Journal of Logic, Language and Information (2025) 34:341–369
https://doi.org/10.1007/s10849-025-09437-9

Exception and typicality, logically framed

Mario Piazza[1] · **Matteo Tesi**[1,2]

Accepted: 16 May 2025 / Published online: 12 June 2025

Abstract
This paper presents a novel proof-theoretic approach to a logic specifically designed to handle exceptions and typicality. Our method extends the classical first-order sequent calculus by incorporating a specialized framework to manage both *negative* extra-logical information (explicit exceptions) and *positive* information (background assumptions). We prove that the resulting sequent calculi satisfy the cut-elimination theorem, thereby ensuring strong analytical properties. Furthermore, we show how this framework effectively models traditional reasoning patterns involving conflicting information and typicality. Finally, we establish a natural correspondence between our approach and the Kraus-Lehmann-Magidor postulates, further grounding our work within established theoretical foundations.

Keywords Proof theory · Extra-logical axioms · Exceptions · Cut elimination

1 Introduction

The notions of exception and typicality are complementary to each other and are omnipresent in scientific theorizing as well as everyday reasoning. From a logical perspective, they have primarily been explored in the context of non-monotonic logics, where exceptions are generally identified *implicitly* through default rules (Reiter, 1980) or axioms in the knowledge base (Giordano et al., 2013, 2020).[1]. In this paper, we propose a novel approach to the issue by incorporating an explicit representation of exception and typicality within sequent-based systems in a first-order language. Indeed, our primary focus is on the representation of the inferential processes that delineate the reasoning of a rational agent when they are presented with finitely many

[1] For a proof-theoretic approach to non-monotonic logics dealing directly with default rules via specific extra-logical rules, see Piazza and Sabatini (2025).

✉ Matteo Tesi
matteo.tesi@tuwien.ac.at; matteo.tesi@sns.it

Mario Piazza
mario.piazza@sns.it

1 Scuola Normale Superiore di Pisa, Pisa Pi, Italy

2 TU Wien, Technische Universitat Wien, Vienna, Austria

pieces of information that introduce exceptions to their prior assumptions. Such exceptions are interpreted as Boolean combinations of atomic propositions, which conflict with the information conveyed by the premise of an inference.

In particular, the distinctive feature of the sequent calculi we are going to introduce is their capability of accommodating both *negative* and *positive* information. By positive information, we mean the information delivered by extra-logical axioms, which are designed to represent *assumptions* in the background knowledge of the agent, along the lines of Makinson's pivotal assumption logic (Makinson, 2005; Piazza & Pulcini, 2016). The idea of integrating sequent calculi with extra-logical axioms dates back to Gentzen's work, and has been pursued in different ways since then Gentzen (1935); Negri and von Plato (1998); Ciabattoni et al. (2008). We exploit in particular the recent technique of conversion of extra-logical axioms into initial sequents discussed in Piazza and Tesi (2024) by decomposing formulas into initial sequents and checking their closure under cut.

By negative information, we mean the kind of information – attached to any sequent $\Gamma \vdash \Delta$ of the system – which *undermines* the derivability of Δ from Γ. Such extra-logical information represents known exceptions and the (finite) sets (of sets) of formulas encoding them are called *control sets* (D'Agostino et al., 2014; Piazza & Pulcini, 2017). As an illustration, let's consider the phenomenon of concurrent enzyme inhibition in molecular biology. The (pseudo-)sequent below decorated with the control set $\{\{I\}\}$:

$$\Gamma, E, S \vdash_{\{\{I\}\}} E{\star}S$$

reads: the enzyme E binds with the substratum S (so as to form the compound $E{\star}S$), *provided that* the inhibitor I is not present in the biochemical environment Γ at the moment of the interaction, i.e. $\{I\} \nsubseteq \Gamma$ (Boniolo et al., 2015, 2021).

The constraints imposed by the control sets are specified at the start by the design of the calculus, but they can be modified throughout the derivation making it more difficult to satisfy the criteria of soundness. In our system, soundness is rephrased not as set inclusion, but rather as underivability of a control set from the assumptions of a sequent. The core idea is that a certain conclusion is rejected whenever the information contained in the antecedent leads to a proof of formulas contained in a control set.

The notions of positive and negative information are thus strongly linked with derivability in a formal system. Positive information, conveyed by extra-logical initial sequents, enhances the system's deductive power, thereby increasing the size of the set of derivable sequents. Conversely, negative information, associated with control sets, restricts the set of derivable sequents in the system by imposing stricter soundness requirements. As is well-known, working in a first-order environment precludes the definition of a refutation calculus to check soundness due to the undecidability of first-order logic. To address this challenge, we reduce the existential (universal) quantifier to finite disjunctions (conjunctions), corresponding to suitable instances of the quantified formulas. This reduction allows the soundness check to be performed using a refutation system for classical propositional logic.

Dealing with extra-logical information while maintaining the analyticity of a given system is a well-known challenge in proof-theory. From both a conceptual and technical perspective, analyticity in the form of the subformula property is desirable. In an analytic system, proofs only contain information that is already present in the conclusion, thus avoiding the use of lemmas. Moreover, analyticity offers various technical benefits, such as the ability to use the calculus to explore the properties of the underlying logic. Indeed, analyticity is key to maintaining tight control over the structure of derivations within the calculus. First, constructing proofs in analytic systems (unlike in axiomatic calculi) is a relatively straightforward process, as it involves applying the rules in a backward manner. Second, analyticity enables the identification of the specific classes of rules applicable to the endsequent, which, in some cases, allows for the determination of the underivability of certain statements. In this context, we leverage the analyticity of our calculi to provide a logical framework for capturing the intransitivity of the notion of typicality (see Subsection 5.2) Achieving analyticity in a first-order setting, which includes quantified expressions, is even more difficult. However, this is the natural framework in which to develop a system that handles exceptions.

In the present paper, we introduce and discuss two different systems: $\mathbf{LEX}^{\mathcal{S}}_{\mathcal{E}}$ and $\mathbf{LET}^{\mathcal{S}}_{\mathcal{E}}$. The main difference between the two systems lies in the treatment of initial sequents. $\mathbf{LEX}^{\mathcal{S}}_{\mathcal{E}}$ allows for purely tautological initial sequents to be equipped with control sets. This feature leads to a (possible) restriction of instances of classically valid theorems, rendering the resulting calculus inherently *subclassical*. In contrast, in system $\mathbf{LET}^{\mathcal{S}}_{\mathcal{E}}$, control sets are permitted only in extra-logical initial sequents. Consequently, prohibitions come into play only when a rational agent handles information that extends beyond pure logic. As a result, the resulting calculus is *supraclassical*, as it preserves the derivability of classical theorems.

Our systems can be considered proof-theoretically well behaved since they allow for a full-fledged cut-elimination theorem. Cut can be eliminated whenever the conclusion is sound, and interestingly, when it is not, the elimination fails. Cut serves as the natural expression of transitivity, which is the central ingredient of taxonomical reasoning, as noted by Girard (1992) who states that "as far as deduction is concerned, the taxonomist is doing hardly more than combining inclusions by transitivity".

To summarize, our proof-theoretic approach to exception and typicality has three main advantages. Firstly, it doesn't require the use of external logical tools like modal operators. Secondly, it preserves the desirable structural properties of classical logic. Specifically, the sequent calculi we use are based on the language and rules of first-order classical logic, but with additional axioms and control sets that expand or restrict the derivability relation. Thirdly, this structural continuity allows us to establish a connection between classical and non-classical reasoning, similar to Makinson's proposal (Makinson, 2003). However, his approach is inherently semantic and relies on the notion of consistency, which can be treated proof-theoretically in classical propositional logic but not in first-order logic, which is undecidable.

As we said, the concepts of exception and typicality are closely related to non-monotonicity and the traditional approach in non-monotonic logic involves dealing with defeasible reasoning. However, there are significant connections between our proof-theoretic approach and non-monotonic logic that warrant further exploration. In particular, we shall show how to rephrase the Kraus-Lehmann-Magidor postulates into the sequent calculi with control sets. The admissibility of the KLM postulates is shown exploiting the admissibility of a restricted version of the cut rule. Interestingly, while some of the postulates hold only when assuming the soundness of the conclusion, others-such as *cautious monotonicity*-are valid under more general circumstances.

This article is organized as follows. In Section 2, we set the stage by introducing in a first-order setting the notion of control set as well as an array of auxiliary notions serving our machinery. Moreover, we briefly discus a refutation system for classical propositional logic (based on hypersequents) in relation to our approach. In Section 3, we first describe a controlled sequent calculus for first-order classical logic, called $\mathbf{LEX}^{\mathcal{S}}$, and then we turn to its expansion $\mathbf{LEX}^{\mathcal{S}}_{\mathcal{E}}$ obtained by adding extra-logical initial sequents (closed under cut) with a (possibly empty) control set attached to them. Section 4 addresses the proof of cut-elimination theorem for $\mathbf{LEX}^{\mathcal{S}}_{\mathcal{E}}$ which requires some delicacy due to the intertwining of a number of factors. Section 5 shows how a logical description of typicality statements can be afforded by designing a further sequent system ($\mathbf{LET}^{\mathcal{S}}_{\mathcal{E}}$), sharing the same rules as $\mathbf{LEX}^{\mathcal{S}}_{\mathcal{E}}$ but with different initial sequents. Section 6 is devoted to make explicit the connection between the Kraus-Lehmann-Magidor postulates and our approach. In Section 7, we draw conclusions and sketch some ideas for future work to investigate.

2 Preliminaries

2.1 Control sets

The language of first-order logic $\mathcal{F}$ contains a denumerable set of variables, a set of constants, a set of functions and a set of n-ary predicates. Terms are defined recursively as usual, the set of terms is denoted by TER; we define closed terms to be terms without variables. The language contains the logical connectives $\neg, \wedge, \vee, \rightarrow$ and the quantifiers $\forall, \exists$; formulas are defined as usual. We will use P, Q, R,... to denote atomic formulas. Boldface capital letters **S**,**T**,... stand for control sets. The operation of substitution of a variable x with a term t in a formula A (or a multiset) is denoted as $A[x/t]$.

As usual, we take multisets to be collections of formulas in which the multiplicity of their occurrences counts and the order does not. With a slight abuse of notation we shall use capital Greek letters $\Gamma, \Delta, \Theta, \ldots$ to denote collections of formulas which, depending on the context, will denote multisets or sets of formulas. We will make explicit the difference whenever needed in order to avoid confusion.

Definition 1 A *control set* **S** is a finite set of finite sets of quantifier-free first-order formulas in the $\{\wedge, \vee, \neg\}$-fragment of the language satisfying the following closure conditions:

(1) If $\{\Gamma, A \wedge B\} \in \mathbf{S}$, then $\{\Gamma, A, B\} \in \mathbf{S}$
(2) If $\{\Gamma, A \vee B\} \in \mathbf{S}$, then $\{\Gamma, A\} \in \mathbf{S}$ and $\{\Gamma, B\} \in \mathbf{S}$
(3) If $\{\Gamma, \neg(A \vee B)\} \in \mathbf{S}$, then $\{\Gamma, \neg A, \neg B\} \in \mathbf{S}$
(4) If $\{\Gamma, \neg(A \wedge B)\} \in \mathbf{S}$, then $\{\Gamma, \neg A\} \in \mathbf{S}$ and $\{\Gamma, \neg B\} \in \mathbf{S}$

Given a set X of propositional formulas in the language $\{\wedge, \vee, \neg\}$, the control set generated by X is the set of formulas obtained by closing the set X under the conditions (1) - (4).

The conditions mentioned above reflect the proper task performed by control sets, which is to express negative information and exceptions. Conjunctive statements express exceptions which form a pair, whereas disjunctive formulas encode a form of exclusive prohibition. Although we are working in a first-order setting, we do not include quantified formulas in control sets. The motivation for this decision is twofold. First, control sets that contain quantifier-free formulas are expressive enough to provide an adequate representation of inferential patterns concerning exceptions. Second, the inclusion of quantified formulas poses an obstacle for formulating a decidable soundness criterion.

However, when looking for a way to regiment the derivability of sequents in a suitably formulated calculus, we immediately encounter an obstacle. Specifically, while validity can be easily checked w.r.t. formulas in control sets, it is impossible to do so for arbitrary sequents in a first-order language, due to the undecidability of first-order logic. Therefore, it becomes necessary to introduce an auxiliary notion to build a bridge between control sets and sequents. In particular, we define the *relativization* operator $\mathfrak{R}$, which replaces quantifiers with finite disjunctions and conjunctions.

Definition 2 Given a multiset of formulas Γ, the *relativization* of Γ with respect to the control set $\mathbf{S}$, notation $\mathfrak{R}(\Gamma, \mathbf{S})$, is obtained by replacing each occurrence of the universal quantifier $\forall x A$ in Γ with $A(t_1) \wedge ... \wedge A(t_n)$ and of the existential quantifier $\exists x B$ with $B(t_1) \vee ... \vee B(t_n)$, where t_i are all the terms in $\mathbf{S}$.

In other words, the relativization operator is commissioned to eliminate the quantifiers by substituting them with finite conjunctions and disjunctions instantiated with the terms occurring in the control sets.

Definition 3 (**Compatibility**) Given a multiset Γ of formulas and a control set $\mathbf{S}$, Γ is *compatible* with $\mathbf{S}$ (in symbols $\Gamma \,||\, \mathbf{S}$) if and only if for every set of formulas Σ in $\mathbf{S}$ it is not the case that $\bigwedge \mathfrak{R}(\Gamma, \mathbf{S}) \rightarrow \bigwedge \Sigma$ is valid.

Compatibility expresses the idea that the formulas occurring in the multiset Γ do not conflict with the information stored in the control set $\mathbf{S}$. Notice that the notion of compatibility is indeed decidable, because it reduces to checking the (in)validity of a statement not involving quantifiers. We shall return on the notion of compatibility in the next section, after the introduction of the sequent calculus. We now state some properties of control sets.

Proposition 2.1 *The following statements hold:*

(1) $\Gamma, A \wedge B \,||\, \mathbf{S}$ *if and only if* $\Gamma, A, B \,||\, \mathbf{S}$.
(2) $\Gamma, A \vee B \,||\, \mathbf{S}$ *if and only if* $\Gamma, A \,||\, \mathbf{S}$ *or* $\Gamma, B \,||\, \mathbf{S}$.
(3) $\Gamma, \forall x A \,||\, \mathbf{S}$ *if and only if* $\Gamma, A[x/t_1], ..., A[x/t_n] \,||\, \mathbf{S}$ *where* $\{t_1, ..., t_n\}$ *are the terms occurring in* $\mathbf{S}$.
(4) $\Gamma, \exists x A \,||\, \mathbf{S}$ *if and only if* $\Gamma, A[x/t_i] \,||\, \mathbf{S}$ *for some* t_i *occurring in* $\mathbf{S}$.

Proof Item (1) is immediate by well-known equivalences which hold in classical propositional logic. With respect to item (2), the direction from left to right follows from the observation that if $\bigwedge \mathfrak{R}(\Gamma, A \vee B, \mathbf{S}) \rightarrow \bigwedge \Sigma$ is not valid (for $\Sigma \in \mathbf{S}$), then $\bigwedge \mathfrak{R}(\Gamma, A, \mathbf{S}) \rightarrow \bigwedge \Sigma$ is not valid or $\bigwedge \mathfrak{R}(\Gamma, B, \mathbf{S}) \rightarrow \bigwedge \Sigma$ is not valid. For the direction from right to left, we argue by contraposition, i.e. we assume that $\Gamma, A \vee B$ is not compatible with $\mathbf{S}$. Hence the formula $\bigwedge \mathfrak{R}(\Gamma, A \vee B, \mathbf{S}) \rightarrow \bigwedge \Sigma$ is valid for some $\Sigma \in \mathbf{S}$. The conclusion follows observing that $\bigwedge \mathfrak{R}(\Gamma, A, \mathbf{S}) \rightarrow \bigwedge \mathfrak{R}(\Gamma, A \vee B, \mathbf{S})$ and $\bigwedge \mathfrak{R}(\Gamma, B, \mathbf{S}) \rightarrow \bigwedge \mathfrak{R}(\Gamma, A \vee B, \mathbf{S})$ are both valid. As regards (3) and (4), they are reduced to the former cases due to the new definition of compatibility. We limit ourselves to discussing claim (3). In particular, it suffices to observe that $\mathfrak{R}(\Gamma, \forall x A, \mathbf{S}) = \mathfrak{R}(\Gamma, A[x/t_1] \wedge ... \wedge A[x/t_n], \mathbf{S})$ and by iterated applications of item (1) we get the desired conclusion. The case of (4) is dual. □

Proposition 2.2 *The following statements hold:*

(1) If $\Gamma \subseteq \Delta$ *and* $\Delta \,||\, \mathbf{S}$*, then* $\Gamma \,||\, \mathbf{S}$
(2) If $\mathbf{S} \subseteq \mathbf{T}$ *and* $\Gamma \,||\, \mathbf{T}$*, then* $\Gamma \,||\, \mathbf{S}$.

Proof We discuss only the claim (1), the other being analogous. We assume that $\Gamma \subseteq \Delta$ and $\Delta \,||\, \mathbf{S}$. The result immediately follows observing that $\mathfrak{R}(\Gamma, \mathbf{S}) \subseteq \mathfrak{R}(\Delta, \mathbf{S})$. □

2.2 Refutation systems

To perform invalidity checks and verify the compatibility of multisets of formulas and control sets, we present to the readers an automated method. Over the past few decades, refutation systems for classical propositional logic have been introduced in various guises. We opt for a refutation system (in Figure 1) based on hypersequents along the lines of systems discussed in Piazza et al. (2023). We recall some basic notions concerning hypersequents for refutation. We denote hypersequents with the capital Latin letters $\mathcal{G}, \mathcal{H}, \mathcal{I}, \ldots$ Following Avron (1991, 1996), a hypersequent is a list of ordinary (one-sided) sequents separated by the bar symbol '$\,|\,$', i.e., $\mathcal{G} = \Gamma_1 \dashv \Delta_1 \,|\, \Gamma_2 \dashv \Delta_2 \,|\, \cdots \,|\, \Gamma_n \dashv \Delta_n$. Hypersequents are considered up to the order of their components, that is $\mathcal{G} \,|\, \mathcal{H} \,|\, \mathcal{I} = \mathcal{G} \,|\, \mathcal{I} \,|\, \mathcal{H}$. We write $\Gamma \dashv \Delta \in \mathcal{G}$ to mean that the sequent $\Gamma \dashv \Delta$ is one of those displayed by $\mathcal{G}$. In case a certain hypersequent consists of a unique component $\Gamma \dashv \Delta$, we simply indicate it with $\Gamma \dashv \Delta$ instead of employing the longer expression $\varnothing \,|\, \Gamma \dashv \Delta$. The intended interpretation of a hypersequent $\Gamma_1 \dashv \Delta_1 \,|\, \Gamma_2 \dashv \Delta_2 \,|\, \cdots \,|\, \Gamma_n \dashv \Delta_n$ in the calculus HCPL is:

$$\bigwedge \Gamma_i \rightarrow \bigvee \Delta_i \ \textit{is not valid for some}\ i \in \{1, ..., n\}.$$

Therefore, the hypersequents admit a metalogical interpretation in terms of disjunctions.[2]

[2] Observe that the use of a hypersequent calculus for refutation comes from the definition of a hypersequent calculus for classical propositional logic where derivations have a linear structure and the reading of hypersequents is instead conjunctive.

AXIOM	
$\dfrac{}{\mathcal{G}\,|\,\Gamma \dashv \Delta}\,ax$	$P \notin \Gamma \cap \Delta$ for every P atomic
LOGICAL RULES	
$\dfrac{\mathcal{G}\,|\,\Gamma \dashv \Delta, A, B}{\mathcal{G}\,|\,\Gamma \dashv \Delta, A \vee B}\,\dashv\vee$	$\dfrac{\mathcal{G}\,|\,\Gamma \dashv \Delta, A\,|\,\Gamma \dashv \Delta, B}{\mathcal{G}\,|\,\Gamma \dashv \Delta, A \wedge B}\,\dashv\wedge$
$\dfrac{\mathcal{G}\,|\,A, B, \Gamma \dashv \Delta}{\mathcal{G}\,|\,A \wedge B, \Gamma \dashv \Delta}\,\wedge\dashv$	$\dfrac{\mathcal{G}\,|\,A, \Gamma \dashv \Delta\,|\,B, \Gamma \dashv \Delta}{\mathcal{G}\,|\,A \vee B, \Gamma \dashv \Delta}\,\vee\dashv$
$\dfrac{\mathcal{G}\,|\,\Gamma \dashv \Delta, A}{\mathcal{G}\,|\,\neg A, \Gamma \dashv \Delta}\,\neg\dashv$	$\dfrac{\mathcal{G}\,|\,A, \Gamma \dashv \Delta}{\mathcal{G}\,|\,\Gamma \dashv \Delta, \neg A}\,\dashv\neg$

Fig. 1 The refutation calculus HCPL for classical propositional logic

Derivations in the hypersequent system have a linear structure (by the design of the rules) and the calculus is sound and complete with respect to the invalid formulas of propositional classical logic in the sense specified by the next theorem.

Theorem 2.3 *For every hypersequent* $\mathcal{G} = \Gamma_1 \dashv \Delta_1 \mid \ldots \mid \Gamma_n \dashv \Delta_n$*:*

$$\mathcal{G} \text{ is derivable if and only if } \Gamma_i \dashv \Delta_i \text{ is not valid for some } i \in \{1, \ldots, n\}.$$

Proof The left-to-right direction is established by induction on the height of the derivation. Initial hypersequents are clearly not valid. We discuss the case of the rule $\dashv \wedge$. Since the premise $\Gamma \dashv A, \Delta \mid \Gamma \dashv B, \Delta$ (omitting hypersequent contexts for the sake of conciseness) is derivable, by induction hypothesis $\Gamma \dashv A, \Delta$ or $\Gamma \dashv B, \Delta$ is not valid. In both cases, we get that $\Gamma \dashv A \wedge B, \Delta$ is not valid.

The right-to-left direction is proved by contradiction. We assume that $\Gamma_i \dashv \Delta_i$ is not valid for some $i \in \{1, \ldots, n\}$. It can be easily shown that - for every rule of the calculus - whenever the conclusion is not valid, so is the premise. If we assume towards a contradiction that $\mathcal{G}$ is not derivable in the refutation calculus HCPL, then its initial hypersequent contains all tautological initial components of the shape $\Pi_j, P \dashv P, \Sigma_j$ for some atomic formula P. As a consequence, we would get that such initial sequent is not valid, which is a contradiction. □

This implies that we can use the refutation calculus for hypersequents to check the compatibility relation. In particular, the following proposition can be easily established.

Proposition 2.4 $\Gamma \,||\, \mathbf{S}$ *if and only if* $\Gamma \dashv \Lambda$ *is derivable in* HCPL *for any* Λ *in* $\mathbf{S}$.

Proof Immediate by the soundness and completeness of the calculus HCPL. □

3 The controlled extra-logical calculus

We start by introducing a controlled sequent calculus $\mathbf{LEX}^{\mathcal{S}}$ for first-order classical logic. $\mathcal{S}$ is a function which assigns a control set to every atomic formula of the language.

3.1 The controlled first-order system

Definition 4 A *controlled sequent* is a syntactic object of the form:

$$\Gamma \vdash_{\mathbf{S}} \Delta$$

where Γ, Δ are finite multisets of formulas and $\mathbf{S}$ is a control set.

The intended reading of a controlled sequent $\Gamma \vdash_{\mathbf{S}} \Delta$ is: from the assumptions contained in Γ, infer a conclusion in Δ, provided that the assumptions in Γ do not entail any information encoded in $\mathbf{S}$.

Remark 1 In previous studies on controlled sequent calculi, sequents were equipped with an additional syntactic device known as a "repository" (or "background sets" in Millson and Straßer (2019)). This was useful in the proof of cut-elimination and conceptually served as a storage unit Piazza and Pulcini (2017); Millson and Straßer (2019). Specifically, controlled sequents were expressed in the following form:

$$\Sigma \mid \Gamma \vdash_{\mathbf{S}} \Delta$$

Overall, repositories transform the *global* notion of a formula's position in a derivation into a *local* one. Removing repositories is not just a conceptual issue, but also technically relevant in first-order logic.

One can see that a first-order system with repositories would not be complete with respect to classical logic, as shown by the following simple derivation:

$$\dfrac{\dfrac{\dfrac{}{\cdot \mid P(x) \vdash P(x)}\text{ax}}{P(x) \mid\ \vdash P(x) \to P(x)}\text{R}\to}{P(x) \mid\ \vdash \forall x (P(x) \to P(x))}\text{R}\forall$$

The end-sequent is clearly derivable if we remove the repository. However, the presence of the repository prevents the application of the R$\forall$ rule as the variable x still appears in the conclusion of the rule (assuming that the freshness condition applies to the whole syntactic structure).

Therefore, we propose a new strategy to eliminate repositories from the controlled calculus. Specifically, we leverage the concept of polarity of a formula in a sequent (as described in Girard (1991); Troelstra and Schwichtenberg (2000)).

Definition 5 Given a formula A, we recursively define positive and negative contexts of A, respectively $\mathcal{P}(A)$ and $\mathcal{N}(A)$ as follows:

- $A \in \mathcal{P}(A)$.

- If $B \wedge C, B \vee C, \forall x B, \exists x B \in \mathcal{P}(A)$, then $B, C \in \mathcal{P}(A)$.
- If $B \wedge C, B \vee C, \forall x B, \exists x B \in \mathcal{N}(A)$, then $B, C \in \mathcal{N}(A)$.
- If $B \rightarrow C \in \mathcal{P}(A)$, then $B \in \mathcal{N}(A)$ and $C \in \mathcal{P}(A)$.
- If $B \rightarrow C \in \mathcal{N}(A)$, then $B \in \mathcal{P}(A)$ and $C \in \mathcal{N}(A)$.
- If $\neg B \in \mathcal{P}(A)$, then $B \in \mathcal{N}(A)$.
- If $\neg B \in \mathcal{N}(A)$, then $B \in \mathcal{P}(A)$.

The positive and negative contexts of a sequent $\Gamma \vdash_{\mathbf{S}} \Delta$ are the positive and negative contexts of $\bigwedge \Gamma \rightarrow \bigvee \Delta$.

Intuitively, a sequent is considered sound when its antecedent aligns with the attached control set. Nevertheless, formulas like implications and negations may shift positions throughout a derivation, necessitating careful tracking. By applying the notions of negative and positive polarity, repositories become unnecessary, as polarities automatically monitor the positioning of formulas in the derivation.

To this aim, we use the notation $\mathfrak{R}(\Gamma \vdash \Delta, \mathbf{S})$, which denotes the sequent obtained by replacing every formula of the form $\exists x A$ and $\forall x A$ with the formulas $A[x/t_1] \vee ... \vee A[x/t_n]$ and $A[x/t_1] \wedge ... \wedge A[x/t_n]$, where $t_1, ..., t_n$ are the terms found in $\mathbf{S}$.

For a controlled sequent $\Gamma \vdash_{\mathbf{S}} \Delta$, we denote $\mathsf{Neg}(\Gamma \vdash \Delta)$ and $\mathsf{Pos}(\Gamma \vdash \Delta)$ as the set of formulas occurring in the negative and positive contexts, respectively, of $\mathfrak{R}(\Gamma \vdash \Delta, \mathbf{S})$. The set of formulas required to check soundness is:

$$\mathsf{Asm}(\Gamma \vdash_{\mathbf{S}} \Delta) = \Gamma^* \cup \{B \mid B \rightarrow C \in \mathsf{Pos}(\Gamma \vdash \Delta)\} \cup \{C \mid B \rightarrow C \text{ or } \neg C \in \mathsf{Neg}(\Gamma \vdash \Delta)\}.$$

Γ^* is the antecedent of the sequent $\mathfrak{R}(\Gamma \vdash \Delta, \mathbf{S})$, which is obtained by relativizing the sequent $\Gamma \vdash_{\mathbf{S}} \Delta$ with respect to the control set $\mathbf{S}$. The definition of the set Asm is based on the idea of tracking the formulas that may appear in the antecedent of a sequent during a derivation. The name Asm stands for *assumptions*. The underlying idea is that, in order to verify soundness, one must ensure that the information stored in the sequents does not conflict with the prohibitions expressed by the control sets.

Definition 6 A sequent $\Gamma \vdash_{\mathbf{S}} \Delta$ is sound whenever $\mathsf{Asm}(\Gamma \vdash_{\mathbf{S}} \Delta) \,||\, \mathbf{S}$.

To assess the soundness of a sequent, we first substitute quantifiers with finite disjunctions and conjunctions. We then examine the set obtained by adding to the antecedent any formulas that may occur in the antecedent at some stage during the derivation.

Example 3.1 Let us consider the sequent: $\forall x R(x) \vdash_{\mathbf{S}} \neg T(a) \vee U(b)$, we denote it by Seq, with $\mathbf{S} = \{\{T(a) \vee U(b)\}, \{T(a)\}, \{U(b)\}\}$. We first compute $\mathfrak{R}(\forall x R \vdash \neg T(a) \vee U(b), \mathbf{S}) = R(a) \wedge R(b) \vdash_{\mathbf{S}} \neg T(a) \vee U(b)$. Next we determine the set of assumptions:

$$\mathsf{Asm}(\mathsf{Seq}) = \{T(a), R(a) \wedge R(b)\}$$

Therefore , since we have that $T(a), R(a) \wedge R(b) \vdash T(a)$ is derivable and $\{T(a)\} \in \mathbf{S}$,

$$\forall x R(x) \vdash_{\mathbf{S}} \neg T(a) \vee U(b)$$

is not sound.

The controlled sequent calculus $\mathbf{LEX}^{\mathcal{S}}$ is presented in Figure 2. The superscript $\mathcal{S}$ references the specific assignment of control sets to any atomic formulas. Initial sequents are restricted to atomic formulas only, and each initial sequent is equipped with a possibly empty control set that is then propagated throughout the derivation. Control sets $\mathcal{S}(P)$ equipped to initial sequents $P \vdash P$ are closed under substitution, in the sense that if $P \vdash_{\mathcal{S}(P)} P$ is an initial sequent, so is $P[x/t] \vdash_{\mathcal{S}(P)[x/t]} P[x/t]$ for every variable and term t. Intuitively, $P \vdash_{\mathcal{S}(P)} P$ means that any set of formulas in $\mathcal{S}(P)$ destroys the information carried by P and renders it unusable. The binary rules allow for the union of two control sets. It is important to note that the propositional rules are multiplicative, and the structural rules of weakening and contraction are also included. The left rule for the existential and the right rule for the universal quantifier come with the usual restriction imposed on the variable y. The variable is *fresh* in the sense that it does not occur in the conclusion of the rule and is also referred to as the *eigenvariable* of the rule.

3.2 Adding extra-logical axioms

In classical logic, extra-logical sequents are equivalent to *complementary* sequents, which are made up solely of atomic formulas and have an empty intersection between the antecedent and the succedent.

Definition 7 An *extra-logical sequent* is a sequent $\Gamma \vdash \Delta$ where Γ and Δ are multisets of closed atomic formulas, i.e. atomic formulas with closed terms, such that $\Gamma \cap \Delta = \emptyset$.

Given a set of sequents $\mathcal{T}$, we indicate with $\mathbf{LK} + \mathcal{T}$ the system which results from taking each sequent in $\mathcal{T}$ as an initial sequent. Let us consider a set of extra-logical sequents $\mathcal{E}$ and attach a control set $\mathbf{S}$, consisting of sets of closed atomic formulas, to each of them. The set thus obtained represents the knowledge base and contains two kinds of information:[3]

- Static assertions, i.e. information regarding hard facts of the world.
- Controlled assertions, i.e. expressing information containing exceptions encoded in the control sets.

To obtain the closure of a set $\mathcal{U}$ of sequents under contraction, we replace every sequent $\Gamma \vdash \Delta$ in $\mathcal{U}$ with a sequent $\Gamma^* \vdash \Delta^*$ where Γ^* and Δ^* contain every formula of Γ and Δ without repetitions.

Definition 8 (Piazza and Tesi (2024)) For every set $\mathcal{E}$ of extra-logical sequents, its closure under cut $\mathcal{E}^*$ which contains ordered pairs $(\Gamma \vdash_{\mathbf{S}} \Delta; \mathcal{A})$ is inductively defined as follows:

(1) For every $\Gamma \vdash_{\mathbf{S}} \Delta \in \mathcal{E}$, $(\Gamma \vdash_{\mathbf{S}} \Delta; \emptyset) \in \mathcal{E}^*$.
(2) If $(\Gamma \vdash_{\mathbf{S}} \Delta, P; \mathcal{A})$ and $(P, \Gamma' \vdash_{\mathbf{S}'} \Delta'; \mathcal{B})$ are in $\mathcal{E}^*$, then if $(\Gamma, \Gamma' \vdash_{\mathbf{S} \cup \mathbf{S}'} \Delta, \Delta'; \mathcal{A} \cup \mathcal{B} \cup \{P\})$ is an extra-logical sequent, add it to $\mathcal{E}^*$, otherwise discard it.

[3] In the final section we will consider also other kinds of information.

Initial Sequents

$$\frac{}{P \vdash_{\mathbb{S}(P)} P}\, Ax$$

Cut rule

$$\frac{\Gamma \vdash_{\mathbf{S}} \Delta, A \qquad A, \Gamma' \vdash_{\mathbf{T}} \Delta'}{\Gamma, \Gamma' \vdash_{\mathbf{S} \cup \mathbf{T}} \Delta, \Delta'}\, \text{Cut}$$

Logical Rules

$$\frac{A, B, \Gamma \vdash_{\mathbf{S}} \Delta}{A \wedge B, \Gamma \vdash_{\mathbf{S}} \Delta}\, L\wedge \qquad \frac{\Gamma \vdash_{\mathbf{S}} \Delta, A \qquad \Gamma' \vdash_{\mathbf{T}} \Delta', B}{\Gamma, \Gamma' \vdash_{\mathbf{S} \cup \mathbf{T}} \Delta, \Delta', A \wedge B}\, R\wedge$$

$$\frac{A, \Gamma \vdash_{\mathbf{S}} \Delta \qquad B, \Gamma' \vdash_{\mathbf{T}} \Delta'}{A \vee B, \Gamma, \Gamma' \vdash_{\mathbf{S} \cup \mathbf{T}} \Delta, \Delta'}\, L\vee \qquad \frac{\Gamma \vdash_{\mathbf{S}} \Delta, A, B}{\Gamma \vdash_{\mathbf{S}} \Delta, A \vee B}\, R\vee$$

$$\frac{\Gamma \vdash_{\mathbf{S}} \Delta, A \qquad B, \Gamma' \vdash_{\mathbf{T}} \Delta'}{A \to B, \Gamma, \Gamma' \vdash_{\mathbf{S} \cup \mathbf{T}} \Delta, \Delta'}\, L \to \qquad \frac{A, \Gamma \vdash_{\mathbf{S}} B, \Delta}{\Gamma \vdash_{\mathbf{S}} \Delta, A \to B}\, R \to$$

$$\frac{\Gamma \vdash_{\mathbf{S}} \Delta, A}{\neg A, \Gamma \vdash_{\mathbf{S}} \Delta}\, L\neg \qquad \frac{A, \Gamma \vdash_{\mathbf{S}} \Delta}{\Gamma \vdash_{\mathbf{S}} \Delta, \neg A}\, R\neg$$

$$\frac{A[x/y], \Gamma \vdash_{\mathbf{S}} \Delta}{\exists x A, \Gamma \vdash_{\mathbf{S}} \Delta}\, L\exists,\ y \text{ fresh} \qquad \frac{\Gamma \vdash_{\mathbf{S}} \Delta, A[x/t]}{\Gamma \vdash_{\mathbf{S}} \Delta, \exists x A}\, R\exists$$

$$\frac{A[x/t], \Gamma \vdash_{\mathbf{S}} \Delta}{\forall x A, \Gamma \vdash_{\mathbf{S}} \Delta}\, L\forall \qquad \frac{\Gamma \vdash_{\mathbf{S}} A[x/y], \Delta}{\Gamma \vdash_{\mathbf{S}} \Delta, \forall x A}\, R\forall,\ y \text{ fresh}$$

Structural Rules

$$\frac{\Gamma \vdash_{\mathbf{S}} \Delta}{A, \Gamma \vdash_{\mathbf{S}} \Delta}\, LW \qquad \frac{\Gamma \vdash_{\mathbf{S}} \Delta}{\Gamma \vdash_{\mathbf{S}} \Delta, A}\, RW$$

$$\frac{A, A, \Gamma \vdash_{\mathbf{S}} \Delta}{A, \Gamma \vdash_{\mathbf{S}} \Delta}\, LC \qquad \frac{\Gamma \vdash_{\mathbf{S}} \Delta, A, A}{\Gamma \vdash_{\mathbf{S}} \Delta, A}\, RC$$

$$\frac{\Gamma \vdash_{\mathbf{S}} \Delta}{\Gamma \vdash_{\mathbf{S} \cup \mathbf{T}} \Delta}\, \sigma$$

Fig. 2 The controlled sequent calculus $\mathbf{LEX}^{\mathbb{S}}$

(3) Closure of the set thus obtained under contraction.
(4) Repeat steps 1 - 3 until there are no more cut formulas.

For every extra-logical sequent $\Gamma \vdash_{\mathbf{S}} \Delta$ such that $(\Gamma \vdash_{\mathbf{S}} \Delta; \mathcal{A}_1), ..., (\Gamma \vdash_{\mathbf{S}} \Delta; \mathcal{A}_n) \in \mathcal{E}^*$, we say that $\bigcup_{1 \leq i \leq n} \mathcal{A}_i$ is the set of background assumptions of $\Gamma \vdash_{\mathbf{S}} \Delta$.

The assignment of the set of background assumptions is indeed functional. We now show that the procedure described above is well-defined. In particular, it is clear that given a finite set of sequents, the procedure terminates.

Theorem 3.1 *Given a finite set of extra-logical sequents $\mathcal{E}$, the algorithm terminates.*

Proof See Piazza and Tesi (2024). □

The above procedure ensures that there is a set of extra-logical initial sequents that are closed under cut. However, it does not provide any information about the compatibility of the sequents involved. On the one hand, it is possible that a cut between two controlled extra-logical sequents introduces an unsound sequent to the set. On the other hand, it is also possible for a cut between two unsound extra-logical sequents (or between an unsound sequent and a sound sequent) to introduce a sound sequent. Therefore, it is necessary to differentiate between extra-logical initial sequents that are *sound* and those that are *unsound*. In order to define a notion of soundness for extra-logical sequents, we need to take into account the background assumptions attached to the sequents.

Definition 9 An extra-logical sequent $\Gamma \vdash_{\mathbf{S}} \Delta$ in $\mathcal{E}^*$ is sound if and only if

$$\{\mathcal{A}, \Gamma \vdash P \mid P \in \Sigma \in \mathbf{S}\} \cap (\mathcal{E}^* \cup \{\Theta \vdash \Lambda \mid P \in \Theta \cap \Lambda\}) = \emptyset$$

where $\mathcal{A}$ is the set of background assumptions of $\Gamma \vdash_{\mathbf{S}} \Delta$.

Observe that the above definition is equivalent to checking soundness of the sequent $\mathcal{A}, \Gamma \vdash_{\mathbf{S}} \Delta$ in a calculus extended with the extra-logical initial sequents $\mathcal{E}^*$. It is immediate to notice that the notion of soundness for extra-logical initial sequents is well-defined, in the sense that every extra-logical initial sequent is either sound or unsound, and decidable.

Example 3.2 To address concerns that the previous definition may appear ad hoc, consider an example of how our framework handles a scenario regarding a Basenji (which are dogs that cannot bark). Let us consider the following pieces of information:

(1) Toby is a Basenji.
(2) If Toby is a Basenji, then Toby is a dog.
(3) If Toby is a Basenji, Toby does not bark.
(4) If Toby is a dog, then Toby barks, unless he is a Basenji.

We use the constant t for Toby and D, B, and K as unary predicates to express that x is a dog, a Basenji, and barks, respectively.

The sequent $\vdash_{\{B(t)\}} K(t)$ expresses the fact that Toby barks unless he is a Basenji. However, an agent who possesses the relevant information about Toby would not realistically conclude that "Toby barks unless he is a Basenji". Such a sequent is indeed included in the set resulting from the closure under cut of the set of formulas, but is not sound because $B(t) \vdash B(t)$, where $\vdash B(t)$ is an extra-logical initial sequent in the knowledge base (corresponding to item (1)).

Remark 2 The present definition of soundness avoids difficulties arising with the previous one which was defined in terms of set inclusion. In particular, we are able to handle cases of this kind:

(1) $\vdash_{\{P\}} Q$
(2) $Q \vdash P$

In this scenario, a cut between the two extra-logical sequents introduces the sequent $\vdash_{\{P\}} P$ with background assumption $\{Q\}$. This sequent asserts a somewhat counterintuitive statement, namely *P holds unless P holds*. Following our approach, the sequent $\vdash_{\{P\}} P$ is not sound, because the sequent: $Q \vdash P$ is of course provable by assumption.

Let us furnish a concrete example of the application of the algorithm.

Example 3.3 Let us consider a set of extra-logical sequents $\mathcal{E}$ comprising: $P(s) \vdash_{\mathbf{S}_1} R(u)$, $P(s) \vdash_{\mathbf{S}_2} Q(u)$, $R(u) \vdash_{\mathbf{S}_3} T(v)$ and $T(v) \vdash_{\mathbf{S}_4} U(s)$, where $\mathbf{S}_4 = \{\{R(u)\}\}$ and the other control sets are empty.

(1) The closure under cut yields the set $\mathcal{E}^*$ containing the sequents $(P(s) \vdash T(v); \{R(u)\})$, $(T(v) \vdash_{\mathbf{S}_4} U(s); \{R(u)\})$, $(P(s) \vdash_{\mathbf{S}_4} U(s); \{R(u), T(v)\})$.
(2) The set of unsound sequents in $\mathcal{E}^*$ is $\{R(u) \vdash_{\mathbf{S}_4} U(s), P(s) \vdash_{\mathbf{S}_4} U(s)\}$. The remaining sequents in $\mathcal{E}^*$ are sound.

Definition 10 (The sequent calculus $\mathbf{LEX}^{\mathbb{S}}_{\mathcal{E}}$). For every set of extra-logical sequents $\mathcal{E}$, we denote by $\mathbf{LEX}^{\mathbb{S}}_{\mathcal{E}}$ the calculus $\mathbf{LEX}^{\mathbb{S}}$ extended by initial sequents of the form:

$$\frac{}{\Gamma \vdash_{\mathbb{S}(\overline{ax_i})} \Delta}\ \overline{ax_i}$$

where $(\Gamma \vdash_{\mathbb{S}(\overline{ax_i})} \Delta; \mathcal{A}) \in \mathcal{E}^*$ and is sound.

Loosely speaking, $\mathbf{LEX}^{\mathbb{S}}_{\mathcal{E}}$ is obtained from $\mathbf{LEX}^{\mathbb{S}}$ by adding extra-logical information converted into equivalent initial extra-logical sequents $\overline{ax_i}$ which come equipped with a (possibly empty) control set.

As is typical for controlled calculi, we distinguish two types of derivations using the concepts of proof and paraproof. Essentially, a proof is a derivation in which each step is sound and the information contained in the sequents is not in conflict with the information provided by the control sets. Derivations - conceived as rooted trees constructed according to the rules regardless of their soundness - are called paraproofs, so it is clear that proofs form a proper subset of paraproofs.

Definition 11 (Proof, paraproof) A *paraproof* is a finite rooted tree where each node is labelled by a controlled sequent, the leaves are initial sequents or extra-logical sequents, and are obtained according to the rules of calculus. A controlled sequent $\Gamma \vdash_{\mathbf{S}} \Delta$ in a paraproof is sound if $\mathcal{A}, \Gamma \vdash_{\mathbf{S}} \Delta$ is sound, where $\mathcal{A}$ is the union of the sets of the background assumptions of the initial sequents in the subderivation above $\Gamma \vdash_{\mathbf{S}} \Delta$. A *paraproof* is a *proof* whenever every controlled sequent in it is sound.

We give an example of an unsound derivation which involves quantifiers.

Example 3.4 Suppose that $P(t) \vdash_{\{\{P(s)\},\{Q(t)\}\}} R(s)$ is an extra-logical initial sequent and consider the following derivation:

$$\dfrac{\overline{P(t) \vdash_{\{\{P(s)\},\{Q(t)\}\}} R(s)}\ \overline{ax}}{\forall x\, P(x) \vdash_{\{\{P(s)\},\{Q(t)\}\}} R(s)}\ \mathrm{L}\forall$$

The following is not a proof because the application of L∀ is not sound. In particular,

$$\{P(s), P(t)\} \subseteq \mathsf{Asm}(\forall x\, P(x) \vdash_{\mathbf{S}} R(s)) \text{ and } P(s), P(t) \vdash P(t) \text{ is provable.}$$

In other words, the application of the rule for the universal quantifier implicitly adds the information $P(s)$ which is in conflict with the prohibitions expressed by the control set.

Remark 3 As anticipated in the Section 2, to check the soundness of sequents derived in the system $\mathbf{LEX}^{\mathcal{S}}_{\mathcal{E}}$ we can use the system HCPL. Since we are dealing with extensions of classical logic, we need to encompass extra-logical axioms in the refutation calculus. To do so, we need to define a system $\mathsf{HCPL}_{\mathcal{E}}$ obtained from HCPL by modifying the initial sequents as follows:

$$\overline{\mathcal{G} \mid \Gamma \dashv \Delta}\ ax$$

where $P \notin \Gamma \cap \Delta$ for every atomic formula P and $(\Gamma \dashv \Delta, \mathcal{A}) \notin \mathcal{E}^*$ for some set of background assumptions $\mathcal{A}$. Furthermore, if a sequent $\Gamma \dashv \Delta$ is not derivable in $\mathsf{HCPL}_{\mathcal{E}}$, it can be easily shown (and we leave it to the reader) that $\Gamma \dashv \Delta$ is valid in classical propositional logic extended with the axioms in $\mathcal{E}$.

4 Structural analysis

4.1 Cut-elimination

In this section, we prove the cut-elimination theorem for the calculus $\mathbf{LEX}^{\mathcal{S}}_{\mathcal{E}}$. We recall that the *degree* of a cut formula is the number of logical symbols occurring in it and the *height* of a derivation is the length of its maximal branch, where branches consisting of one node are considered to have height 0. We assume that the derivations are *pure*, meaning that an eigenvariable in applications of the rules $R\forall$ and $L\exists$ occurs only in the subtree above the inference. The combination of controlled sequents with extra-logical axioms in $\mathbf{LEX}^{\mathcal{S}}_{\mathcal{E}}$, as well as the explicit presence of contraction, make the task of discerning a cut-elimination procedure a delicate and detail-intensive one.

The idea of the proof. The procedure follows the standard Gentzen-style approach, i.e. we remove topmost instances of the cut rule from the derivations. However, there are some subtleties to consider. In particular, to eliminate instances of cuts we need to be sure that the removal of cuts between sound derivations (proofs) does not introduce unsound derivations, otherwise we would not be in the position to apply the induction hypothesis. Contrary to previous controlled sequent calculi (Piazza & Pulcini, 2017), the new definition of compatibility (and, therefore, of soundness) does not yield that derivations of sound sequents are sound. Indeed, binary rules may lead from unsound premises to sound conclusions.[4] Cut elimination holds because it is carried out working with derivations in which every inferential step is sound.

To start with, we prove the admissibility of the rule of substitution in a way similar to the case of the standard calculi for classical logic.

Lemma 4.1 *The rule of substitution:*

$$\frac{\Gamma \vdash_{\mathbf{S}} \Delta}{\Gamma[x/t] \vdash_{\mathbf{S}[x/t]} \Delta[x/t]}\ Sub[x/t]$$

is height-preserving admissible in $\mathbf{LEX}^{\mathcal{S}}_{\mathcal{E}}$.

Proof The proof is by induction on the height of the derivations. The case of initial sequent follows from the definition of control sets.

If $n > 0$ the most delicate cases involve quantifiers. We consider the situation in which the last rule applied is $L\exists$:

$$\frac{A[y/z], \Gamma \vdash_{\mathbf{S}} \Delta}{\exists y A, \Gamma \vdash_{\mathbf{S}} \Delta}\ \mathrm{L}\exists$$

If $y = x$, then the conclusion follows by applying the induction hypothesis and then L∃ again. If $y \neq x$, then we proceed thus:

$$\dfrac{\dfrac{\dfrac{\dfrac{A[y/z], \Gamma \vdash_{\mathbf{S}} \Delta}{A[y/z][z/u], \Gamma \vdash_{\mathbf{S}} \Delta}\ \text{IH, } u \text{ fresh not in } t}{A[y/u], \Gamma \vdash_{\mathbf{S}} \Delta}\ \equiv}{(A[x/t])[y/u], \Gamma[x/t] \vdash_{\mathbf{S}[x/t]} \Delta[x/t]}\ \text{IH}}{\exists y(A[x/t]), \Gamma[x/t] \vdash_{\mathbf{S}[x/t]} \Delta[x/t]}\ \mathrm{L}\exists$$

The induction hypothesis can be applied twice due to the preservation of the height of the derivation. In particular, the first substitution of z with a fresh variable u is done to prevent the case that t contains z. The fact that $A[y/u][x/t] = A[x/t][y/u]$ is entailed from the fact that u is a fresh variable. □

We show that the cut rule is admissible in $\mathbf{LEX}^{\mathcal{S}}_{\mathcal{E}}$ if we restrict ourselves to considering a certain class of derivations. We show how to eliminate the Gcut, a suitable generalization of the cut rule introduced in Ono (2019):

$$\frac{\Gamma \vdash_{\mathbf{S}} \Delta, A \qquad A, \Pi \vdash_{\mathbf{T}} \Lambda}{\Gamma, \Pi_A \vdash_{\mathbf{T} \cup \mathbf{S}} \Delta_A, \Lambda}\ Gcut$$

[4] To witness this, consider the following example: $P \vee Q \vdash_{\{\{Q\}\}} P \vee Q$. The sequent is indeed sound, but one of the premises - $Q \vdash_{\{\{Q\}\}} P \vee Q$ - obtained by (one possible) application of the rule L∨ is not.

where Δ_A and Π_A are obtained from Δ and Π by deleting some (possibly none) occurrences of A.

Remark 4 The cut rule is not generally admissible, because it is not soundness preserving. More precisely, there are instances of paraproofs in which cuts between proper axioms cannot be eliminated. However, this particular characteristic is actually desirable, as it provides a fresh perspective on transitivity when dealing with the concept of exceptions. In particular, by equating transitivity with cut-elimination, we can assert that transitivity only falters when the interaction between two premises results in an unsound conclusion. Of course, full cut-elimination can be obtained for paraproofs by extending the system with all the initial sequents in $\mathcal{E}^*$, including the unsound ones.

We show how to eliminate certain instances of the Gcut rule, which we call *safe*.

Definition 12 An instance of Gcut between two conclusions of proofs π_1 and π_2:

$$\dfrac{\begin{array}{c}\vdots\ \pi_1\\ \Gamma \vdash_{\mathbf{S}} \Delta, A\end{array} \qquad \begin{array}{c}\vdots\ \pi_2\\ A, \Pi \vdash_{\mathbf{T}} \Lambda\end{array}}{\Gamma, \Pi_A \vdash_{\mathbf{T} \cup \mathbf{S}} \Delta_A, \Lambda}\ Gcut$$

is *safe* whenever π_1 and π_2 do not contain Gcuts and $A \to A, \mathcal{A}, \mathcal{B}, \Gamma, \Pi_A \vdash_{\mathbf{T} \cup \mathbf{S}} \Delta_A, \Lambda$ is sound, where $\mathcal{A}$ and $\mathcal{B}$ are the multisets of background assumptions of the extra-logical initial sequents in π_1 and π_2, respectively.

Proposition 4.2 *For every unary rule, if the conclusion $\Gamma \vdash_{\mathbf{S}} \Delta$ is sound, then the premise is sound.*

Proof For every unary rule ρ and every conclusion $\Gamma \vdash_{\mathbf{S}} \Delta$ and premise $\Gamma' \vdash_{\mathbf{S}'} \Delta'$ we have that $\mathsf{Asm}(\mathfrak{R}(\Gamma' \vdash \Delta', \mathbf{S}')) \subseteq \mathsf{Asm}(\mathfrak{R}(\Gamma \vdash \Delta, \mathbf{S}))$ and $\mathbf{S}' \subseteq \mathbf{S}$. The desired conclusion thus follows from the Proposition 2.2. □

The above result shows that the unary rules of the calculus are invertible w.r.t. soundness in the sense that, whenever the conclusion is sound, so is the premise. On the contrary, rules (not even the unary ones) are not soundness preserving in general. In particular, only unary propositional rules preserve the soundness of the sequents. The cut rule, which is not explicitly part of the calculus, is - in general - neither invertible w.r.t soundness, nor soundness preserving.

Theorem 4.3 *(Gcut-elimination) Given Gcut-free proofs π_1 and π_2 of $\Gamma \vdash_{\mathbf{S}} \Delta, A$ and of $A, \Pi \vdash_{\mathbf{T}} \Lambda$, respectively, there is a proof of $\Gamma, \Pi_A \vdash_{\mathbf{T} \cup \mathbf{S}} \Delta_A, \Lambda$ and $A \to A, \mathcal{A}, \mathcal{B}, \Gamma, \Pi_A \vdash_{\mathbf{T} \cup \mathbf{S}} \Delta_A, \Lambda$ is sound, where $\mathcal{A}$ and $\mathcal{B}$ are the multisets of background assumptions of the extra-logical initial sequents in π_1 and π_2, respectively.*

Proof We remove topmost cuts and the proof runs by double induction, with main induction on the degree of the Gcut formula and secondary induction hypothesis on the sum of the heights of the derivations of the premises of the Gcut. We consider the following two main cases:

(1) The Gcut formula is not principal in one of the premises or it is principal in the application of a structural rule.
(2) The Gcut formula is principal in both premises of the Gcut.

We discuss the two items separately.

The Gcut formula is not principal or principal in a structural rule. The Gcut can be permuted upwards and removed by the secondary induction hypothesis, verifying at each reduction step that the soundness of the Gcuts is preserved due to Proposition 4.2. We detail the case in which the last rule applied is a binary rule, for example $L \rightarrow$.

$$\dfrac{\Gamma \vdash_{\mathbf{T}} \Delta, A \qquad \dfrac{\Gamma' \vdash_{\mathbf{T}'} \Delta', B \qquad A, C, \Gamma'' \vdash_{\mathbf{T}''} \Delta''}{A, B \rightarrow C, \Gamma', \Gamma'' \vdash_{\mathbf{T}' \cup \mathbf{T}''} \Delta', \Delta''}\,\mathrm{L}\rightarrow}{B \rightarrow C, \Gamma, (\Gamma', \Gamma'')_A \vdash_{\mathbf{T} \cup \mathbf{T}' \cup \mathbf{T}''} (\Delta)_A, \Delta', \Delta''}\,\mathrm{Gcut}$$

We suppose for brevity that A does not occur in Γ' (otherwise we apply twice the secondary induction hypothesis). Hence we proceed as follows:

$$\dfrac{\Gamma' \vdash_{\mathbf{T}'} \Delta', B \qquad \dfrac{\Gamma \vdash_{\mathbf{T}} \Delta, A \qquad A, C, \Gamma'' \vdash_{\mathbf{T}''} \Delta''}{C, \Gamma, \Gamma''_A \vdash_{\mathbf{T} \cup \mathbf{T}''} \Delta_A, \Delta', \Delta''}\,\mathrm{Gcut}}{B \rightarrow C, \Gamma, \Gamma', \Gamma''_A \vdash_{\mathbf{T} \cup \mathbf{T}' \cup \mathbf{T}''} \Delta_A, \Delta', \Delta''}\,\mathrm{L}\rightarrow$$

The Gcut is removed by the secondary induction hypothesis on the sum of the height of the derivations. The derivation of $C, \Gamma, \Gamma''_A \vdash_{\mathbf{T} \cup \mathbf{T}''} \Delta_A, \Delta', \Delta''$ is sound by the application of the induction hypothesis. Furthermore, the conclusion is easily seen to be sound due to the soundness of $\mathcal{A}, \mathcal{B}, A \rightarrow A, B \rightarrow C, \Gamma, \Gamma', \Gamma'' \vdash_{\mathbf{T} \cup \mathbf{T} \cup \mathbf{T}''} \Delta, \Delta', \Delta''$. Hence the entire derivation is a proof, because every sequent occurring therein is sound.

The Gcut formula is principal in both the premises of the cut. We limit ourselves to discussing two cases, one concerning extra-logical initial sequents and the other the quantifiers.

- If both the premises are extra-logical sequents, we have:

$$\dfrac{\overline{\Gamma \vdash_{\mathcal{S}(\overline{ax}_i)} \Delta, P}\,\overline{ax} \qquad \overline{P, \Gamma' \vdash_{\mathcal{S}(\overline{ax}_j)} \Delta'}\,\overline{ax}}{\Gamma, \Gamma' \vdash_{\mathcal{S}(\overline{ax}_i) \cup \mathcal{S}(\overline{ax}_j)} \Delta, \Delta'}\,Gcut$$

  Here, the conclusion is an extra-logical sequent as well and the conclusion is an instance of $\overline{ax}$ because by assumption $P \rightarrow P, \mathcal{A}, \mathcal{B}, \Gamma, \Gamma' \vdash_{\mathcal{S}(\overline{ax}_i) \cup \mathcal{S}(\overline{ax}_j)} \Delta, \Delta'$ is sound.
- We discuss the case of the universal quantifier. We have:

$$\dfrac{\dfrac{\Gamma \vdash_{\mathbf{S}} \Delta, A[x/y]}{\Gamma \vdash_{\mathbf{S}} \Delta, \forall x A}\,\mathrm{R}\forall \qquad \dfrac{A[x/t], \Gamma' \vdash_{\mathbf{T}} \Delta'}{\forall x A, \Gamma' \vdash_{\mathbf{T}} \Delta'}\,\mathrm{L}\forall}{\Gamma, \Gamma'_{\forall x A} \vdash_{\mathbf{S} \cup \mathbf{T}} \Delta_{\forall x A}, \Delta'}\,\mathrm{Gcut}$$

  There are three subcases: if neither Δ nor Γ' contain $\forall x A$, if one of them contains $\forall x A$ or if both contain $\forall x A$. We deal with the last one, which is the most general. We proceed as follows:

$$
\dfrac{
\dfrac{
\dfrac{\dfrac{\Gamma \vdash_{\mathbf{S}} \Delta, A[x/y] \qquad \forall x A, \Gamma' \vdash_{\mathbf{T}} \Delta'}{\Gamma, \Gamma'_{\forall x A} \vdash_{\mathbf{S}\cup\mathbf{T}} \Delta_{\forall x A}, \Delta', A[x/y]}\,\text{Gcut}}{\Gamma, \Gamma'_{\forall x A} \vdash_{\mathbf{S}\cup\mathbf{T}} \Delta_{\forall x A}, \Delta', A[x/t]}\,\text{Sub}\,[y/t]
\qquad
\dfrac{\Gamma \vdash_{\mathbf{S}} \Delta, \forall x A \qquad A[x/t], \Gamma' \vdash_{\mathbf{T}} \Delta'}{\Gamma, A[x/t], \Gamma'_{\forall x A} \vdash_{\mathbf{S}\cup\mathbf{T}} \Delta_{\forall x A}, \Delta'}\,\text{Gcut}
}{\Gamma^2, \Gamma'^2_{\forall x A} \vdash_{\mathbf{S}\cup\mathbf{T}} \Delta^2_{\forall x A}, \Delta'^2}\,\text{Gcut}
}{\Gamma, \Gamma'_{\forall x A} \vdash_{\mathbf{S}\cup\mathbf{T}} \Delta_{\forall x A}, \Delta'}\,\text{several contractions}
$$

The first two Gcuts are removed by secondary induction hypothesis on the sum of the heights of the derivations. The latter Gcut is removed by primary induction hypothesis on the degree of the cut formula. In all the three cases the soundness of the conclusion is obtained by the induction hypothesis and by the soundness of $\mathcal{A}, \mathcal{B}, \forall x A \rightarrow \forall x A, \Gamma, \Gamma'_{\forall x A} \vdash_{\mathbf{S}\cup\mathbf{T}} \Delta_{\forall x A}, \Delta'$. Furthermore, the step of substitution does not upset the soundness of the resulting sequent. Indeed, $\Gamma, \Gamma'_{\forall x A} \vdash_{\mathbf{S}\cup\mathbf{T}} \Delta_{\forall x A}, \Delta', A[x/y]$ is sound by induction hypothesis and the conclusion of $Sub[y/t]$, i.e., $\Gamma, \Gamma'_{\forall x A} \vdash_{\mathbf{S}\cup\mathbf{T}} \Delta_{\forall x A}, \Delta', A[x/t]$, is sound due the soundness of $\mathcal{A}, \mathcal{B}, \forall x A \rightarrow \forall x A, \Gamma, \Gamma'_{\forall x A} \vdash_{\mathbf{S}\cup\mathbf{T}} \Delta_{\forall x A}, \Delta'$. The case of the existential quantifier is dual and we omit the details.

□

4.2 A comparison with first-order logic

The approach presented in this paper relies on two crucial components: the inclusion of extra-logical axioms and their interaction with control sets. This combination allows for the expression of complex information incorporating both prohibitions and exceptions.

The calculus $\mathbf{LEX}^{\mathcal{S}}_{\mathcal{E}}$ manipulates sequents of the form $\Gamma \vdash_{\mathbf{E}} \Delta$ as its basic syntactic elements. We would like to specify that our approach, despite being related to classical logic, presents significant differences. In particular, there is not *prima facie* a way to translate a controlled sequent into a formula of the language of first-order logic, thus making soundness internal with respect to the system.

Our solution has two advantages. Firstly, it is interesting to explore the relationship between soundness and normalization. Control sets enable the preservation of soundness throughout the cut-elimination process by keeping track of prohibitions in derivations and proposing a direct cut-elimination strategy.

Secondly, the calculus $\mathbf{LEX}^{\mathcal{S}}_{\mathcal{E}}$ extends classical logic by incorporating initial sequents that encode extra-logical information. The sequents are closed under (safe instances of) cut, and therefore, modulo the soundness of the conclusion, a cut-elimination theorem is ensured, resulting in a strong subformula property. The addition of extra-logical axioms is handled through initial sequents, thus maintaining the soundness of the system.

Claims of provability and unprovability are harnessed to design calculi for non-monotonic logics. Nevertheless, when dealing with first-order languages, this approach encounters limitations due to the inherent undecidability of first-order logic (although it is worth mentioning that results in this sense have been achieved in Giordano et al.

(2020)). We overcome this difficulty by relativizing the quantified formulas instantiating them with the terms occurring in the control sets.

5 Reasoning about typicality

5.1 A sequent calculus for typicality

In order to incorporate reasoning about typicality information into our framework, we need to modify the framework that has been adopted thus far. To start with, we restrict ourselves to considering first-order languages without function symbols. We can conceptualize a typicality statement as a kind of "iceberg of generality" with the concealed part encoding exceptions. For example, consider the statement *typically, adults work*. Its concealed part encompasses exceptions such as "unless they are students", "unless they are unemployed", "unless they are retired", "unless they are on holiday" and so on.[5] Due to the formulation of these kind of statements, we must expand our notion of extra-logical sequent so as to include formulas containing quantifiers.

Definition 13 A *typicality axiom* is any sentence of the form $\forall x(P_1(x) \wedge ... \wedge P_m(x) \rightarrow Q_1(x) \vee ... \vee Q_n(x))$, where P_i and Q_j are unary predicates. Each typicality axiom comes equipped with an *exception list* $E_1(x) \vee ... \vee E_l(x)$, where $E_k(x)$ is an atomic formula for every $i \in \{1, ..., l\}$.

We are aware that listing exceptions is not a common trend in non-monotonic reasoning. However, we recall that we aim to represent inferential patterns of a rational agent who is able to handle a *finite* set of exceptions. The idea is that $E_1(x) \vee ... \vee E_l(x)$ expresses a list of possible conflicting pieces of information which can invalidate the inclusion. We introduce a class of extra-logical axioms, i.e. typicality axioms $\mathbb{TA}$, which contains typicality axioms and formulas built from atoms and connectives (but not quantifiers), i.e. extra-logical axioms as previously described. We now discuss an algorithm of conversion of typicality axioms into initial sequents.

Definition 14 Given a set X of axioms in $\mathbb{TA}$, we generate a set of initial sequents $\mathcal{T}_X$ as follows:

(1) For every formula A in X we distinguish two subcases. If A is an extra-logical axiom, then proceed as in Section 3. If A is a typicality axiom of the form $\forall x(P_1(x) \wedge ... \wedge P_m(x) \rightarrow Q_1(x) \vee ... \vee Q_n(x))$, then we take all the initial sequents of the form $P_1(t), ..., P_m(t) \vdash Q_1(t), ..., Q_n(t)$, for every term t of the language.
(2) We associate to every initial sequent $P_1(t), ..., P_m(t) \vdash Q_1(t), ..., Q_n(t)$ obtained by a typicality axiom the control set **S** which is generated by the singleton $\{E_1(t) \vee ... \vee E_l(t)\}$, where $E_1(x) \vee ... \vee E_l(x)$ is the exception list associated to the typicality axiom.

[5] This, on our view, marks a major difference between typicality sentences and generic ones like "adults work", "birds fly", which do not have such concealed part covered by the adverb "typically". Thus we disagree with the claim that if one adds "typically" to a generic sentence one get at most a slight variation of meaning (Krifka et al.., 1995).

Definition 15 Given a set X of axioms in $\mathbb{TA}$ and its associated set of initial sequents $\mathcal{T}_X$, its closure under cut $\mathcal{T}^*_X$ which contains ordered pairs $(\Gamma \vdash_{\mathbf{S}} \Delta; \mathcal{A})$ is inductively defined:

(1) For every $\Gamma \vdash_{\mathbf{S}} \Delta \in \mathcal{T}^*_X$, then $(\Gamma \vdash_{\mathbf{S}} \Delta; \emptyset) \in \mathcal{T}^*_X$.
(2) If $(\Gamma \vdash_{\mathbf{S}} \Delta, P; \mathcal{A})$ and $(P, \Gamma' \vdash_{\mathbf{S}'} \Delta'; \mathcal{B})$ in $\mathcal{T}^*_X$, then if $(\Gamma, \Gamma' \vdash_{\mathbf{S} \cup \mathbf{S}'} \Delta, \Delta'; \mathcal{A} \cup \mathcal{B} \cup \{P\})$ is an extra-logical sequent, add it to $\mathcal{T}^*_X$, otherwise discard it.
(3) Closure under contraction.
(4) Repeat steps (2)-(3).

We can show the following:

Proposition 5.1 *The membership of the set* $\mathcal{T}^*_X$ *is decidable.*

Proof The set of terms occurring in the initial sequents obtained by extra-logical axioms is finite: $\{t_1, ..., t_n\}$. To compute the closure under cut it suffices to consider a finite number of initial sequents obtained from the typicality axioms, namely those instantiated with a term t_i, where $i \in \{1, ..., n\}$. Next, for every term in the language $s \notin \{t_1, ..., t_n\}$ (including variables), we only need to consider the finite number of initial sequents obtained by cuts between initial sequents resulting from typicality axioms. □

As noted by a reviewer, the set of initial sequents generated after applying the algorithm in Definition 15 may be infinite. However, the procedure terminates in the sense that for any term t instantiating the sequents derived from typicality axioms, there exists a finite number of cuts to be performed. For any given sequent, it is decidable whether it belongs to the closure under cut and contraction of $\mathcal{T}^*_X$. This ensures that the extra-logical information can be handled in an effective and controlled manner. Notice that the restriction to languages not containing function symbols is crucial to obtain the termination of the procedure.[6] The soundness of the sequents in the set $\mathcal{T}^*_X$ is defined as in the case of the set $\mathcal{E}^*_X$.

Example 5.1 Let us consider the set of formulas $\{\forall x(P(x) \to Q(x)), Q(r) \to R(u) \wedge T(u)\}$. This yields the set $\mathcal{T}^*_X = \{P(t) \vdash_{\mathbf{S}} Q(t) \,|\, t \in TER\} \cup \{Q(r) \vdash_{\mathbf{S}} R(u); Q(r) \vdash_{\mathbf{S}} T(u); P(r) \vdash_{\mathbf{S}} R(u); P(r) \vdash_{\mathbf{S}} T(u)\}$.

We introduce a new sequent system called $\mathbf{LET}^{\mathcal{S}}_{\mathcal{T}}$, which shares the same rules as $\mathbf{LEX}^{\mathcal{S}}$ but has different initial sequents. The initial sequents of $\mathbf{LET}^{\mathcal{S}}_{\mathcal{T}}$ are of two kinds:

- *logical* initial sequents: $P \vdash P$.
- *typicality* initial sequents: $\Gamma \vdash_{\mathbf{S}} \Delta$ where $(\Gamma \vdash \Delta; \mathcal{A}) \in \mathcal{T}^*$ and $\mathbf{S}$ is a (possibly empty) control set.

In order to prove the cut-elimination theorem some auxiliary lemmas are needed. In particular, we start by proving the substitution lemma.

[6] To witness the impossibility of this, consider the case of a monadic predicate N, a constant 0 and a unary function symbol s together with an axioms $N(0)$ and $\forall x(N(x) \to N(s(x)))$ to get an undecidable theory.

Lemma 5.2 *The rule:*

$$\frac{\Gamma \vdash_{\mathbf{S}} \Delta}{\Gamma[x/t] \vdash_{\mathbf{S}[x/t]} \Delta[x/t]}\ \mathit{Sub[x/t]}$$

is height-preserving admissible for every variable x and term t in the system $\mathbf{LET}^{\mathcal{S}}_{\mathcal{T}}$.

Proof The proof is by induction on the height of the derivations. We need to consider the case when $\Gamma \vdash_{\mathbf{S}} \Delta$ is an extra-logical or typicality initial sequent. We distinguish two subcases: whether it contains free variables or not. In the second case, the substitution has no effect on the sequent. In the first case, we can assume that $\Gamma \vdash_{\mathbf{S}} \Delta$ is of the form $P_1(x), \ldots, P_m(x) \vdash_{\mathbf{S}} Q_1(x), \ldots, Q_n(x)$, where $\mathbf{S}$ only contains sets of formulas of the form $E_1(x) \vee ... \vee E_k(x)$, with $n \geqslant 1$. By definition of the conversion algorithm, the sequent $\Gamma[x/t] \vdash_{\mathbf{S}[x/t]} \Delta[x/t]$ is also an inclusion sequent.

□

Safe Gcuts are defined as in the case of system $\mathbf{LEX}^{\mathcal{S}}_{\mathcal{E}}$. We now articulate the cut-elimination proof.

Theorem 5.3 *Every proof in* $\mathbf{LET}^{\mathcal{S}}_{\mathcal{T}}$ *of* $\Gamma \vdash_{\mathbf{S}} \Delta$ *containing safe Gcuts has a cut-free proof.*

Proof We follow the same argument as for $\mathbf{LEX}^{\mathcal{S}}_{\mathcal{E}}$, but limit ourselves to considering the case where one of the premises of the cut is an extra-logical or typicality initial sequent. To establish soundness we follow the arguments outlined for the system $\mathbf{LEX}^{\mathcal{S}}_{\mathcal{E}}$ and so we omit the details. Suppose the left premise of the cut is a proper initial sequent:

$$\dfrac{\dfrac{}{P_1(s_1), ..., P_m(s_m) \vdash_{\mathbf{S}} Q_1(t_1), ..., Q_n(t_n)}\ \overline{ax} \qquad Q_n(t_n), \Pi \vdash_{\mathbf{T}} \Lambda}{P_1(s_1), ..., P_m(s_m), \Pi \vdash_{\mathbf{S} \cup \mathbf{T}} Q_1(t_1), ..., Q_{n-1}(t_{n-1}), \Lambda}\ \text{Gcut}$$

The argument is by induction on the height of the right premise of the cut. If it is an axiomatic initial sequent, then there is nothing to prove. If it is a proper initial sequent, then the conclusion of cut is a typicality sequent in $\mathcal{T}^{*}_{X}$. In every other case we permute the cut upwards, applying the substitution lemma whenever required in order to avoid clashes of variables. For instance:

$$\dfrac{\dfrac{}{P_1(s_1), ..., P_m(s_m) \vdash_{\mathbf{S}} Q_1(t_1), ..., Q_n(t_n)}\ \overline{ax} \qquad \dfrac{Q_n(t_n), A[z/y], \Pi \vdash_{\mathbf{T}} \Lambda}{Q_n(t_n), \exists z A, \Pi \vdash_{\mathbf{T}} \Lambda}\ \text{L}\exists}{\exists x A, P_1(s_1), ..., P_m(s_m), \Pi \vdash_{\mathbf{S} \cup \mathbf{T}} Q_1(t_1), ..., Q_{n-1}(t_{n-1}), \Lambda}\ \text{Gcut}$$

We assume that y occurs among the free variables in $P_1(s_1), ..., P_m(s_m) \vdash_{\mathbf{S}} Q_1(t_1), ..., Q_n(t_n)$ (the other case is easier), then we proceed as follows:

$$\dfrac{\dfrac{\dfrac{}{P_1(s_1), ..., P_m(s_m) \vdash_{\mathbf{S}} Q_1(t_1), ..., Q_n(t_n)}\ \overline{ax} \qquad \dfrac{Q(t_n), A[z/y], \Pi \vdash_{\mathbf{T}} \Lambda}{Q(t_n), A[z/u], \Pi \vdash_{\mathbf{T}} \Lambda}\ \mathit{Sub}[y/u]}{A[z/u], P_1(s_1), ..., P_m(s_m), \Pi \vdash_{\mathbf{S} \cup \mathbf{T}} Q_1(t_1), ..., Q_{n-1}(t_{n-1}), \Lambda}\ \text{Gcut}}{\exists z A, P_1(s_1), ..., P_m(s_m), \Pi \vdash_{\mathbf{S} \cup \mathbf{T}} Q_1(t_1), ..., Q_{n-1}(t_{n-1}), \Lambda}\ \text{L}\exists$$

where u is fresh and the cut is removed invoking the secondary induction hypothesis on the sum of the height of the derivations. □

We have obtained a calculus which enjoys cut-elimination and in which typicality statements are transformed into initial sequents equipped with a control set.

Remark 5 Our interpretation of typicality statements is similar in spirit to that of default logic. A (normal) default rule is a rule of the form:

$$\frac{B(x) \ : \ MF(x)}{F(x)}\ \text{Default}$$

where M means "it is consistent to assume". So, the interpretation of the default above is: *if x is a bird and it is consistent to assume that x flies, infer that x flies*. In the setting of controlled sequent calculi this translates to an initial sequent of the form

$$B(x) \vdash_{\mathbf{S}} F(x)$$

and all its substitution instances, where $\mathbf{S}$ is the control set generated by a disjunction of negative pieces of information. The negative information is internalized in the calculus by the control set attached to the proper initial sequent. The soundness condition for controlled sequents can then be rephrased as expressing a form of consistency. Indeed, consider the extra-logical sequent $B(x) \vdash_{\mathbf{S}} F(x)$, where $\mathbf{S}$ contains a set of exceptions, for example $\{O(x) \vee P(x)\}$, expressing that *x is an ostrich* or *x is a penguin*, respectively. In particular, it amounts to verifying that the sequent:

$$B(x) \vdash O(x) \vee P(x)$$

is not derivable. This, due to well-known equivalence results, corresponds to verifying the underivability of the sequent:

$$B(x), \neg O(x), \neg P(x) \vdash$$

Hence our approach consists in checking the compatibility (or the consistency) of the knowledge base with respect to a list of specified exceptions.

We wish to draw attention to some specific features of our approach. First, our presentation of typicality statements carefully distinguishes between typicality statements and universal assertions. In particular, as expected, we can easily show that universal statements imply typicality ones, but not vice versa.

Proposition 5.4 *If* $\vdash \forall x(P(x) \to Q(x))$ *is derivable with* P, Q *unary predicates, then so is* $P(t) \vdash_{\mathbf{S}} Q(t)$ *for any term t and any control set* $\mathbf{S}$*, provided that for every* $\Lambda \in \mathbf{S}$*,* $\Lambda \neq \{P(t)\}$*.*

Proof We proceed as follows:

$$\dfrac{\dfrac{\vdash \forall x(P(x) \to Q(x)) \qquad \forall x(P(x) \to Q(x)), P(t) \vdash Q(t)}{P(t) \vdash Q(t)}\ \text{Gcut}}{P(t) \vdash_{\mathbf{S}} Q(t)}\ \sigma$$

The conclusion is sound by the assumption on the control set **S**. □

The converse does not hold. Indeed, let us consider the typicality statement *usually birds fly unless they are penguins or ostriches* and consider the following derivation:

$$\dfrac{\dfrac{B(x) \vdash_{\{\{O(x)\vee P(x)\},\{O(x)\},\{P(x)\}\}} F(x)}{\vdash_{\{\{O(x)\vee P(x)\},\{O(x)\},\{P(x)\}\}} B(x) \to F(x)}\,\mathrm{R}{\to}}{\vdash_{\{\{O(x)\vee P(x)\},\{O(x)\},\{P(x)\}\}} \forall x(B(x) \to F(x))}\,\mathrm{R}\forall$$

in which $B(x)$ and $F(x)$ stand for x is a bird and x flies, respectively and $O(x)$ and $P(x)$ express the exceptions. It is immediate to notice that the application of the rule R∀ is not correct because it violates the freshness of variable x.

Another interesting point to note is that the sequent calculus $\mathbf{LET}_{\mathbb{T}}^{\mathbb{S}}$ does not fundamentally depart from classical logic. Its non-monotonic features arise within a classical deductive framework through the interplay between positive information conveyed by axioms and negative information conveyed by control sets. This can be seen as an attempt to reconcile classical and typicality reasoning without resorting to more complex modal operators, as put forward by some other approaches (Giordano et al., 2013, 2020).

5.2 Another look at Tweety the bird

To illustrate the dynamics of the calculus, we offer a coherent representation of Reiter's famous example of Tweety the bird (Reiter, 1978, 1980, 1987), emphasizing the notion of typicality involved.

(1) Typically, birds fly.
(2) Penguins do not fly.
(3) Tweety is a bird.

Being informed that Tweety is a bird, an agent concludes that Tweety flies. However, she promptly retracts this conclusion if the information that Tweety is a penguin is added. According to our treatment of typicality axioms, we want to expand our base calculus with the following statements in the class $\mathbb{TA}$:

Formulas	Sequents	Interpretation
$\forall x(B(x) \to F(x))$ exc. $\mathbf{S} = \{P(x) \vee O(x) \vee T(x)\}$	$\{B(s) \vdash_{\mathbf{S}[s/x]} F(s) \mid s \in TER\}$	Every bird flies unless it is a penguin, a turkey, or an ostrich
$\forall x(P(x) \to \neg F(x))$	$\{P(s), F(s) \vdash \ \mid s \in TER\}$	Penguins do not fly
$\forall x(P(x) \to B(x))$	$\{P(s) \vdash B(s) \mid s \in TER\}$	Penguins are birds

In the terminology of Reiter and Criscuolo (1983), the first item expresses a "prototypical fact" and so it is accompanied by an exception list, whereas the second and the third express “hard” facts and so no exception is envisaged. Let us now observe the following.

Proposition 5.5 *The sequent* $P(t) \vdash_{\mathbf{S}} F(t) \wedge \neg F(t)$ *is derivable as a paraproof in* $\mathbf{LET}^{\mathbb{S}}_{\mathbb{J}}$ *with cut, but it does not admit a proof in* $\mathbf{LET}^{\mathbb{S}}_{\mathbb{J}}$.

Proof Let us observe first that $P(t) \vdash_{\mathbf{S}} F(t) \wedge \neg F(t)$ is derivable as a paraproof in $\mathbf{LET}^{\mathbb{S}}_{\mathbb{J}}$ extended with cut.

$$\dfrac{\dfrac{\dfrac{}{P(t) \vdash B(t)}\overline{ax} \qquad \dfrac{}{B(t) \vdash_{\mathbf{S}} F(t)}\overline{ax}}{P(t) \vdash_{\mathbf{S}} F(t)}\text{Gcut} \qquad \dfrac{\dfrac{}{P(t), F(t) \vdash}\overline{ax}}{P(t) \vdash \neg F(t)}\text{R}\neg}{\dfrac{P(t), P(t) \vdash_{\mathbf{S}} F(t) \wedge \neg F(t)}{P(t) \vdash_{\mathbf{S}} F(t) \wedge \neg F(t)}\text{LC}}\text{R}\wedge$$

Notice that the application of Gcut is not safe because it introduces an unsound sequent. Moreover, there cannot be a proof of sequent $P(t) \vdash_{\mathbf{S}} F(t) \wedge \neg F(t)$, since the endsequent is unsound. □

Summing up, we are able to account for a situation in which:

(1) *Typically, birds fly*, because flight is characteristic of the type “birds";
(2) Tweety, despite being a bird, does not fly as a conclusion obtained from an extra-logical piece of information;
(3) The derivation of a contradiction which would arise classically from the fact that Tweety should fly as a bird, but does not fly because it is a penguin, is rejected.

This is a simple but telling example of an application of the method of extra-logical axioms alongside control sets. Within the framework of controlled sequent calculi, the Tweety example illustrates a close connection between non-monotonicity and paraconsistency.

5.3 The intransitivity of “typically”

A key aspect of reasoning about typicality – as noted by Reiter and Criscuolo (Reiter & Criscuolo, 1983) – is that the notion of typicality (as conveyed by the adverb “typically") fails in general to express transitivity.
Consider, indeed, the syllogism below, which mirrors the structure of a syllogism in the mode of Barbara:

(1) Typically, A’s are B’s.
(2) Typically, B’s are C’s.
(3) Typically, A’s are C’s.

The conclusion often turns out to be unsound. For example,

(1) Typically, university students are adults.
(2) Typically, adults are employed.
(3) Typically, university students are employed.

University students tend to be not employed, yet the default reasoning endorses the conclusion that they are typically employed (Reiter & Criscuolo, 1983).

Our solution for blocking transitivity is systematic and purely of proof-theoretic character.[7] As we have already remarked, our systems are not thoroughly transitive. In particular, cut-elimination fails in presence of proper axioms which lead to unsound conclusions

Let us consider the example above. Let A, E and S be unary predicates corresponding to *adult, employed*, and *student*, respectively. Then, we consider the system $\mathbf{LET}_{\mathcal{T}}^{\mathcal{S}}$ including the typicality axioms (1) – (2) where (1) comes equipped with the exception list $\mathbf{S}_1 = \{C(x)\}$, where $C(x)$ stands for *x is a child prodigy* and (2) comes equipped with the exception list $\mathbf{S}_2 = \{S(x) \vee R(x)\}$ which stands for *x is a student* and *x is retired*, respectively. If we compute the set $\mathcal{T}_X^*$ we obtain the set:

$$\{S(t) \vdash_{\mathbf{S}_1} A(t) \mid t \in TER\}$$
$$\cup\{A(t) \vdash_{\mathbf{S}_2} E(t) \mid t \in TER\} \cup \{S(t) \vdash_{\mathbf{S}_1 \cup \mathbf{S}_2} E(t) \mid t \in TER\}.$$

where TER is the set of terms of the language. A closer inspection reveals that the set of extra-logical initial sequents added to the system $\mathbf{LET}_{\mathcal{T}}^{\mathcal{S}}$ is $\{S(t) \vdash_{\mathbf{S}_1} A(t) \mid t \in TER\} \cup \{A(t) \vdash_{\mathbf{S}_2} E(t) \mid t \in TER\}$, simply because every sequent in $\{S(t) \vdash_{\mathbf{S}_1 \cup \mathbf{S}_2} E(t) \mid t \in TER\}$ is unsound as $\{S(t)\}$ is in every control set.

The following proposition is an immediate consequence of the previous observation.

Proposition 5.6 *The sequent* $S(t) \vdash_{\mathbf{S}_1 \cup \mathbf{S}_2} E(t)$ *is not derivable as a proof* $\mathbf{LET}_{\mathcal{T}}^{\mathcal{S}}$.

Proof The sequent cannot be derived as a proof, because it is not sound. The sequent is only derivable as a paraproof with the cut:

$$\frac{S(t) \vdash_{\mathbf{S}_1} A(t) \qquad A(t) \vdash_{\mathbf{S}_2} E(t)}{S(t) \vdash_{\mathbf{S}_1 \cup \mathbf{S}_2} E(t)}\ \text{Gcut}$$

which is not safe. □

6 Recovering the KLM postulates

The Kraus-Lehmann-Magidor (KLM) postulates define a set of rules that characterize the concept of derivability under normal circumstances (Kraus et al., 1990). Our work does not focus on non-monotonic reasoning as traditionally intended, since we explicitly consider exceptions. Specifically, the use of control sets to represent exceptions leads to a relativization of the postulates with respect to a given control set.

In this section, we rephrase the KLM postulates in terms of the derivability condition that arises from our sequent system $\mathbf{LET}_{\mathcal{T}}^{\mathcal{S}}$, and examine the extent to which we can draw comparisons. We interpret the non-monotonic consequence relation $\mid\!\sim$ as $\vdash \mathbf{S}$ for a controlled set $\mathbf{S}$.

[7] For a diagrammatic representation of this intransitivity phenomenon using Venn diagrams, see Bhattacharjee and Mario (2025).

- **Reflexivity.** Reflexivity is expressed by $A \mathrel{|\!\sim} A$ for every formula A. We have to observe that $A \vdash A$ is derivable in $\mathbf{LET}^{\mathbb{S}}_{\mathbb{T}}$ for every A, since tautological initial sequents are not equipped with a control set. Hence, we can affirm that $A \vdash_{\mathbf{S}} A$ is derivable as long as it is sound and that the last rule applied in the derivation is the rule σ.
- **Left logical equivalence.** Left logical equivalence can be interpreted as a rule asserting that:

$$\frac{\vdash A \leftrightarrow B \qquad A \mathrel{|\!\sim} C}{B \mathrel{|\!\sim} C}$$

By the cut admissibility theorem this rule can be simulated in $\mathbf{LET}^{\mathbb{S}}_{\mathbb{T}}$ as follows whenever the sequent $A \rightarrow A, B \vdash_{\mathbf{S}} C$ is sound:

$$\frac{\dfrac{\dfrac{\vdash A \leftrightarrow B}{\vdash B \rightarrow A}\,\text{Adm rule}}{B \vdash A}\,\text{Adm rule} \qquad A \vdash_{\mathbf{S}} C}{B \vdash_{\mathbf{S}} C}\,\text{Gcut}$$

where *Gcut* can be applied due to Theorem 5.3 and the two steps *Adm rule* denote applications of the invertibility of the rules of classical propositional logic.
- **Right weakening.** The rule of right weakening is:

$$\frac{\vdash A \rightarrow B \qquad C \mathrel{|\!\sim} A}{C \mathrel{|\!\sim} B}$$

Once again this rule can be simulated using an essential step of cut.

$$\frac{C \vdash_{\mathbf{S}} A \qquad \dfrac{\vdash A \rightarrow B \qquad A \rightarrow B, A \vdash B}{A \vdash B}\,\text{Gcut}}{C \vdash_{\mathbf{S}} B}\,\text{Gcut}$$

The rule needs to be equipped with the side condition that the sequent $A \rightarrow A, C \vdash_{\mathbf{S}} B$ is sound.
- **Cut.** Cut is expressed as:

$$\frac{A \mathrel{|\!\sim} B \qquad B, A \mathrel{|\!\sim} C}{A \mathrel{|\!\sim} C}$$

The cut rule can be easily simulated as follows:

$$\frac{\dfrac{A \vdash_{\mathbf{S}} B \qquad B, A \vdash_{\mathbf{T}} C}{A, A \vdash_{\mathbf{S} \cup \mathbf{T}} C}\,\text{Gcut}}{A \vdash_{\mathbf{S} \cup \mathbf{T}} C}\,\text{LC}$$

provided $B \rightarrow B, A \vdash_{\mathbf{S} \cup \mathbf{T}} C$ is sound.
- **Cautious monotonicity.** Cautious monotonicity is expressed by the rule:

$$\frac{A \mathrel{|\!\sim} B \qquad A \mathrel{|\!\sim} C}{A, B \mathrel{|\!\sim} C}$$

Cautious monotonicity is a desirable feature for a non-monotonic logic because it suggests that if B and C normally follow from A, then we can conclude that C follows from B and A. In general, the rule can be simulated with a step of $\mathrm{L}W$ whenever the end-sequent is sound.

A weaker version of cautious monotonicity may be seen to hold in the following form:

$$\frac{A \vdash B \qquad A \vdash \bigwedge_{i \leq n} B_i \qquad A \mid\!\sim C}{A, B \mid\!\sim C}\ \mathrm{CM}^-$$

where $B_1, \ldots, B_n$ are the formulas in the set $\{C \mid C \to D \text{ or } \neg C \in \mathsf{Neg}(B)\}$. Indeed, let us assume that B follows classically from A and that the sequent $A \vdash_{\mathbf{S}} C$ is derivable for some control set $\mathbf{S}$. Let us suppose - towards a contradiction - that $A, B \mid\!\sim C$ is not sound, hence, according to the definition, there is a derivation of:

$$A, B, \mathsf{Asm}(A, B \vdash C), \Pi \vdash \bigwedge \Sigma$$

where Π is the set of background assumptions associated with the derivation of $A \mid\!\sim C$ and $\Sigma \in \mathbf{T}$. Hence we get:

$$\dfrac{\dfrac{A \vdash \bigwedge_{i \leq n} B_i \qquad \dfrac{A \vdash B \qquad A, B, \mathsf{Asm}(A, B \vdash C), \Pi \vdash \bigwedge \Sigma}{A, \mathsf{Asm}(A, B \vdash C), \Pi \vdash \bigwedge \Sigma}\ \mathrm{Gcut}}{A, A^n, \mathsf{Asm}(A \vdash C), \Pi \vdash \bigwedge \Sigma}\ \text{several Gcuts}}{A, \mathsf{Asm}(A \vdash C), \Pi \vdash \bigwedge \Sigma}\ \mathrm{LC}$$

Hence we get that $A \mid\!\sim C$ is not sound, which is a contradiction. In other words, monotonicity is acceptable if we restrict ourselves to adding the classical consequences in the antecedent.

7 Conclusions

In this paper, we have introduced a proof-theoretic approach to a logic that deals with exceptions and typicality. This approach combines the method of control sets as presented in Piazza and Pulcini (2017) with the decomposition method outlined in Piazza and Tesi (2024). The outcome is a controlled sequent calculus for first-order logic, augmented with extra-logical axioms.

Our proposal successfully meets four key criteria. First, the introduced sequent systems are cut-free, making them entirely analytic. Second, they exhibit flexibility by accommodating both positive and negative information within a classical deductive framework. Third, these calculi are equipped to handle conflicts between different pieces of information as well as paraconsistency phenomena. Lastly, one of these systems is tailored to address the intricate concept of typicality, which is notoriously challenging to formalize, without resorting to external modal operators. In particular, we provide a proof-theoretic explanation for the non-transitive nature of typicality information.

Numerous potential directions for future research are available. One promising avenue involves enhancing control sets to express counterorders, which would allow them to decrease in size under specific conditions and unblock previously blocked derivations. This could be accomplished by formulating rules that, under specific premises, permit the removal of formulas from the control sets. This development is desirable, as it would facilitate the modeling of cases where the introduction of new information can transform unprovable sequents into provable ones. Another interesting possibility could be to avoid the use of the relativization operator $\mathfrak{R}$. In this context, we use it in order to circumvent the undecidability of first-order logic which would prevent the possibility to perform soundness checks. To eliminate it would be desirable to work with certain decidable fragments of first-order logics such as the guarded or the two variable fragments. Doing so, we would not need to resort to any preprocessing of formulas in the sequents.

Moreover, it would be advantageous to expand the current approach to allow for a local soundness check. In particular, it would be intriguing to devise a technique for monitoring background assumptions as derivations progress. In this regard, it might be beneficial to reintroduce repositories to keep track of the formulas and to manage potential conflicts arising from the information stored in the control sets. However, a challenge arises due to the permissibility of axioms containing free variables, which clashes with the freshness conditions essential for applying the rules R$\forall$ and L$\exists$.

Lastly, it would be valuable to explore the study of substructural controlled sequent calculi that incorporate proper axioms. In particular, investigating controlled systems based on non-commutative sequent calculi could lead to interesting connections with categorial grammars. Specifically, it would be valuable to develop a technique for managing proper axioms in a non-commutative setting and subsequently leverage control sets to articulate prohibitions within a linguistic framework.

Acknowledgements Mario Piazza acknowledges support from the MIUR PRIN Project "Proof and Understanding in Mathematics (PUMa). Purity of methods, simplicity, and explanation in mathematical reasoning", Bando 2022, Protocol Number 2022TYNY32. Matteo Tesi acknowledges support from FWF project I 6372-N "Logical methods for Deontic Explanations" (LoDEx) and from the Project MSCA IEF 101152658 "Structures for modal and deontic logics" (REMODEL). The authors would also like to thank two anonymous reviewers for their careful reading of the paper and their insightful comments.

Funding Open access funding provided by Scuola Normale Superiore within the CRUI-CARE Agreement.

References

Avron, A. (1991). Hypersequents, logical consequence and intermediate logics for concurrency. *Annals of Mathematics and Artificial Intelligence, 4*(3–4), 225–248.

Avron, A. (1996). The method of hypersequents in the proof theory of propositional non-classical logics. *In Logic: from foundations to applications,* pages 1–32. Clarendon Press.
Bhattacharjee, R., & Piazza, M. (2025). Venn diagrams and belief revision. *Draft*.
Boniolo, G., D'Agostino, M., Piazza, M., & Pulcini, G. (2015). Adding logic to the toolbox of molecular biology. *European Journal for Philosophy of Science, 5*(3), 399–417.
Boniolo, G., D'Agostino, M., Piazza, M., & Pulcini, G. (2021). Molecular biology meets logic: Context-sensitiveness in focus. *Foundations of Science,* pages 1–19.
Ciabattoni, A., Galatos, N., & Terui, K. (2008). From axioms to analytic rules in nonclassical logics. *In 23rd Annual IEEE Symposium on Logic in Computer Science,* pages 229–240.
D'Agostino, M., Piazza, M., & Pulcini, G. (2014). A logical calculus for controlled monotonicity. *J. Applied Logic, 12*(4), 558–569.
Gentzen, G. (1935). Untersuchungen über das logische Schliessen. *Math. Zeitschrift, 39*, 176–210.
Giordano, L., Gliozzi, V., Lieto, A., Olivetti, N., & Pozzato, G. L. (2020). Reasoning about typicality and probabilities in preferential description logics. *Applications and Practices in Ontology Design, Extraction, and Reasoning, 49*, 145.
Giordano, L., Gliozzi, V., Olivetti, N., & Pozzato, G. L. (2013). A non-monotonic description logic for reasoning about typicality. *Artificial Intelligence, 195*, 165–202.
Girard, J.-Y. (1991). A new constructive logic: classic logic. *Mathematical Structures in Computer Science, 1*(3), 255–296.
Girard, J.-Y. (1992). Logic and exceptions: a few remarks. *Journal of Logic and Computation, 2*(2), 111–118.
Kraus, S., Lehmann, D., & Magidor, M. (1990). Nonmonotonic reasoning, preferential models and cumulative logics. *Artificial intelligence, 44*(1–2), 167–207.
Krifka, M., Pelletier, F. J., Carlson, G., ter Meulen, A., & A., G. Chierchia, and G. Link. (1995). Genericity: An introduction. In G. Carlson and J. Pelletier (eds.), editors. *The Generic Book,* pages 1–124. The University of Chicago Press.
Makinson, D. (2003). Bridges between classical and nonmonotonic logic. *Logic Journal of the IGPL, 11*(1), 69–96.
Makinson, D. (2005). Bridges from classical to nonmonotonic logic. *King's College,*.
Millson, J., & Straßer, C. (2019). A logic for best explanations. *Journal of Applied Non-Classical Logics, 29*(2), 184–231.
Negri, S., & von Plato, J. (1998). Cut elimination in the presence of axioms. *Bulletin of Symbolic Logic, 4*(4), 418–435.
Ono, H. (2019). Proof Theory and Algebra in Logic. *Springer*.
Piazza, M., & Pulcini, G. (2016). Uniqueness of axiomatic extensions of cut-free classical propositional logic. *Logic Journal of the IGPL, 24*(5), 708–718.
Piazza, M., & Pulcini, G. (2017). Unifying logics via context-sensitiveness. *J. Log. Comput., 27*(1), 21–40.
Piazza, M., Pulcini, G., & Tesi, M. (2023). Linear logic in a refutational setting. *Journal of Logic and Computation, 34*(7), 1249–1273.
Piazza, M., & Sabatini, A. (2025). Hypersequent calculi for propositional default logics. *ACM Transactions on Computational Logic*. https://doi.org/10.1145/3725849
Piazza, M., & Tesi, M. (2024). Analytic calculi with extra-logical information. *Journal of Logic and Computation,* pages 1–31.
Reiter, R. (1978). On reasoning by default. *American Journal of Computational Linguistics,* pages 29–37.
Reiter, R. (1980). A logic for default reasoning. *Artificial Intelligence, 13*, 81–132.
Reiter, R. (1987). Nonmonotonic reasoning. *Annual Review of Computer Science, 2*, 147–186.
Reiter, R., & Criscuolo, G. (1983). Some representational issues in default reasoning. *In Computational Linguistics,* pages 15–27. Elsevier.
Troelstra, A. S., & Schwichtenberg, H. (2000). Basic proof theory, volume 43. *Cambridge University Press*.