen as the same object, which can be just expressed in different ways, neither is given preference over the other. My plan for future work on this topic is to use $\lambda^{2Int}$ to create also a term-annotated version of `SC2Int`. Thus, I would be able, just like I did in Chapter

2, to compare sense and denotation across ND and SC, now in a bilateralist setting.

In conclusion, although the aforementioned chapters are (or are planned to be) independent publications, the central theme of proof-theoretic semantics is present throughout this thesis and guides the argumentation essentially. While considering different and rather specific applications of the question about meaning and identity of proofs, my aim was to retain a coherent underlying framework at the same time. It is clear, though - as I briefly touched on in the introduction - that there are also other, very different approaches when it comes to conceptions about proof identity or bilateralism and thus, it will be interesting for future research to see what new insights those will give us in PTS and whether and how they will be compatible with the present framework.

# References


Ayhan, S. (2020). *A cut-free sequent calculus for the bi-intuitionistic logic 2Int.* (Unpublished Note) https://arxiv.org/abs/2009.14787.

Ayhan, S. (2021a). Uniqueness of logical connectives in a bilateralist setting. In M. Blicha & I. Sedlár (Eds.), *The Logica Yearbook 2020* (pp. 1–16). London: College Publications.

Ayhan, S. (2021b). What is the meaning of proofs? A Fregean distinction in proof-theoretic semantics. *Journal of Philosophical Logic*, *50*, 571-591. https://doi.org/10.1007/s10992-020-09577-2

Baader, F., & Nipkow, T. (1998). *Term rewriting and all that.* Cambridge: Cambridge University Press.

Barendregt, H. (1992). Lambda calculi with types. In S. Abramsky, D. M. Gabbay, & T. S. E. Maibaum (Eds.), *Handbook of logic in computer science* (Vol. 2, pp. 117–309). Oxford: Oxford University Press.

Barendregt, H., & Ghilezan, S. (2000). Lambda terms for natural deduction, sequent calculus and cut elimination. *Journal of Functional Programming*, *10(1)*, 121-134.

Belnap, N. D. (1962). Tonk, plonk and plink. *Analysis*, *22*(6), 130-134.

Brandom, R. B. (2000). *Articulating reasons: An introduction to inferentialism.* Cambridge, MA: Harvard University Press.

Cook, R. T. (2005). What's wrong with tonk(?). *Journal of Philosophical Logic*, *34*, 217–226. https://doi.org/10.1007/s10992-004-7805-x

Curry, H. B. (1934). Functionality in combinatory logic. *Proceedings of the National Academy of Science USA*, *20*(11), 584–690.

Curry, H. B., & Feys, R. (1958). *Combinatory logic.* Amsterdam: North Holland.

de Groote, P. (1999). On the strong normalisation of natural deduction with permutation-conversions. In P. Narendran & M. Rusinowitch (Eds.), *Rewriting techniques and applications: RTA 1999* (pp. 45–59). Berlin/Heidelberg: Springer.

Došen, K. (2003). Identity of proofs based on normalization and generality. *Bulletin of Symbolic Logic*, *9*, 477–503.

Došen, K. (2008). *Cut elimination in categories.* Berlin: Springer.

Došen, K., & Schroeder-Heister, P. (1985). Conservativeness and uniqueness. *Theoria*, *51*, 159–173.

Došen, K., & Schroeder-Heister, P. (1988). Uniqueness, definability and interpolation. *The Journal of Symbolical Logic*, *53*(2), 554–570.

Drobyshevich, S. (2019). Tarskian consequence relations bilaterally: some familiar notions. *Synthese*. `https://doi.org/10.1007/s11229-019-02267-w`.

Dummett, M. (1973). *Frege: Philosophy of language.* New York: Harper & Row.

Dummett, M. (1975). The philosophical basis of intuitionistic logic. In H. E. Rose & J. C. Shepherdson (Eds.), *Logic colloquium '73* (pp. 5–40). Amsterdam: North-Holland.

Dummett, M. (1976). What is a theory of meaning? (II). In G. Evans & J. McDowell (Eds.), *Truth and meaning: Essays in semantics* (pp. 67–137). Oxford: Clarendon Press.

Dummett, M. (1981). *Frege: Philosophy of language* (2nd ed.). London: Duckworth.

Dummett, M. (1991). *The logical basis of metaphysics*. London: Duckworth.

Duží, M., Jespersen, B., & Materna, P. (2010). *Procedural semantics for hyperintensional logic: Foundations and applications of transparent intensional logic*. Berlin: Springer.

Ekman, J. (1994). *Normal proofs in set theory* (Ph.D. dissertation). University of Göteborg.

Ekman, J. (1998). Propositions in propositional logic provable only by indirect proofs. *Mathematical Logic Quarterly*, *44*, 69–91. https://doi.org/10.1002/malq.19980440105

Evans, G. (1982). *The varieties of reference*. Oxford: Clarendon Press.

Ferguson, T. M. (2020). Explicit analyses of proof/refutation interaction for constructible falsity and Heyting—Brouwer logic. *Journal of Logic and Computation*, *30*(8), 1505–1540. https://doi.org/10.1093/logcom/exaa047

Ferreira, F. (2008). The co-ordination principles: A problem for bilateralism. *Mind*, *117*(468), 1051–1057.

Fitch, G., & Nelson, M. (2018). Singular propositions. In E. N. Zalta & U. Nodelman (Eds.), *The Stanford Encyclopedia of Philosophy* (Spring 2018 ed.). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/spr2018/entries/propositions-singular/

Francez, N. (2014a). Bilateralism in proof-theoretic semantics. *Journal of Philosophical Logic*, *43*(2-3), 239–259.

Francez, N. (2014b). Bilateral relevant logic. *The Review of Symbolic Logic*, *7*(2), 250–272.

Francez, N. (2015). *Proof-theoretic semantics*. London: College Publications.

Francez, N. (2017). On harmony and permuting conversions. *Journal of Applied Logic*, *21*, 14–23.

Francez, N. (2019). Bilateralism, trilateralism, multilateralism and poly-sequents. *Journal of Philosophical Logic*, *48*, 245–262.

Francez, N., & Dyckhoff, R. (2012). A note on harmony. *Journal of Philosophical Logic*, *41*, 613–628. https://doi.org/10.1007/s10992-011-9208-0

Frege, G. (1892). Über Sinn und Bedeutung. *Zeitschrift für Philosophie und philosophische Kritik*, *NF 100*, 25–50.

Frege, G. (1948[1892]). Sense and reference. *The Philosophical Review*, *57*(3), 209–230.

Frege, G. (1979). *Posthumous writings*. Oxford: Basil Blackwell.

Frege, G. (1993[1918/1919]). Die Verneinung. Eine logische Untersuchung. In G. Patzig (Ed.), *Gottlob Frege. Logische Untersuchungen* (4th ed., pp. 54–71). Göttingen: Vandenhoeck & Ruprecht.

Friedman, H. (1975). Equality between functionals. In R. Parikh (Ed.), *Logic colloquium: Lecture notes in mathematics* (Vol. 453, pp. 23–37). Berlin/Heidelberg: Springer.

Gabbay, M. (2017). Bilateralism does not provide a proof theoretic treatment of classical logic (for technical reasons). *Journal of Applied Logic*, *25*, 108–122.

Gentzen, G. (1964[1935]). Investigations into logical deduction. *American Philosophical Quarterly*, *1*(4), 288–306.

Gheorghiu, A., & Pym, D. (Submitted). From proof-theoretic validity to base-extension semantics for intuitionistic propositional logic.

Gibbard, P. (2002). Price and Rumfitt on rejective negation and classical logic. *Mind*, *111*, 297–303.

Girard, J.-Y. (1989). *Proofs and types*. Cambridge: Cambridge University Press.

Goré, R. (2000). Dual intuitionistic logic revisited. In R. Dyckhoff (Ed.), *Automated reasoning with analytic tableaux and related methods. tableaux 2000* (pp. 252–267). Berlin: Springer-Verlag.

Griffin, T. G. (1990). A formulae-as-types notion of control. In *Symposium on principles of programming languages* (pp. 47–58). ACM Press.

Hacking, I. (1979). What is logic? *The Journal of Philosophy*, *76)*(6), 285–319.

Herbelin, H. (1994). A lambda-calculus structure isomorphic to Gentzen-style sequent calculus structure. *Annual Conference for Computer Science Logic*, 61–75.

Hindley, J. R. (1997). *Basic simple type theory*. Cambridge: Cambridge University Press.

Hindley, J. R., & Seldin, J. P. (2008). *Lambda-Calculus and Combinators: an Introduction*. Cambridge: Cambridge University Press.

Hjortland, O. T. (2014). Speech acts, categoricity, and the meanings of logical connectives. *Notre Dame Journal of Formal Logic*, *55*(4), 445–467.

Horty, J. (2007). *Frege on definitions: A case study of semantic content*. Oxford: Oxford University Press.

Horwich, P. (2005). The Frege-Geach Point. *Philosophical Issues*, *15*, 78–93.

Howard, W. A. (1980[1969]). The formulae-as-types notion of construction. In J. P. Seldin & J. R. Hindley (Eds.), *To H. B. Curry: Essays on combinatory logic, lambda calculus and formalism* (pp. 479–490). London/New York: Academic Press.

Humberstone, L. (2000). The revival of rejective negation. *Journal of Philosophical*

*Logic*, *29*, 331–381.

Humberstone, L. (2011). *The connectives.* Cambridge, MA/London: MIT Press.

Humberstone, L. (2019). Priest on negation. In C. Başkent & T. M. Ferguson (Eds.), *Graham Priest on dialetheism and paraconsistency. Outstanding contributions to logic* (Vol. 18). Springer.

Humberstone, L. (2020a). Explicating logical independence. *Journal of Philosophical Logic*, *49*, 135–218.

Humberstone, L. (2020b). Sentence connectives in formal logic. In E. N. Zalta & U. Nodelman (Eds.), *The Stanford Encyclopedia of Philosophy* (Spring 2020 ed.). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/spr2020/entries/connectives-logic/

Kamide, N., & Wansing, H. (2012). Proof theory of Nelson's paraconsistent logic: A uniform perspective. *Theoretical Computer Science*, *415*, 1–38.

Kowalski, T., & Ono, H. (2017). Analytic cut and interpolation for bi-intuitionistic logic. *The Review of Symbolic Logic*, *10(2)*, 259–283.

Kreisel, G. (1971). A survey of proof theory II. In J. E. Fenstad (Ed.), *Proceedings of the Second Scandinavian Logic Symposium* (pp. 109–170). Amsterdam: North Holland.

Kremer, M. (1988). Logic and meaning: The philosophical significance of the sequent calculus. *Mind*, *97*, 50–72.

Kürbis, N. (2016). Some comments on Ian Rumfitt's bilateralism. *Journal of Philosophical Logic*, *45*, 623–644.

Kürbis, N. (2019). *Proof and falsity: A logical investigation.* Cambridge: Cambridge University Press.

Kürbis, N. (in press). Supposition: A problem for bilateralism. *Bulletin of the Section of Logic*.

Lindley, S. (2007). Extensional rewriting with sums. In S. R. D. Rocca (Ed.), *Typed lambda calculi and applications: TLCA 2007* (pp. 255–271). Berlin/Heidelberg: Springer.

López-Escobar, E. G. K. (1972). Refutability and elementary number theory. *Indigationes Mathematicae*, *75*(4), 362–374.

Makinson, D. (2014). On an inferential semantics for classical logic. *Logic Journal of the IGPL*, *22*(1), 147–154.

Martin-Löf, P. (1975). About models for intuitionistic type theories and the notion of definitional equality. In S. Kanger (Ed.), *Proceedings of the Third Scandinavian Logic Symposium* (pp. 81–109). Amsterdam: North Holland.

Martin-Löf, P. (1996). On the meanings of the logical constants and the justifications of the logical laws. *Nordic Journal of Philosophical Logic*, *1*(1), 11–60.

Martin-Löf, P. (2021). The sense/reference distinction in constructive semantics. *Bulletin of Symbolic Logic*, *27*(4), 501–513.

Muskens, R. (2005). Sense and the computation of reference. *Linguistics and Philosophy*, *28*(4), 473–504.

Naibo, A., & Petrolo, M. (2015). Are uniqueness and deducibility of identicals the same? *Theoria*, *81*, 143–181.

Negri, S., & von Plato, J. (2001). *Structural proof theory.* Cambridge/New York: Cambridge University Press.

Nelson, D. (1949). Constructible falsity. *The Journal of Symbolic Logic*, *14(1)*, 16–26.

Parigot, M. (1992). $\lambda\mu$-calculus: An algorithmic interpretation of classical natural deduction. In A. Voronkov (Ed.), *Logic programming and automated reasoning* (Vol. 624 of Lecture Notes in Artificial Intelligence, pp. 190–201). Springer.

Parigot, M. (1993). Classical proofs as programs. In G. Gottlob, A. Leitsch, & D. Mundici (Eds.), *Computational logic and proof theory* (Vol. 713 of Lecture Notes in Computer Science, pp. 263–276). Springer.

Pfenning, F. (2000). Structural cut elimination: I. intuitionistic and classical logic. *Information and Computation*, *157*, 84–141.

Postniece, L. (2010). *Proof theory and proof search of bi-intuitionistic and tense logic* (Unpublished doctoral dissertation). The Australian National University, Canberra.

Pottinger, G. (1977). Normalization as a homomorphic image of cut-elimination. *Annals of Mathematical Logic*, *12*, 323–357.

Prawitz, D. (1965). *Natural deduction: A proof-theoretical study.* Stockholm: Almqvist & Wiksell.

Prawitz, D. (1971). Ideas and results in proof theory. In J. E. Fenstad (Ed.), *Proceedings of the Second Scandinavian Logic Symposium* (pp. 235–307). Amsterdam: North Holland.

Prawitz, D. (1973). Towards a foundation of a general proof theory. In P. Suppes, L. Henkin, A. Joja, & G. C. Moisil (Eds.), *Logic, methodology, and philosophy of science IV* (pp. 225–250). Amsterdam: North Holland.

Prawitz, D. (1974). On the idea of a general proof theory. *Synthese*, *27*, 63–77.

Prawitz, D. (2006). Meaning approached via proofs. *Synthese*, *148*, 507–524.

Price, H. (1983). Sense, assertion, Dummett and denial. *Mind*, *92*(366), 161–173.

Prior, A. N. (1960). The runabout inference-ticket. *Analysis*, *21*(2), 38–39.

Rauszer, C. (1974). A formalization of the propositional calculus of H-B logic. *Studia Logica*, *33*(1), 23–34.

Read, S. (2010). General-elimination harmony and the meaning of the logical constants. *Journal of Philosophical Logic*, *39*, 557–576. https://doi.org/10.1007/s10992-010-9133-7

Restall, G. (1997). *Extending intuitionistic logic with subtraction.* (Online Notes) https://consequently.org/papers/extendingj.pdf

Restall, G. (2005). Multiple conclusions. In P. Hajek, L. Valdes-Villanueva, & D. Westerstahl (Eds.), *Logic, methodology and philosophy of science: Proceedings of the twelfth international congress* (pp. 189–205). London: King's College Publications.

Restall, G. (2013). Assertion, denial and non-classical theories. In K. Tanaka, F. Berto, E. Mares, & F. Paoli (Eds.), *Paraconsistency: Logic and applications* (pp. 81–99). Springer.

Restall, G. (2017). *Proof terms for classical derivations.* (Article in progress) https://consequently.org/papers/proof-terms.pdf

Restall, G. (2021). *Speech acts & the quest for a natural account of classical proof.* (Article in progress) https://consequently.org/papers/speech-acts-for-classical-natural-deduction.pdf

Restall, G. (in press). Structural rules in natural deduction with alternatives. *Bulletin of the Section of Logic*.

Ripley, D. (2011). Negation, denial, and rejection. *Philosophy Compass*, *6*, 622–629.

Ripley, D. (2013). Paradoxes and failures of cut. *Australasian Journal of Philosophy*, *91*(1), 139–164. https://doi.org/10.1080/00048402.2011.630010

Ripley, D. (2015). Anything goes. *Topoi*, *34*, 25–36. https://doi.org/10.1007/s11245-014-9261-8

Ripley, D. (2017). Bilateralism, coherence, warrant. In F. Moltmann & M. Textor (Eds.), *Act-based conceptions of propositional content: Contemporary and historical perspectives* (pp. 307–324). Oxford: Oxford University Press.

Ripley, D. (2020a). Denial. In V. Déprez & M. T. Espinal (Eds.), *The Oxford Handbook of Negation* (pp. 47–57). Oxford: Oxford University Press.

Ripley, D. (2020b). Strong normalization in core type theory. In I. Sedlár & M. Blicha (Eds.), *The Logica Yearbook 2019* (pp. 111–130). London: College Publications.

Rumfitt, I. (2000). 'Yes' and 'No'. *Mind*, *109*(436), 781–823.

Sandqvist, T. (2009). Classical logic without bivalence. *Analysis*, *69*(2), 211–218.

Sandqvist, T. (2012). The subformula property in classical natural deduction established constructively. *The Review of Symbolic Logic*, *5*(4), 710–719.

Sandqvist, T. (2015). Base-extension semantics for intuitionistic sentential logic. *Logic Journal of the IGPL*, *23*(5), 719–731.

Schroeder-Heister, P. (2009). Sequent calculi and bidirectional natural deduction: On the proper basis of proof-theoretic semantics. In M. Peliš (Ed.), *The logica yearbook 2008* (pp. 245–259). London: College Publications.

Schroeder-Heister, P. (2012a). The categorical and the hypothetical: a critique of some fundamental assumptions of standard semantics. *Synthese*, *187*(3), 925–942.

Schroeder-Heister, P. (2012b). Proof-theoretic semantics, self-contradiction, and the format of deductive reasoning. *Topoi*, *31*, 77–85.

Schroeder-Heister, P. (2016). Open problems in proof-theoretic semantics. In T. Piecha & P. Schroeder-Heister (Eds.), *Advances in proof-theoretic semantics* (pp. 253–283). Springer.

Schroeder-Heister, P. (2022). Proof-theoretic semantics. In E. N. Zalta & U. Nodelman (Eds.), *The Stanford Encyclopedia of Philosophy* (Winter 2022 ed.). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/spr2018/entries/proof-theoretic-semantics/

Schroeder-Heister, P., & Tranchini, L. (2017). Ekman's paradox. *Notre Dame Journal of Formal Logic*, *58*(4), 567–581.

Schroeder-Heister, P., & Tranchini, L. (2018). How to Ekman a Crabbé-Tennant. *Synthese*. https://doi.org/10.1007/s11229-018-02018-3

Smiley, T. (1996). Rejection. *Analysis*, *56*(1), 1–9.

Sørensen, M., & Urzyczyn, P. (2006). *Lectures on the Curry-Howard Isomorphism.* Amsterdam: Elsevier Science.

Spade, P., & Panaccio, C. (2019). William of Ockham. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Sring 2019 ed.). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/spr2019/entries/ockham/

Statman, R. (1983). Lambda-definable functionals and beta-eta-conversion. *Archiv für Mathematische Logik*, *23*, 21–26.

Steinberger, F. (2011). Why conclusions should remain single. *Journal of Philosophical Logic*, *40*, 333–355.

Straßburger, L. (2019). The problem of proof identity, and why computer scientists should care about Hilbert's 24th problem. *Philosophical Transactions of the Royal Society A*, *377*(2140). https://doi.org/10.1098/rsta.2018.0038

Sundholm, G. (1994). Proof-theoretical semantics and fregean identity criteria for propositions. *The Monist*, *77*(3), 294–314.

Tennant, N. (1978). *Natural logic.* Edinburgh: Edinburgh University Press.

Tennant, N. (1982). Proof and paradox. *Dialectica*, *36*, 265–296.

Tennant, N. (1995). On paradox without self-reference. *Analysis*, *55*(3), 199–207.

Tennant, N. (2021). Normalizability, cut eliminability and paradox. *Synthese*, *199*(Suppl 3), S597–S616. https://doi.org/10.1007/s11229-016-1119-8

Tranchini, L. (2015). Harmonising harmony. *The Review of Symbolic Logic*, *8*(3), 411–423. https://doi.org/10.1017/S1755020315000179

Tranchini, L. (2016). Proof-theoretic semantics, paradoxes and the distinction between sense and denotation. *Journal of Logic and Computation*, *26*(2), 495–512.

Tranchini, L. (2018). Stabilizing quantum disjunction. *Journal of Philosophical*

*Logic*, *47*, 1029–1047.

Tranchini, L. (2021). Proof-theoretic harmony: towards an intensional account. *Synthese*, *198*(Suppl 5), S1145–S1176. https://doi.org/10.1007/s11229-016-1200-3

Troelstra, A., & Schwichtenberg, H. (2000). *Basic proof theory* (2nd ed.). Cambridge: Cambridge University Press.

Urban, C. (2014). Revisiting Zucker's work on the correspondence between cut-elimination and normalisation. In L. Pereira, E. Haeusler, & V. de Paiva (Eds.), *Advances in natural deduction: A celebration of Dag Prawitz's work* (pp. 31–50). Dordrecht: Springer.

van Dalen, D. (2004). *Logic and structure* (4th ed.). Berlin/Heidelberg: Springer.

von Kutschera, F. (1968). Die Vollständigkeit des Operatorensystems $\{\neg, \vee, \supset\}$ für die Intuitionistische Aussagenlogik im Rahmen der Gentzensematik. *Archiv für mathematische Logik und Grundlagenforschung*, *11*, 3–16.

von Plato, J. (2000). A problem of normal form in natural deduction. *Mathematical Logic Quarterly*, *46*(1), 121–124.

von Plato, J., & Siders, A. (2012). Normal derivability in classical natural deduction. *The Review of Symbolic Logic*, *5*(2), 205–211.

Wansing, H. (1993). *The logic of information structures.* Berlin and Heidelberg: Springer.

Wansing, H. (2000). The idea of a proof-theoretic semantics and the meaning of the logical operations. *Studia Logica*, *64*, 3–20.

Wansing, H. (2006). Connectives stranger than tonk. *Journal of Philosophical Logic*, *35*, 653–660. https://doi.org/10.1007/s10992-006-9025-z

Wansing, H. (2010). Proofs, disproofs, and their duals. In V. Goranko, L. Beklemishev, & V. Sheht (Eds.), *Advances in modal logic* (Vol. 8, pp. 483–505). College Publications.

Wansing, H. (2016a). Falsification, natural deduction and bi-intuitionistic logic. *Journal of Logic and Computation*, *26*(1), 425–450.

Wansing, H. (2016b). Natural deduction for bi-connexive logic and a two-sorted typed lambda-calculus. *IFCoLog Journal of Logics and their Applications*, *3*(3), 413–439.

Wansing, H. (2016c). On split negation, strong negation, information, falsification, and verification. In K. Bimbó (Ed.), *J. Michael Dunn on information based logics. Outstanding contributions to logic* (Vol. 8, pp. 161–189). Springer.

Wansing, H. (2017). A more general general proof theory. *Journal of Applied Logic*, *25*, 23–46.

Wansing, H. (in press). A note on synonymy in proof-theoretic semantics. In T. Piecha & K. Wehmeier (Eds.), *Peter Schroeder-Heister on Proof-Theoretic Semantics. Outstanding contributions to logic.* Springer.

Widebäck, F. (2001). *Identity of proofs.* Stockholm: Almquist & Wiksell International.

Wittgenstein, L. (2006[1953]). *Werkausgabe Band 1: Tractatus logico-philosophicus - Tagebücher 1914-1916 - Philosophische Untersuchungen.* Frankfurt a. M.: Suhrkamp.

Wójcicki, R. (1979). Referential matrix semantics for propositional calculi. *Bulletin of the Section of Logic*, *8*(4), 170–176.

Zucker, J. (1974). The correspondence between cut-elimination and normalization. *Annals of Mathematical Logic*, *7*, 1–112.

Zucker, J., & Tragesser, R. (1978). The adequacy problem for inferential logic. *Journal of Philosophical Logic*, *7*, 501–516.

# Deutsche Zusammenfassung

Das übergeordnete Thema, in dem sich meine Dissertation bewegt, ist *beweistheoretische Semantik*. Dieser Ausdruck wurde von Schroeder-Heister auf einer Konferenz im Jahr 1991 geprägt, wodurch er eine bestimmte Fokussierung in dem übergeordneten Gebiet der *allgemeinen Beweistheorie* ausdrücken wollte (vgl. Schroeder-Heister, 2022). Diese wurde zu Beginn der 1970er Jahre von Vertretern wie Kreisel (1971), Prawitz (1971; 1973) und Martin-Löf (1975) stark vorangetrieben. Die Annahme einer allgemeinen Beweistheorie ist hierbei – in Opposition zu z. B. Hilberts Auffassung -, dass Beweise nicht nur ein ‚Mittel zum Zweck' o. ä. seien, sondern dass sie interessante Objekte an sich darstellten, die einer genaueren Untersuchung wert seien. Der Fokus liegt dann hierbei z. B. nicht mehr nur auf der Frage, was bewiesen werden kann, sondern eher darauf, *wie* etwas bewiesen werden kann, also auf den Beweisregeln eines Systems statt auf den Theoremen. Beweistheoretische Semantik kann nun in diesem Kontext als die Beschäftigung mit einer Semantik von Beweisen aufgefasst werden. Zum anderen ist beweistheoretische Semantik aber auch eine Semantik *durch* Beweise, in dem Sinne, dass die zentrale Frage sich mit der Bedeutung der logischen Konnektive beschäftigt und hierbei davon ausgegangen wird, dass diese durch die Regeln, die ihren Gebrauch in einem Beweissystem bestimmen, gegeben wird. Damit wendet sich beweistheoretische Semantik gegen die ‚traditionelle' Antwort auf diese Frage, welche die Bedeutung der Konnektive durch modelltheoretische Begriffe charakterisiert.

In diesem Sinne ist beweistheoretische Semantik damit in das Gebiet des *Inferentialismus* einzuordnen, eine philosophische Position, die man auf den Wittgenstein'schen Ausspruch in seinen philosophischen Untersuchungen zurückführen kann, dass „[d]ie Bedeutung eines Wortes [...] sein Gebrauch in der Sprache" (Wittgenstein, 2006[1953], p. 43) ist. Der Inferentialismus, der besonders von Brandom (vgl. 2000) vertreten und vorangebracht wurde, ist damit eine semantische Position, die davon ausgeht, dass die Bedeutung sprachlicher Ausdrücke durch die Art und Weise ihres Gebrauchs und ihrer Interaktion mit anderen Ausdrücken bestimmt wird. Der Vorteil einer solchen semantischen Position, und damit auch der Vorteil von beweistheoretischer Semantik gegenüber modelltheoretischer Semantik, wird meist damit begründet, dass sie ohne Annahmen über metaphysisch kontroverse Konzepte, wie Wahrheit, mögliche Welten, etc., auskommt. Eine solche ‚sparsamere' Metaphysik wird, nach dem Prinzip von *Ockhams Rasiermesser*, allgemeinhin in der Philosophie als vorteilhaft gesehen.

Ein Kernproblem, mit dem sich beweistheoretische Semantik befassen muss, ist, welche Beweisregeln überhaupt zulässig sind, um als bedeutungsgebend für logische Konnektive gelten zu können. Paradigmatisch wird hierbei zur Illustrierung meist das Konnektiv `tonk` herangezogen, welches in Priors (1960) äußerst einflussreichem

Aufsatz „The Runabout Inference-Ticket“ eingeführt wurde, mit dem Ziel einen solchen beweistheoretischen Ansatz in Bezug auf Semantik ad absurdum zu führen. Prior argumentiert, dass, wenn Beweisregeln tatsächlich alles wären, was wir für eine Semantik der Konnektive bräuchten, uns nichts davon abhalten würde, ein Konnektiv, wie `tonk`, für bedeutungsvoll zu halten, welches mit folgenden Regeln gebraucht werden darf: Von beliebiger Aussage $A$ darf auf $A$ `tonk` $B$ geschlossen werden und von $A$ `tonk` $B$ darf auf beliebige Aussage $B$ geschlossen werden. Damit sind die Regeln für `tonk` eine Mischung aus einer der Einführungsregeln für die Disjunktion sowie einer der Beseitigungsregeln der Konjunktion und insofern problematisch, als dass sie eine Ableitung von beliebigem $B$ aus beliebigem $A$ erlauben würden. Man könnte, mit anderen Worten, also alles aus allem ableiten und somit würde `tonk` die Konsequenzrelation trivialisieren. Mit diesem Einwand Priors hat beweistheoretische Semantik die nicht-triviale Aufgabe bekommen, bestimmte Merkmale für Regeln zu geben, welche, grob gesagt, möglichst die Konnektive einschließen, von denen wir intuitiv annehmen, dass sie akzeptabel sind (wie z. B. die Konjunktion, Implikation, etc.),[99] während die inakzeptablen Konnektive, wie `tonk`, ausgeschlossen werden sollten. Diese Merkmale werden allgemein unter dem Begriff der Harmonie zwischen Einführungs- und Beseitigungsregeln[100] eines Konnektivs zusammengefasst: Die Regeln eines Konnektivs müssen untereinander harmonisch sein, damit sie als bedeutungsgebend akzeptiert werden können. Was genau diese Harmonie dabei ausmacht, ist allerdings nicht unkontrovers in der Literatur. In der Arbeit skizziere ich einige besonders prominente Ansätze, ohne dabei ins Detail zu gehen (s. Kapitel 1.2). Nachdem ich nun einen Überblick über das allgemeine Thema, in dem meine Arbeit verortet werden kann, gegeben habe, werde ich im Folgenden die einzelnen Aufsätze, die meine kumulative Dissertation ausmachen, zusammenfassen.

In dem Aufsatz „What is the meaning of proofs? A Fregean distinction in proof-theoretic semantics” (Kapitel 2) ist das Ziel, eine Unterscheidung zwischen Sinn und Denotation von Beweisen zu geben, in Anlehnung an Freges Konzeption dazu im Falle von einzelnen Wörtern sowie Sätzen, wie er sie in seinem berühmten Aufsatz „Über Sinn und Bedeutung“ darlegt (s. Frege, 1892). Während über die Denotation von Beweisen in der Literatur zur allgemeinen Beweistheorie schon einiges geschrieben wurde, ist dies nämlich für eine davon abweichende Bedeutung, also eines Sinnes von Beweisen, nicht der Fall. Tranchini hat zu diesem Thema einen interessanten Aufsatz (s. Tranchini, 2016) verfasst mit der Motivation zwischen Beweisen zu unterscheiden, die zum einen sowohl Sinn als auch Denotation haben, zum

[99]Damit soll nicht gesagt werden, dass diese Konnektive vollkommen unkontrovers seien. Natürlich gibt es auch bei unseren „Standard-Konnektiven“ Diskussionen darüber, welche davon als primitiv im System angesehen werden sollten, welche Regeln tatsächlich am besten sind, um sie zu charakterisieren, etc.

[100]In einem System des natürlichen Schließens, welches das bevorzugte Format eines Beweissystems in beweistheoretischer Semantik ist. Es gibt aber auch wichtige Vertreter, die für die Überlegenheit von Sequenzenkalkül-Systemen argumentieren. S. Kapitel 1.3.

anderen zwar Sinn aber keine Denotation, was für Tranchini auf beweistheoretische Paradoxien zutreffen würde, und an dritter Stelle Beweise, die weder Sinn noch Denotation haben, was nämlich dann „Beweise“ wären, in denen ein Konnektiv, wie `tonk`, vorkommt. Der grundlegende Gedanke einer Frege'schen Unterscheidung von Sinn und Denotation von Beweisen ist, dass es verschiedene Möglichkeiten gibt, eine Ableitung desselben Beweises zu liefern. Ein Standardbeispiel hierfür wären zwei Ableitungen, von denen die eine in nicht-normaler Form und die andere in ihrer jeweiligen Normalform vorliegt. Als Normalform einer Ableitung versteht man eine Ableitung, welche keine redundanten Schrittfolgen in dem Sinne enthält, dass auf die Anwendung einer Einführungsregel eines Konnektivs direkt eine Anwendung der Beseitigungsregel für dasselbe Konnektiv folgt. Reduktionen sind hierbei gewisse Transformationen, welche uns für ein bestimmtes Konnektiv vorgeben, wie eine Ableitung, welche eine solche redundante Schrittfolge enthält, in eine Ableitung ohne diese Schrittfolge gebracht werden kann. Hierbei ist es üblich zu unterscheiden (s. z. B. Kreisel, 1971; Martin-Löf, 1975; Prawitz, 1971) zwischen einem *Beweis* als dem zugrundeliegenden Objekt (aufgefasst als mentale Entität im Sinne der intuitionistischen Tradition) und einer *Ableitung* als dessen jeweiliger sprachlicher Repräsentation. Da die Ableitung in Normalform die direkteste Art der Darstellung des Beweises ist, kann man argumentieren, dass diese die Denotation am besten wiedergibt. Wir können also sagen, dass in einem Fall, in dem wir zwei Ableitungen haben, von denen eine die Normalform der anderen ist, die Denotation der beiden Ableitungen die gleiche ist, während ihr Sinn unterschiedlich ist. Ein solches Konzept findet z. B. auch bei Girard (1989) Erwähnung, ist aber in der beweistheoretischen Standardliteratur ansonsten nicht weiter entwickelt worden. Tranchini (2016) stellt ausführlicher dar, wie eine solche Unterscheidung im Kontext der beweistheoretischen Semantik sinnvoll angewendet werden kann. Er argumentiert, dass eine Ableitung nur dann einen Sinn haben kann, wenn es der Fall ist, dass alle in ihr angewandten Regeln über Reduktionen verfügen (im Gegensatz zu z. B. den Regeln für `tonk`). Dies begründet er damit, dass die Reduktionen das sind, was eine Ableitung in ihre normale Form, d. h. ihre Denotation, überführt. Die Reduktionen sind also der Weg zur Denotation von Beweisen, was einer Frege'schen Konzeption von Sinn zu entsprechen scheint. In meinem Aufsatz baue ich auf Tranchinis Idee auf, indem ich das Vorhandensein von Reduktionen als notwendige Bedingung dafür, dass einer Ableitung Sinn zugesprochen werden kann, übernehme. Gleichzeitig entwickle ich dieses Konzept aber auch weiter, indem ich es erstens auf ein Beweissystem, welches mit Lambdatermen annotiert ist, übertrage und zweitens eine konkrete Definition von Denotation und Sinn einer Ableitung gebe. Während die Denotation einer Ableitung in solchen Systemen die normale Form des Lambdaterms ist, der die Konklusion der Ableitung annotiert (hier: der Endterm), argumentiere ich, dass der Sinn einer Ableitung als die Menge aller Lambdaterme, die in der Ableitung vorkommen, aufgefasst werden sollte. Dies wird dadurch motiviert, dass diese die in der Ableitung verwendeten Operationen der Ableitung widerspiegeln. Man kann an ihnen die Art und Weise, wie der Endterm aufgebaut wird, ablesen und damit zeigen sie den Weg, der zur Denotation führt. Dieses Vorgehen hat zwei Vorteile. Zum einen, dass wir dadurch nicht nur von Gleichheit sprechen können, wenn es um die Denotation geht, d. h. um die Identität der Beweise, sondern auch von Gleichheit des Sinnes, also die Frage der Synonymie von Beweisen. Zum anderen können wir durch die Verwendung der Lambdaterme Sinn und Denotation über verschiedene Arten von Beweissystemen hinweg vergleichen, z. B. zwischen Systemen des natürlichen Schließens und Sequenzenkalkül-Systemen. Zwei Beweise sind genau dann identisch, aber nicht synonym, wenn die Normalform des Endterms identisch ist, die Terme allerdings, die innerhalb der Ableitung auftauchen, nicht identisch sind. Dies kann z. B. in dem oben erwähnten Fall vorkommen, in dem wir zwei unterschiedliche Ableitungen haben, die aber dieselbe Normalform aufweisen. Es kann aber auch dadurch zustande kommen, dass es zwei Ableitungen mit demselben Endterm gibt, bei denen der Unterschied darin liegt, dass die Reihenfolge, in der die Regeln in der Ableitung angewendet werden, anders ist. Dadurch wird der – im Endeffekt identische - Endterm dann nämlich auf unterschiedliche Weise aufgebaut. Die Art und Weise, wie wir zum gleichen Ergebnis gelangen, ist eine andere und damit, so kann man in Frege'scher Denkweise argumentieren, ist die Art des Gegebenseins, also der Sinn, ein anderer. Zum anderen kann es den Fall geben, dass sowohl Denotation als auch Sinn gleich sind, also Synonymie vorliegt. Der Vorteil an dem hier vorgeschlagenen System ist dabei, dass dies nicht automatisch mit nur den Fällen zusammenfällt, in denen die sprachliche Repräsentation exakt die gleiche ist, sondern es, wie bei Frege auch, durchaus den Fall geben kann, dass unterschiedliche Zeichen nicht zu einem Unterschied im Sinne führen. Dies ist dann der Fall, wenn die Menge der Lambdaterme, die in der Ableitung verwendet werden, exakt gleich ist, aber die sprachliche Repräsentation nicht gleich ist, wie z. B., wenn die eine Ableitung im natürlichen Schließen vorliegt, während die andere in einer Sprache des Sequenzenkalküls gegeben ist.

Der Aufsatz „What are good reductions? Perspectives from proof-theoretic semantics and type theory", welcher Kapitel 3 ausmacht, schließt inhaltlich direkt an den vorhergehenden an. Ich gehe hierin der Frage nach, welche Reduktionen von Ableitungen als zulässig anerkannt werden sollten und welche nicht. Die Motivation dafür liegt besonders im Zusammenhang zwischen Reduktionen und dem Sinn von Ableitungen begründet, welchen ich im ersten Aufsatz herausgestellt habe. Wenn man die Auffassung vertritt, dass eine Ableitung nur dann Bedeutung haben kann, wenn die in ihr verwendeten Regeln Reduktionen haben, dann scheint es legitim zu sein, ein konkretes Konzept dessen, was überhaupt als Reduktion gelten darf und was nicht, zu fordern. Der Hintergrund ist, dass teilweise argumentiert wird, dass es nicht nur die ‚üblichen' Reduktionen gibt, welche an die Regeln für die jeweiligen Konnektive gebunden sind, sondern, dass alles, was in irgendeiner Weise einen redundanten Umweg aus einer Ableitung eliminiert, als (zulässige) Reduktion angesehen werden kann. Ein Beispiel für eine solche Reduktion wird in (Ekman, 1994, 1998) gegeben und die Konsequenzen davon für die beweistheoretische Konzeption von Paradoxien aufgezeigt, welche wären, dass diese Konzeption aufgegeben werden müsste. Schroeder-Heisters und Tranchinis (2017) Antwort darauf ist, dass es die zugrundeliegende Ekman-Reduktion ist, welche problematisch ist und nicht die Konzeption von Paradoxien, indem sie zeigen, dass die Akzeptanz der Ekman-Reduktion in einem System zur Folge hätte, dass ein auf Reduktionen basierendes Konzept von Identität zwischen Beweisen trivialisiert würde. In dem Fall, so zeigen sie, müssten nämlich alle Ableitungen derselben Formel identifiziert werden, was im üblichen Verständnis von Beweistheorie äußerst kontraintuitiv erscheint. Während ich Schroeder-Heisters und Tranchinis Schlussfolgerungen sehr wohl zustimme, versuche ich ihren Ansatz zu verallgemeinern, um ihn noch weiter anwendbar zu machen. Dafür zeige ich zum einen, warum die Ekman-Reduktion auch dann als nicht akzeptabel angesehen werden kann, wenn man ein auf Reduktionen basierendes Konzept von Identität zwischen Beweisen ablehnt, was nämlich einige einflussreiche Autoren machen (s. z. B. Tennant, 2021). Darauf aufbauend gebe ich dann zum anderen ein allgemeines Kriterium, an dem für jegliche Kandidaten, die als Reduktion scheinen können, geprüft werden kann, ob dieses gilt oder nicht. Dieses Kriterium nenne ich „schwache Subjektreduktion" und es gilt, wie auch in meinem ersten Aufsatz, für Terme des Lambdakalküls, mit welchen das Beweissystem annotiert werden kann. Es besagt, dass es für eine Ableitung des Terms $t$ vom Typ $A$, wobei $t$ auf Term $t'$ reduziert werden kann und Subjektreduktion nicht gilt, nicht der Fall sein darf, dass $t'$ von einem zufälligen Typ $B$ ist. Dies ist für die Ekman-Reduktion aber genau der Fall und ich zeige außerdem, dass es für eine Reduktion des Nonsens-Konnektivs `tonk` der Fall wäre, würden wir eine solche für Lambdaterme konstruieren. Aus philosophischen Gründen (und im natürlichen Schließen technischen Gründen) ist es aber Konsens, dass es keine Reduktion für `tonk` geben kann, weswegen jedes formale Kriterium, welches für die Zulässigkeit von Reduktionen gegeben wird, eine solche Reduktion auch ausschließen sollte, was mit meinem Kriterium möglich ist.

Kapitel 4 beinhaltet schließlich den Aufsatz „A cut-free sequent calculus for the bi-intuitionistic logic `2Int`". Ich stelle hier den von mir entworfenen Sequenzenkalkül `SC2Int` vor, welcher das Pendant zu Wansings (2016a; 2017) System des natürlichen Schließens, `N2Int`, für die Logik `2Int` darstellt. Das System ist im Stile des intuitionistischen Kalküls `G3ip` (s. Negri & von Plato, 2001) und ich beweise in dem Aufsatz die Zulässigkeit der strukturellen Regeln *Weakening*, *Contraction* und *Cut*. Letzteres ist eine besonders wichtige Eigenschaft von Sequenzenkalkül-

Beweissystemen. Durch den Beweis von Schnitteliminierung erhält man nämlich weitere wichtige Korollarien, wie die Konsistenz des Systems oder das Subformel-Theorem. Der Beweis der Schnitteliminierung geht dabei über Wansings (2017) Ergebnisse für `N2Int` hinaus, für welches er einen Beweis des Normalformtheorems erbringt. Ein Normalformtheorem für ein System des natürlichen Schließens besagt, dass, wenn eine Formel $A$ ableitbar ist in dem System, es auch eine Ableitung von $A$ gibt, welche in Normalform ist. Der Beweis von Schnitteliminierung korrespondiert allerdings im natürlichen Schließen zu einem Beweis von Normalisierung, welches eine stärkere Eigenschaft als das ist, was ein Normalformtheorem besagt. Bei einem Beweis von Normalisierung bzw. Schnitteliminierung wird eine Prozedur gegeben, mit der man tatsächlich in der Lage ist, jedwede Ableitung in Normalform zu bringen bzw. Anwendungen der Schnittregel aus ihr zu eliminieren.

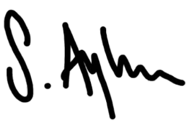
Der nächste Aufsatz, welcher Kapitel 5 bildet, trägt den Titel „Uniqueness of Logical Connectives in a Bilateralist Setting“. Die Frage der Einzigartigkeit ist die Frage, ob ein Konnektiv durch die Regeln, die seine Verwendung in Beweisen bestimmen, so charakterisiert ist, dass es höchstens ein Konnektiv gibt, das diese spezifische Rolle in Inferenzen spielt. Der übliche Weg, dies zu testen, ist ein ‘Nachahmer’-Konnektiv aufzustellen, welches von exakt denselben Regeln bestimmt wird, und zu zeigen, dass Formeln, die diese Konnektive enthalten, gegenseitig ableitbar sind. Es gibt schon einige wichtige Arbeiten zu dem Thema (besonders hervorzuheben ist hier Humberstone (2011; 2019; 2020b)), in welchen problematische Merkmale bestimmter Logiken aufgezeigt werden, die zum Scheitern von Einzigartigkeit einiger Konnektive in diesen Systemen führen, und außerdem auch Abwandlungen bzw. Verfeinerungen der Bedingungen, durch die wir die Eigenschaft der Einzigartigkeit sichern können. Ich untersuche dieses Phänomen nun im Kontext bilateraler Beweissysteme, genauer gesagt, in den Beweissystemen für die Logik `2Int`, *N2Int* und *SC2Int*, die insofern bilateral sind, als dass sie zwei Ableitbarkeitsrelationen aufweisen: eine für Beweisbarkeit und eine für duale Beweisbarkeit. Mein Ziel ist es, zu zeigen, dass die Probleme, die in einem bilateralen System auftreten, anders sind als die bisher entdeckten Probleme, die dazu führen, dass Einzigartigkeit in bestimmten Systemen nicht gewährleistet ist. Abschließend schlage ich eine Modifikation der bisherigen Charakterisierung von Einzigartigkeit vor, die es uns ermöglicht, diese Eigenschaft in bilateralen Systemen zu prüfen. Diese Modifikation besteht darin, dass in einem bilateralen System die gegenseitige Ableitbarkeit nicht nur für eine Ableitbarkeitsrelation gelten muss, damit gesichert ist, dass das Konnektiv einzigartig ist, sondern dass dies für beide Ableitbarkeitsrelationen gelten muss.

Der letzte Teil in Kapitel 6 ist als Aufsatz geplant, aber noch nicht eingereicht und trägt den vorläufigen Titel „Meaning and identity of proofs in a bilateralist setting: A two-sorted typed $\lambda$-calculus for proofs and refutations“. In gewisser Weise laufen alle Stränge der vorherigen Aufsätze in diesem Aufsatz zusammen, da es zum einen wieder um ein bilaterales System geht, zum anderen um Sinn und Denotation in diesem System und dafür natürlich auch wieder die Reduktionen von Bedeutung sind. Konkret wird in diesem Fall eine Erweiterung des Lambdakalküls entworfen, welche geeignet ist, sowohl Beweise als auch Widerlegungen für die Logik `2Int` zu annotieren. Für dieses System, welches ich $\lambda^{2Int}$ nenne, werden zunächst einige Eigenschaften bewiesen, die standardmäßig als wichtig für Lambdakalküle angesehen werden, wie z. B. Subjektreduktion. Darauffolgend stelle ich ein *Dualitätstheorem* auf und beweise es, welches besagt, dass für jeden Beweis (bzw. für jede Widerlegung) einer Formel $A$, welche/r in dem System gegeben werden kann, auch eine Widerlegung (bzw. ein Beweis) der sogenannten *dualen Formel* von $A$ gegeben werden kann. Im Anschluss argumentiere ich dann dafür, dass wir die Terme und ihre jeweiligen dualen Terme, die über das Dualitätstheorem generiert werden, identifizieren sollten. Das würde dazu führen, dass wir für jeden Beweis (bzw. für jede Widerlegung) sagen können, dass es eine Widerlegung (bzw. einen Beweis) gibt, sodass gilt, dass sie dieselbe Denotation haben. Das jeweilige Konstrukt, was dem Beweis/der Widerlegung zugrunde liegt, ist dasselbe, der Sinn ist allerdings voneinander verschieden, da die Art und Weise, wie das Konstrukt uns gegeben wird, unterschiedlich ist. Damit hätten wir eine, aus bilateraler Sicht, sehr wünschenswerte Gleichheit zwischen Beweisen und Widerlegungen, da weder das eine noch das andere als grundlegender angesehen wird, sondern beide als gleichwertige (da gleiche) Objekte gesehen werden.

# Eidesstattliche Erklärung

| | |
|---|---|
| Name: | Sara Ayhan |
| Geburtsdatum: | 11.07.1990 |
| Geburtsort: | Datteln |

Hiermit versichere ich an Eides statt, dass ich die eingereichte Dissertation selbstständig und ohne unzulässige fremde Hilfe verfasst, andere als die in ihr angegebene Literatur nicht benutzt und dass ich alle ganz oder annähernd übernommenen Textstellen sowie verwendete Grafiken, Tabellen und Auswertungsprogramme kenntlich gemacht habe. Außerdem versichere ich, dass die vorgelegte elektronische mit der schriftlichen Version der Dissertation übereinstimmt und die Abhandlung in dieser oder ähnlicher Form noch nicht anderweitig als Promotionsleistung vorgelegt und bewertet wurde.



Sara Ayhan

Melbourne, den 17.03.2023

Sara Ayhan

*Curriculum Vitae*

✉ *sara.ayhan@rub.de*

## Education

2018 - 2023 **PhD in Philosophy,** Ruhr University Bochum
Thesis: *Meaning and identity of proofs in (bilateralist) proof-theoretic semantics*,
Grade: *summa cum laude*
Supervisors: Prof. Dr. Heinrich Wansing, Prof. Dr. Greg Restall

2015 - 2018 **M.A. Philosophy (single major),** Ruhr University Bochum, (2-year program)
Thesis: *Proof-theoretic semantics and paradoxes*, Award for best Master's thesis of the faculty
Supervisor: Prof. Dr. Heinrich Wansing

2010 - 2015 **First State Exam** Philosophy, English & History, University of Siegen
(5-year program)
Thesis: *Donald Davidson's Conception of Truth*
Supervisor: Prof. Dr. Richard Schantz

## Employment

2018 - 2023 **Research Assistant,** *Institute for Philosophy I - Logic and Epistemology*, Ruhr University Bochum
(Prof. Dr. Heinrich Wansing)

2013 - 2015 **Student Assistant,** Chair of History of Philosophy & Philosophy of Language, University of Siegen
(Prof. Dr. Richard Schantz)

## Publications

Ayhan, S. (forthcoming). What are acceptable reductions? Perspectives from proof-theoretic semantics and type theory. *Australasian Journal of Logic.*

Ayhan, S. & Wansing, H. (forthcoming). On synonymy in proof-theoretic semantics. The case of 2Int. *Bulletin of the Section of Logic.*
https://czasopisma.uni.lodz.pl/bulletin/article/view/14910/18176.

Ayhan, S. (2021). What is the Meaning of Proofs? A Fregean Distinction in Proof-Theoretic Semantics. *Journal of Philosophical Logic* 50, 571–591.
http://link.springer.com/article/10.1007/s10992-020-09577-2.

Ayhan, S. (2021). Uniqueness of Logical Connectives in a Bilateralist Setting. In *The Logica Yearbook 2020* ed. by Igor Sedlár & Martin Blicha, pp. 1–16, College Publications.

Ayhan, S. (2020). A cut-free sequent calculus for the bi-intuitionistic logic `2Int`. Unpublished Manuscript, https://arxiv.org/abs/2009.14787.