The Graded Lambek Calculus

Aubrey Bryant
School of Computer and Cyber Sciences
Augusta University
Augusta, Georgia, USA
aubreylbryant@gmail.com

Harley Eades III
School of Computer and Cyber Sciences
Augusta University
Augusta, Georgia, USA
harley.eades@gmail.com

One gap in the literature on functional programming and software verification is the application of non-commutative logic. This gap is easily seen by the lack of programming languages and proof assistants that incorporate non-commutative reasoning. To help close this gap we investigate the combination of non-commutativity and graded necessity modalities. This marriage offers a surprisingly expressive system capable of encoding a large number of substructural logics in both commutative and non-commutative formalizations. We propose a new graded modal logic called the Graded Lambek Calculus which comes with the ability for the prover/programmer to declare when hypotheses/inputs are allowed to be exchanged, and in what direction they are allowed to travel. We then show that Graded Lambek Calculus can encode the mixture of commutative/non-commutative graded modal logic, left/right-commutative graded modal logic, and a number of other systems.

1 Introduction

Suppose we want to verify the correctness of the `leftpad` program, a recent benchmark for software verification techniques. This is a program that pads a vector to the left with a given symbol. Given an expressive enough system we can intrinsically verify that `leftpad` is correct. For example, the following is the type of `leftpad` in the Granule programming language [7]:

```
leftpad: forall  {t:Type,m n:Nat}.{m >= n} => □_{m-n} t -> N m -> Vec n t -> Vec m t
```

In the above, `n` and `m` are natural numbers called **grades**. These label types using a modal operator called a **graded necessity modality** [2, 7, 1, 2, 5], $\Box_{\mathtt{r}}$, where the grade `r` is a member of a semiring, and it constrains the usage of inputs of type `t`[1]. For example, we can use natural numbers to specify how many times an input can be used, but grades can range over more than just natural numbers. The type `t` is the type of the symbol we are padding the input vector with. The type of `leftpad` states that if we start with a vector of size `n`, and pad a symbol to the left of the vector `(m - n)`-times, then we will end up with a vector of size `m`. This type intrinsically checks nearly every property of `leftpad`, except one. It does not prevent the input vector from being reordered; thus, if we could control the order of data, then we could also intrinsically check that the output vector is in the same order as the input vector, effectively obtaining a fully verified `leftpad` without writing a single proof. However, marrying graded necessity modalities with non-commutativity is an open problem.

In this paper, we propose to combine non-commutativity and graded necessity modalities through the Lambek calculus [6] resulting in a system we call the **Graded Lambek Calculus**. However, just adding graded necessity modalities to the Lambek calculus is not our only contribution; we also show how to control exchange using the graded necessity modality, resulting in a very expressive system.

[1] Graded necessity modalities are a refinement of Girard's [3] of-course modality; in fact, instantiating the semiring with the singleton semiring results in the graded necessity modality being isomorphic to the of-course modality.

Submitted to:
LINEARITY & TLLA 2020

Graded logics and type systems are parameterized by a mathematical structure that captures the usage of data within the system. In the most basic of systems this structure is a monoid, but in the presence of structural rules this structure is a semiring (or a structure closely related). The structural rules weakening, contraction, and exchange are traditionally defined as follows:

$$\frac{\Gamma_1,\Gamma_2 \vdash B}{\Gamma_1,A,\Gamma_2 \vdash B}\,\textbf{weak} \qquad \frac{\Gamma_1,A,A,\Gamma_2 \vdash B}{\Gamma_1,A,\Gamma_2 \vdash B}\,\textbf{contr} \qquad \frac{\Gamma_1,A_1,A_2,\Gamma_2 \vdash B}{\Gamma_1,A_2,A_1,\Gamma_2 \vdash B}\,\textbf{ex}$$

Compare these with the **graded structural rules**. Graded weakening and contraction are as follows:

$$\frac{(\gamma_1,\gamma_3)\odot(\Gamma_1,\Gamma_3)\vdash B \quad \gamma_2\odot\Gamma_2\vdash}{(\gamma_1,0*\gamma_2,\gamma_3)\odot(\Gamma_1,\Gamma_2,\Gamma_3)\vdash B}\,\textbf{weak} \qquad \frac{(\gamma_1,\gamma,\gamma',\gamma_3)\odot(\Gamma_1,\Gamma,\Gamma,\Gamma_3)\vdash B}{(\gamma_1,\gamma+\gamma',\gamma_3)\odot(\Gamma_1,\Gamma,\Gamma_3)\vdash B}\,\textbf{contr}$$

Weakening states that an unused hypothesis is marked with a 0, and if a hypothesis A is graded with a grade of r_1 and r_2, then the usage of A should be recorded as r_1+r_2, and thus, contraction corresponds to the addition of the parameterized semiring. Now exchange is a bit more subtle.

We want to control the exchange of hypotheses using the grades, but this means that we must extend our notion of semiring with a means of indicating which grades are allowed to be exchanged, and which are not. For this, we introduce the **exchange tag** on grades, $e : \mathscr{R} \longrightarrow \mathscr{R}_\perp$, where $(\mathscr{R},1,*,0,+,\leq)$ is a partially-ordered semiring and $\mathscr{R}$ is the set of grades. The exchange tag can then be used to mark grades as exchangable resulting in the following rules:

$$\frac{(\gamma_1,e(r_1),r_2,\gamma_2)\odot(\Gamma_1,A_1,A_2,\Gamma_2)\vdash B}{(\gamma_1,r_2,e(r_1),\gamma_2)\odot(\Gamma_1,A_2,A_1,\Gamma_2)\vdash B}\,\mathrm{ex}_\Rightarrow \qquad \frac{(\gamma_1,r_1,e(r_2),\gamma_2)\odot(\Gamma_1,A_1,A_2,\Gamma_2)\vdash B}{(\gamma_1,e(r_2),r_1,\gamma_2)\odot(\Gamma_1,A_2,A_1,\Gamma_2)\vdash B}\,\mathrm{ex}_\Leftarrow$$

These are the rules proposed by de Paiva and Eades [8], but our system also supports the exchange rule where both grades have to be marked as exchangeable, which naturally arises from an adjoint model between non-commutative linear logic and linear logic proposed by Jiang et. al [4]. We call the system with the exchange tag and the above exchange rules the **Graded Lambek Calculus with biexchange**, but this design hints at a generalization.

The exchange tag is actually induced by an **exchange labeling** $\varepsilon : \mathscr{R} \longrightarrow \{\top,\perp\}$ where $e(r) = r$ when $\varepsilon(r)$ is defined. Now this can be extended to a labeling $\varepsilon : \mathscr{R} \longrightarrow \{\mathsf{left},\mathsf{right},\mathsf{both},\perp\}$ where $e(r) = r$ when $\varepsilon(r) \in \{\mathsf{left},\mathsf{right},\mathsf{both}\}$. Using this new exchange tag we can not only decide when a graded hypotheses is exchangeable, but in which direction they are allowed to travel in the context using the following more general exchange rules:

$$\frac{(\gamma_1,r,s,\gamma_2)\odot(\Gamma_1,A,B,\Gamma_2)\vdash C \quad \varepsilon(r)\in\{\mathsf{right},\mathsf{both}\}}{(\gamma_1,s,r,\gamma_2)\odot(\Gamma_1,B,A,\Gamma_2)\vdash C}\,\mathbf{ex}_\Rightarrow \qquad \frac{(\gamma_1,r,s,\gamma_2)\odot(\Gamma_1,A,B,\Gamma_2)\vdash C \quad \varepsilon(s)\in\{\mathsf{left},\mathsf{both}\}}{(\gamma_1,s,r,\gamma_2)\odot(\Gamma_1,B,A,\Gamma_2)\vdash C}\,\mathbf{ex}_\Leftarrow$$

This more general system we call the **Graded Lambek Calculus**. Using this system we can mix several notions of commutativity like commutative linear logic, non-commutative linear logic, left commutative linear logic, and right commutative linear logic. The labeling and induced exchange tag is similar to Pruiksma et al.'s [9] modes in Adjoint Logic.

The semiring $(\mathbb{N}\sqcup\mathbb{N},\iota_1(1),*,\iota_1(0),+,\leq,\varepsilon)$ where $\varepsilon(\iota_2(r)) = \mathsf{both}$, but $\varepsilon(\iota_1(r)) = \perp$, combines non-commutative graded linear logic with commutative graded linear logic. However, the semiring $(\mathbb{N},1,*,0,+,\leq,\varepsilon)$ where $\varepsilon(r) = \mathsf{right}$ is right commutative graded linear logic. Intuitionistic linear logic can be obtained from the semiring $(\{\infty\},\infty,(\lambda x.\lambda y.\infty),\infty,(\lambda x.\lambda y.\infty),(\lambda x.\lambda y.\top),(\lambda x.\mathsf{both}))$. Similarly, non-commutative intuitionistic affine logic can be obtained from the semiring $(\{0,1,\star\},1,*,0,+,\leq,(\lambda x.\perp))$, where we use $\star$ as an undefined grade that defaults to non-linear logic.

A more interesting example is intervals. Suppose $(\mathscr{R}, 1_\mathscr{R}, *_\mathscr{R}, 0_\mathscr{R}, +_\mathscr{R}, \leq_\mathscr{R})$ is a semiring. Then the set of intervals $\{(r,s) | r \in \mathscr{R}, s \in \mathscr{R}, r \leq_\mathscr{R} s\}$ defines a semiring with operations:

$$(r_1,s_1) + (r_2,s_2) = (r_1 +_\mathscr{R} r_2, s_1 +_\mathscr{R} s_2)$$
$$(r_1,s_1) * (r_2,s_2) = (\sqcap_\mathscr{R}(M), \sqcup_\mathscr{R}(M)), \text{ where } M = \{(r_1 *_\mathscr{R} r_2), (r_1 *_\mathscr{R} s_2), (s_1 *_\mathscr{R} r_2), (s_1 *_\mathscr{R} s_2))\}$$

The functions $\sqcup_\mathscr{R}$ and $\sqcap_\mathscr{R}$ are least and greatest upper bound respectively. The units are $(1_\mathscr{R}, 1_\mathscr{R})$ and $(0_\mathscr{R}, 0_\mathscr{R})$. Taking $\mathscr{R} = \mathbb{N} \cup \{\infty\}$ we can define $!A = \Box_{(0,\infty)}A$, and define affine types to be $\Box_{(0,1)}A$.

Our specific contributions are: (1) a sequent calculus for the Graded Lambek Calculus with biexchange and a proof of cut elimination (Section 2), and (2) a generalization of the sequent calculus into the Graded Lambek Calculus with a new proof of cut elimination, and natural deduction and term assignment formalizations (Section 3, Appendix A).

Acknowledgments. The authors would like to thank Dominic Orchard for useful conversations on earlier drafts of this paper.

2 Graded Lambek Calculus with Biexchange

We begin with the introduction of the Graded Lambek Calculus with biexchange which is the fusion of the Lambek calculus, graded necessary modalities, and the fine grained control of commutativity using graded modalities. Before introducing the rules of the system, we first must introduce the mathematical structure governing graded modalities.

2.1 Semirings

Every hypothesis in the systems we present here will be annotated with a grade to track its usage within proofs. As we define the rules, it becomes necessary to use multiplication to properly calculate the duplication of hypothesis during composition of proofs, e.g. in the cut rule. This suggests that the structure of grades forms at least a monoid $(\mathscr{R}, 1, *)$, where $\mathscr{R}$ is the set of available grades. Then when factoring in the structural rules, as we saw in the introduction, weakened hypotheses need to be annotated with the additive unit, because contraction corresponds to adding grades together. This then implies that the structure of grades must be a semiring $(\mathscr{R}, 1, *, 0, +)$. Finally, placing an ordering on the grades allows for hypotheses to move between grades. We have now arrived at the definition of semirings.

Definition 2.1. *A **semiring** is a tuple $(\mathscr{R}, 1, *, 0, +, \leq)$ consisting of a multiplicative partially-ordered monoid $(\mathscr{R}, *, 1, \leq)$, and an additive partially-ordered monoid $(\mathscr{R}, 0, +, \leq)$, such that the following additional axioms hold:*

- *(Absorption) for any $r \in \mathscr{R}$, $r * 0 = 0 = 0 * r$,*
- *(Left distributivity) for any $r_1, r_2, r_3 \in \mathscr{R}$, $r_1 * (r_2 + r_3) = (r_1 * r_2) + (r_1 * r_3)$, and*
- *(Right distributivity) for any $r_1, r_2, r_3 \in \mathscr{R}$, $(r_2 + r_3) * r_1 = (r_2 * r_1) + (r_3 * r_1)$.*

The structure of a semiring accounts for the structural rules for weakening and contraction, but to also control exchange we extend semirings with the ability to declare grades as exchangable.

Definition 2.2. *A **semiring with bidirectional exchange (biexchange)** $(\mathscr{R}, 1, *, 0, +, \leq, \varepsilon)$ is a semiring $(\mathscr{R}, 1, *, 0, +, \leq)$ with a relation $\varepsilon : \mathscr{R} \longrightarrow \{\top, \bot\}$ called the **exchange labeling**. Lastly, every semiring operation must preserve the exchange label when an operand is exchangable, and the following holds:*

(Commutativity) if $\varepsilon(r_1 + r_2)$ or $\varepsilon(r_2 + r_1)$, then $\varepsilon(r_1 + r_2) = \varepsilon(r_2 + r_1)$ and $r_1 + r_2 = r_2 + r_1$.

This induces a partial function called the **exchange tag**, $e : \mathscr{R} \longrightarrow \mathscr{R}_\perp$ *where* $e(r_1) = r_1$ *when* $\varepsilon(r_1)$.

The definition of the exchange label implies the following.

Lemma 2.3 (0 is exchangable)**.** *For any semiring with biexchange* $(\mathscr{R}, 1, *, 0, +, \leq, \varepsilon)$*, if there exists an* $r \in \mathscr{R}$ *such that* $\varepsilon(r)$*, then* $\varepsilon(0)$ *holds.*

Proof. Suppose $\varepsilon(r)$ for some $r \in \mathscr{R}$, then $\varepsilon(r) = \varepsilon(0 * r) = \varepsilon(0)$. □

This result corresponds to the notion that one can introduce an unused hypothesis through weakening at any location in the context.

The exchange tag is used to mark grades as exchangable, and will be used by both the prover and the inference rules of the system. For example, consider the exchange rules:

$$\frac{(\gamma_1, e(r_1), r_2, \gamma_2) \odot (\Gamma_1, A_1, A_2, \Gamma_2) \vdash B}{(\gamma_1, r_2, e(r_1), \gamma_2) \odot (\Gamma_1, A_2, A_1, \Gamma_2) \vdash B}\,\mathrm{ex}_{\Rightarrow} \qquad \frac{(\gamma_1, r_1, e(r_2), \gamma_2) \odot (\Gamma_1, A_1, A_2, \Gamma_2) \vdash B}{(\gamma_1, e(r_2), r_1, \gamma_2) \odot (\Gamma_1, A_2, A_1, \Gamma_2) \vdash B}\,\mathrm{ex}_{\Leftarrow}$$

Here we mark grades as exchangeable in the premises and the conclusions of each rule. These markings imply an implicit side condition that $\varepsilon(r_1)$ and $\varepsilon(r_2)$ must hold, because if they do not hold, then these rules are not usable in proofs, because grade vectors must contain elements of $\mathscr{R}$ and not $\mathscr{R}_\perp$.

2.2 The Sequent Calculus

We now turn to the sequent calculus for the Graded Lambek Calculus with biexchange. The system is parameterized by a semiring with biexchange $(\mathscr{R}, 1, *, 0, +, \leq, \varepsilon)$. The syntax of the system is as follows:

$$\begin{array}{rl}
\text{(Formulas)} & A, B, C, D ::= I \mid A \triangleright B \mid A \rightharpoonup B \mid B \leftharpoonup A \mid \Box_r A \\
\text{(Contexts)} & \Gamma ::= \emptyset \mid A \mid \Gamma_1, \Gamma_2 \\
\text{(Grade Vectors)} & \gamma ::= \emptyset \mid r \mid \gamma_1, \gamma_2 \\
\text{(Judgments)} & J ::= \gamma \odot \Gamma \vdash \mid \gamma \odot \Gamma \vdash A
\end{array}$$

The inference rules can be found in Figure 1. In the Graded Lambek Calculus with biexchange every hypothesis is graded by pairing a context Γ with a grade vector γ where a **graded context**, denoted $\gamma \odot \Gamma$, is a point-wise assignment of each grade in γ with its corresponding hypothesis in Γ. A well-formed graded context, $\gamma \odot \Gamma \vdash$, is one where γ and Γ have equal lengths. The inference rules are designed to calculate the proper usage constraint for each hypotheses as proofs are derived. In fact, each operation and unit of the semiring appear in one or more inference rules.

Graded hypotheses pair hypotheses with their usage information. The type of usage that is being tracked depends on the semiring one instantiates the system with, but in the general case the inference rules of a graded system must calculate what this usage is for each hypothesis. There are lots of different notions of data-usage tracking, but to help us understand the rules, let us suppose the semiring ranges over $(\mathbb{N}, 1, *, 0, +, \leq, \varepsilon)$. Thus, grades stand for how many times hypotheses are used in proofs.

The graded structural rules are a natural place to begin our discussion of the inference rules:

$$\frac{(\gamma_1, \gamma_3) \odot (\Gamma_1, \Gamma_3) \vdash B \quad \gamma_2 \odot \Gamma_2 \vdash}{(\gamma_1, 0 * \gamma_2, \gamma_3) \odot (\Gamma_1, \Gamma_2, \Gamma_3) \vdash B}\,\textbf{weak} \qquad \frac{(\gamma_1, \gamma, \gamma', \gamma_3) \odot (\Gamma_1, \Gamma, \Gamma, \Gamma_3) \vdash B}{(\gamma_1, \gamma + \gamma', \gamma_3) \odot (\Gamma_1, \Gamma, \Gamma_3) \vdash B}\,\textbf{contr}$$

The rule **weak** adds a context, Γ_2, of unused hypotheses into the context of the premise, but their usage constraint, the grade, must indicate that those hypotheses are unused, and thus, we grade them all with a zero; here $0 * \gamma_2$ is scalar multiplication of the grade vector γ_2. We use a vector γ_2 and require $\gamma_2 \odot \Gamma_2 \vdash$ to ensure we add the proper number of zeros to the grade vector in the conclusion. Contraction, the rule

$$\frac{}{\emptyset \odot \emptyset \vdash}\,\textbf{empty} \qquad \frac{\gamma_1 \odot \Gamma_1 \vdash}{(\gamma_1, r) \odot (\Gamma_1, A) \vdash}\,\textbf{ext} \qquad \frac{}{1 \odot A \vdash A}\,\textbf{id} \qquad \frac{}{\emptyset \odot \emptyset \vdash I}\,\mathbf{I_R} \qquad \frac{\gamma_1 \odot \Gamma_1 \vdash A \quad \gamma_2 \odot \Gamma_2 \vdash B}{(\gamma_1, \gamma_2) \odot (\Gamma_1, \Gamma_2) \vdash A \rhd B}\,\rhd_{\mathbf{R}}$$

$$\frac{(\gamma_1, r, r, \gamma_2) \odot (\Gamma_1, A, B, \Gamma_2) \vdash C}{(\gamma_1, r, \gamma_2) \odot (\Gamma_1, A \rhd B, \Gamma_2) \vdash C}\,\rhd_{\mathbf{L}} \qquad \frac{(\gamma, 1) \odot (\Gamma, A) \vdash B}{\gamma \odot \Gamma \vdash A \rightharpoonup B}\,\rightharpoonup_{\mathbf{R}}$$

$$\frac{\gamma_2 \odot \Gamma_2 \vdash A \quad (\gamma_1, r, \gamma_3) \odot (\Gamma_1, B, \Gamma_3) \vdash C}{(\gamma_1, r * \gamma_2, \gamma_3) \odot (\Gamma_1, A \rightharpoonup B, \Gamma_2, \Gamma_3) \vdash C}\,\rightharpoonup_{\mathbf{L}} \qquad \frac{(1, \gamma) \odot (A, \Gamma) \vdash B}{\gamma \odot \Gamma \vdash B \leftharpoonup A}\,\leftharpoonup_{\mathbf{R}}$$

$$\frac{\gamma_2 \odot \Gamma_2 \vdash A \quad (\gamma_1, r, \gamma_3) \odot (\Gamma_1, B, \Gamma_3) \vdash C}{(\gamma_1, r * \gamma_2, r, \gamma_3) \odot (\Gamma_1, \Gamma_2, B \leftharpoonup A, \Gamma_3) \vdash C}\,\leftharpoonup_{\mathbf{L}} \qquad \frac{\gamma \odot \Gamma \vdash A}{r * \gamma \odot \Gamma \vdash \Box_r A}\,\Box_{\mathbf{R}} \qquad \frac{(\gamma_1, r, \gamma_2) \odot (\Gamma_1, A, \Gamma_2) \vdash B}{(\gamma_1, 1, \gamma_2) \odot (\Gamma_1, \Box_r A, \Gamma_2) \vdash B}\,\Box_{\mathbf{L}}$$

$$\frac{(\gamma_1, e(r), s, \gamma_2) \odot (\Gamma_1, A, B, \Gamma_2) \vdash C}{(\gamma_1, s, e(r), \gamma_2) \odot (\Gamma_1, B, A, \Gamma_2) \vdash C}\,\mathbf{ex}_{\Rightarrow} \qquad \frac{(\gamma_1, r, e(s), \gamma_2) \odot (\Gamma_1, A, B, \Gamma_2) \vdash C}{(\gamma_1, e(s), r, \gamma_2) \odot (\Gamma_1, B, A, \Gamma_2) \vdash C}\,\mathbf{ex}_{\Leftarrow}$$

$$\frac{(\gamma_1, \gamma_3) \odot (\Gamma_1, \Gamma_3) \vdash B \quad \gamma_2 \odot \Gamma_2 \vdash}{(\gamma_1, 0 * \gamma_2, \gamma_3) \odot (\Gamma_1, \Gamma_2, \Gamma_3) \vdash B}\,\textbf{weak} \qquad \frac{(\gamma_1, \gamma, \gamma', \gamma_3) \odot (\Gamma_1, \Gamma, \Gamma, \Gamma_3) \vdash B}{(\gamma_1, \gamma + \gamma', \gamma_3) \odot (\Gamma_1, \Gamma, \Gamma_3) \vdash B}\,\textbf{contr}$$

$$\frac{\gamma_2 \odot \Gamma_2 \vdash A \quad (\gamma_1, r, \gamma_3) \odot (\Gamma_1, A, \Gamma_3) \vdash C}{(\gamma_1, r * \gamma_2, \gamma_3) \odot (\Gamma_1, \Gamma_2, \Gamma_3) \vdash C}\,\textbf{cut} \qquad \frac{\gamma_1 \odot \Gamma \vdash A \quad \gamma_1 \leq \gamma_2}{\gamma_2 \odot \Gamma \vdash A}\,\textbf{approx}$$

Figure 1: The Graded Lambek Calculus with biexchange

contr, on the other hand, merges two copies of a context, here Γ, and thus, to account for this merge we must sum their grade vectors; the operation $\gamma + \gamma'$ is vector pointwise addition. Thus, in the setting of the natural numbers we can see that a grade stands for how many times a hypothesis can be used.

These quantities must also be accounted for during composition of proofs. Consider the cut rule:

$$\frac{\gamma_2 \odot \Gamma_2 \vdash A \quad (\gamma_1, r, \gamma_3) \odot (\Gamma_1, A, \Gamma_3) \vdash C}{(\gamma_1, r * \gamma_2, \gamma_3) \odot (\Gamma_1, \Gamma_2, \Gamma_3) \vdash C}\,\textbf{cut}$$

When we cut A by the premise $\gamma_2 \odot \Gamma_2 \vdash A$ the hypotheses in Γ_2 will need to be duplicated to reconstruct each A that is needed to prove C using **contr**. Thus, we must multiple each grade in γ_2 to account for this duplication. Now since we must use multiplication in cut, in order for the axiom rule, **id**, to be the idenity of composition, we grade the only hypothesis mentioned in **id** with the multiplicative identity:

$$\frac{}{1 \odot A \vdash A}\,\textbf{id}$$

This is why we denote the multiplicative unit using a 1 to indicate "linear" even when the unit may not strictly mean linear; e.g., in semirings that allow non-linear usage of hypotheses like the semiring for non-linear intuitionistic logic – the singleton semiring whose operations preserve the point.

The prover makes use of grades through the **graded necessity modalities**, denoted $\Box_r A$, and whose left and right rules are as follows:

$$\frac{(\gamma_1, r, \gamma_2) \odot (\Gamma_1, A, \Gamma_2) \vdash B}{(\gamma_1, 1, \gamma_2) \odot (\Gamma_1, \Box_r A, \Gamma_2) \vdash B}\,\Box_{\mathbf{L}} \qquad \frac{\gamma \odot \Gamma \vdash A}{r * \gamma \odot \Gamma \vdash \Box_r A}\,\Box_{\mathbf{R}}$$

Graded necessity modalities internalize the grades placed on hypotheses. The left rule makes this explicit: the grade r is pushed into the language of the system, and replaced with a 1 explicitly showing the transformation from a non-linear hypothesis to a linear one. The right rule promotes the conclusion A into a graded conclusion, but we must also promote the grades in γ to account for the potential duplication of those hypotheses. Intuitively, this rule can be thought of as getting A ready to be cut against some other graded hypothesis.

Similarly, implication internalizes composition of proofs, and this is made explicit by the left rule:

$$\frac{(\gamma,1)\odot(\Gamma,A)\vdash B}{\gamma\odot\Gamma\vdash A\rightharpoonup B}\;\rightharpoonup_{\mathbf{R}}\qquad\frac{\gamma_2\odot\Gamma_2\vdash A\quad(\gamma_1,r,\gamma_3)\odot(\Gamma_1,B,\Gamma_3)\vdash C}{(\gamma_1,r,r*\gamma_2,\gamma_3)\odot(\Gamma_1,A\rightharpoonup B,\Gamma_2,\Gamma_3)\vdash C}\;\rightharpoonup_{\mathbf{L}}$$

In the second premise we can see that B may be used r-times in the proof of C, and so, when we discharge A onto B, essentially hiding a cut, we must account for the number of times we will need to reprove A, and we do this by multiplying the grades in its context by r. The rules for left-implication are similar.

Our system takes the left and right exchange rules as primitive:

$$\frac{(\gamma_1,e(r),s,\gamma_2)\odot(\Gamma_1,A,B,\Gamma_2)\vdash C}{(\gamma_1,s,e(r),\gamma_2)\odot(\Gamma_1,B,A,\Gamma_2)\vdash C}\;\mathbf{ex}_{\Rightarrow}\qquad\frac{(\gamma_1,r,e(s),\gamma_2)\odot(\Gamma_1,A,B,\Gamma_2)\vdash C}{(\gamma_1,e(s),r,\gamma_2)\odot(\Gamma_1,B,A,\Gamma_2)\vdash C}\;\mathbf{ex}_{\Leftarrow}$$

Using these rules we can derive biexchange:

$$\frac{(\gamma_1,e(r),e(s),\gamma_2)\odot(\Gamma_1,A,B,\Gamma_2)\vdash C}{(\gamma_1,e(s),e(r),\gamma_2)\odot(\Gamma_1,B,A,\Gamma_2)\vdash C}\;\mathbf{ex}_{\Leftrightarrow}$$

In addition, taking left and right exchange as primitive suggests an extension where we can control the presence of each rule individually; see Section 3. However, we can also take $\mathbf{ex}_{\Leftrightarrow}$ as primitive, but it does not lead to the extensions we have in mind. The remaining rules of the system can now be understood using the intuition given here. One rule the astute reader may have realized is missing is the left rule for the identity proposition:

$$\frac{\gamma_1,\gamma_2\odot\Gamma_1,\Gamma_2\vdash A}{\gamma_1,0,\gamma_2\odot\Gamma_1,I,\Gamma_2\vdash A}\;\mathbf{I_L}$$

In the above rule, we are forced to give I a grade of 0. If we grade I with anything else, cut elimination will fail. Thus, this rule is a special case of the weakening rule.

We conclude this section with an example proof of $(\Box_{e(2)}A\rhd\Box_1 B)\rightharpoonup A\rhd B\rhd A$:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{1\odot A\vdash A}\;\mathbf{id}\qquad\dfrac{\dfrac{}{1\odot B\vdash B}\;\mathbf{id}\qquad\dfrac{}{e(1)\odot A\vdash A}\;\mathbf{id}}{(1,e(1))\odot(B,A)\vdash B\rhd A}\;\rhd_{\mathbf{R}}}{(1,1,e(1))\odot(A,B,A)\vdash A\rhd B\rhd A}\;\rhd_{\mathbf{R}}}{(1,e(1),1)\odot(A,A,B)\vdash A\rhd B\rhd A}\;\mathbf{ex}_{\Leftarrow}}{(1+e(1),1)\odot(A,B)\vdash A\rhd B\rhd A}\;\mathbf{contr}}{(e(2),1)\odot(A,B)\vdash A\rhd B\rhd A}}{(e(2),1)\odot(A,\Box_1 B)\vdash A\rhd B\rhd A}\;\Box_{\mathbf{L}}}{(1,1)\odot(\Box_{e(2)}A,\Box_1 B)\vdash A\rhd B\rhd A}\;\Box_{\mathbf{L}}}{1\odot(\Box_{e(2)}A\rhd\Box_1 B)\vdash A\rhd B\rhd A}\;\rhd_{\mathbf{L}}}{\emptyset\odot\emptyset\vdash(\Box_{e(2)}A\rhd\Box_1 B)\rightharpoonup A\rhd B\rhd A}\;\rightharpoonup_{\mathbf{R}}$$

The above example shows that when we have $e(r)$ in the grade vector, we are allowed to treat it as just r, but with a label. Thus, we can use the axiom rule when we have $e(1)$, because $e(1) = 1$ in this example. We also see that the grade vectors are an implementation detail and the prover uses the graded necessity modality to specify the data-usage constraints for each hypothesis. If no modality is used, then the system defaults to linear usage.

2.3 Proof Theory of the Sequent Calculus

We have proven several properties of the Graded Lambek calculus with biexchange. We proved cut elimination for Graded Lambek calculus with biexchange with left and right exchange, but also for the system where these rules were replaced with biexchange. The cut elimination property is the same as the cut elimination proof for non-linear intuitionistic logic, but where the grade vectors must be equivalent across proof transformations.

Theorem 2.4 (Cut Elimination)**.** *If Π_1 is a proof of $\gamma_1 \odot \Gamma_1 \vdash A$, and Π_1 steps to a proof Π_2 of $\gamma_2 \odot \Gamma_2 \vdash A$ using the cut-elimination procedure, then $\gamma_1 = \gamma_2$.*

An interesting implication of the proof of cut elimination is that precisely tracking the usage of each hypothesis results in the realization that cut elimination fails for the typical rule for contraction:

$$\frac{(\gamma_1, r_1, r_2, \gamma_3) \odot (\Gamma_1, A, A, \Gamma_3) \vdash B}{(\gamma_1, r_1 + r_2, \gamma_3) \odot (\Gamma_1, A, \Gamma_3) \vdash B}\ \textbf{contr}$$

Cutting against this rule on the right results in a stuck proof, because when we commute the cut up the derivation we have to cut against two copies of the same context, but this rule is not general enough to handle this, because the order of hypotheses must be respected.

We also prove that every judgment implies a well-formed graded context.

Lemma 2.5 (Well-Formed Contexts)**.** *If $\gamma \odot \Gamma \vdash A$, then $\gamma \odot \Gamma \vdash$.*

This shows that we can think of γ as an additional constraint on Γ, and their structural relationship is maintained during proof construction. We proved several auxiliary results required by the above two main results, but we omit them due to space. Lastly, we can define natural deduction and term assignment formalizations for the Graded Lambek Calculus with biexchange, but instead of presenting those systems here, we present a more general set of systems.

3 The Graded Lambek Calculus

What we will call the **Graded Lambek Calculus** is a more general system than Graded Lambek Calculus with biexchange where left and right exchange can be precisely controlled through the graded necessity modality resulting in Graded Lambek Calculus supporting not just the mixture of non-commutative linear logic with commutative linear logic, but the mixture of these two in addition to left and right commutative linear logic. Definition 2.2 can be generalized into the following.

Definition 3.1. *A **semiring with exchange** $(\mathscr{R}, 1, *, 0, +, \leq, \varepsilon)$ is a semiring $(\mathscr{R}, 1, *, 0, +, \leq)$ with a relation $\varepsilon : \mathscr{R} \longrightarrow \{\mathsf{left}, \mathsf{right}, \mathsf{both}, \perp\}$ called the **exchange labeling**. Every semiring operation must preserve the exchange label when an operand is exchangable, and the following holds:*

(Commutativity) if $\varepsilon(r_1 + r_2), \varepsilon(r_2 + r_1) \in \{\mathsf{left}, \mathsf{right}, \mathsf{both}\}$, then $r_1 + r_2 = r_2 + r_1$.

The labeling induces the **exchange tag**, $e : \mathscr{R} \longrightarrow \mathscr{R}_{\perp}$ *where* $e(r_1) = r_1$ *when* $\varepsilon(r_1) \in \{\mathsf{left}, \mathsf{right}, \mathsf{both}\}$. *We will denote* $e(r_1)$ *by* $\mathsf{left}\,(r_1)$ *when* $\varepsilon(r_1) = \mathsf{left}$, *and similarly for* right *and* both.

Using this generalization we can now define the general system.

Definition 3.2. ***The Graded Lambek Calculus*** *consists of the same rules as the ones in Figure 1, but where the exchange rules are replaced by the following exchange rules:*

$$\frac{(\gamma_1, r, s, \gamma_2) \odot (\Gamma_1, A, B, \Gamma_2) \vdash C \quad \varepsilon(s) \in \{\mathsf{left}, \mathsf{both}\}}{(\gamma_1, s, r, \gamma_2) \odot (\Gamma_1, B, A, \Gamma_2) \vdash C}\ \boldsymbol{ex}_{\Leftarrow} \qquad \frac{(\gamma_1, r, s, \gamma_2) \odot (\Gamma_1, A, B, \Gamma_2) \vdash C \quad \varepsilon(r) \in \{\mathsf{right}, \mathsf{both}\}}{(\gamma_1, s, r, \gamma_2) \odot (\Gamma_1, B, A, \Gamma_2) \vdash C}\ \boldsymbol{ex}_{\Rightarrow}$$

The prover will then use the partial functions left, right, or both to mark grades with the proper direction of commutativity. For example, consider the refined proof of $(\Box_{\mathsf{left}\,(2)} A \triangleright \Box_1 B) \rightharpoonup A \triangleright B \triangleright A$:

$$
\begin{prooftree}
\AxiomC{}
\RightLabel{$\mathbf{id}$}
\UnaryInfC{$1 \odot A \vdash A$}
\AxiomC{}
\RightLabel{$\mathbf{id}$}
\UnaryInfC{$1 \odot B \vdash B$}
\AxiomC{}
\RightLabel{$\mathbf{id}$}
\UnaryInfC{$\mathsf{left}\,(1) \odot A \vdash A$}
\BinaryInfC{$(1, \mathsf{left}\,(1)) \odot (B, A) \vdash B \triangleright A$}
\RightLabel{$\triangleright_{\mathbf{R}}$}
\BinaryInfC{$(1, 1, \mathsf{left}\,(1)) \odot (A, B, A) \vdash A \triangleright B \triangleright A$}
\RightLabel{$\mathbf{ex}_{\Leftarrow}$}
\UnaryInfC{$(1, \mathsf{left}\,(1), 1) \odot (A, A, B) \vdash A \triangleright B \triangleright A$}
\RightLabel{$\mathbf{contr}$}
\UnaryInfC{$(1 + \mathsf{left}\,(1), 1) \odot (A, B) \vdash A \triangleright B \triangleright A$}
\doubleLine
\UnaryInfC{$(\mathsf{left}\,(2), 1) \odot (A, B) \vdash A \triangleright B \triangleright A$}
\RightLabel{$\Box_{\mathbf{L}}$}
\UnaryInfC{$(\mathsf{left}\,(2), 1) \odot (A, \Box_1 B) \vdash A \triangleright B \triangleright A$}
\RightLabel{$\Box_{\mathbf{L}}$}
\UnaryInfC{$(1, 1) \odot (\Box_{\mathsf{left}\,(2)} A, \Box_1 B) \vdash A \triangleright B \triangleright A$}
\RightLabel{$\triangleright_{\mathbf{L}}$}
\UnaryInfC{$1 \odot (\Box_{\mathsf{left}\,(2)} A \triangleright \Box_1 B) \vdash A \triangleright B \triangleright A$}
\RightLabel{$\rightharpoonup_{\mathbf{R}}$}
\UnaryInfC{$\emptyset \odot \emptyset \vdash (\Box_{\mathsf{left}\,(2)} A \triangleright \Box_1 B) \rightharpoonup A \triangleright B \triangleright A$}
\end{prooftree}
$$

Here we mark the hypothesis A as left commutative allowing for the use of the rule $\mathbf{ex}_{\Leftarrow}$.

Just as we did for the Graded Lambek Calculus with biexchange we proved cut-elimination for this system, and developed natural deduction and term assignment formalizations (Appendix A).

4 Conclusion and Future Work

We started this paper with an example of intrinsically verifying the correctness of `leftpad`. Using non-commutative tensor products and sum types we can model lists by $\mathsf{List} A := I + A \triangleright \mathsf{List} A$. Now that we have shown that it is possible to combine non-commutativity with graded necessity modalities we can extend the Granule programming language [7] – a new language based on linear logic and graded modalities being studied and developed by the second author and Dominic Orchard – to support the above type for lists, resulting in the full correctness of `leftpad`.

We have shown that graded necessity modalities can be extended to support fine-grained control of commutativity using a simple labeling of the grades on hypotheses. These results open the door for further study into the application of non-commutativity and left/right-commutativity within functional programming and software verification, an area of which is lacking exploration within the research community. To help close this gap we plan to implement this system as an extension of the Granule programming language. Using this new implementation we will study applications of non-commutativity in the area of software verification and functional programming.

References

[1] Aloïs Brunel, Marco Gaboardi, Damiano Mazza & Steve Zdancewic (2014): **A Core Quantitative Coeffect Calculus**. In Zhong Shao, editor: *Programming Languages and Systems*, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 351–370.

[2] Marco Gaboardi, Shin-ya Katsumata, Dominic A Orchard, Flavien Breuvart & Tarmo Uustalu (2016): **Combining effects and coeffects via grading.** In: *ICFP*, pp. 476–489.

[3] Jean-Yves Girard (1987): **Linear logic**. *Theoretical Computer Science* 50(1), pp. 1 – 101.

[4] Jiaming Jiang, Harley Eades & Valeria de Paiva (2019): **On the Lambek Calculus with an Exchange Modality**. *Electronic Proceedings in Theoretical Computer Science* 292, pp. 43–89.

[5] Shin-ya Katsumata (2018): **A Double Category Theoretic Analysis of Graded Linear Exponential Comonads**. In Christel Baier & Ugo Dal Lago, editors: *Foundations of Software Science and Computation Structures*, Springer International Publishing, Cham, pp. 110–127.

[6] Joachim Lambek (1958): **The Mathematics of Sentence Structure**. *The American Mathematical Monthly* 65(3), pp. 154–170. Available at `http://www.jstor.org/stable/2310058`.

[7] Domnic Orchard, Vilem-Benjamin Liepelt & Harley Eades III (2019): **Quantitative program reasoning with graded modal types**. *Proc. ACM Program. Lang. 3, ICFP, Article 110*, p. 30.

[8] Valeria de Paiva & Harley Eades (2018): **Dialectica Categories for the Lambek Calculus**. In Sergei Artemov & Anil Nerode, editors: *Logical Foundations of Computer Science*, Springer International Publishing, Cham, pp. 256–272.

[9] Klaas Pruiksma, William Chargin, Frank Pfenning & Jason Reed (2018): **Adjoint Logic**. Unpublished Draft: http://www.cs.cmu.edu/ fp/papers/adjoint18b.pdf.

Appendix

A Natural Deduction and Term Assignment Formalizations

$$\frac{}{1 \odot A \vdash A}\,\mathbf{id} \qquad \frac{}{\emptyset \odot \emptyset \vdash I}\,\mathbf{I_I} \qquad \frac{(\gamma_1, r, s, \gamma_2) \odot (\Gamma_1, A, B, \Gamma_2) \vdash C \quad \varepsilon(s) \in \{\mathsf{left}, \mathsf{both}\}}{(\gamma_1, s, r, \gamma_2) \odot (\Gamma_1, B, A, \Gamma_2) \vdash C}\,\mathbf{ex}_{\Leftarrow}$$

$$\frac{(\gamma_1, r, s, \gamma_2) \odot (\Gamma_1, A, B, \Gamma_2) \vdash C \quad \varepsilon(r) \in \{\mathsf{right}, \mathsf{both}\}}{(\gamma_1, s, r, \gamma_2) \odot (\Gamma_1, B, A, \Gamma_2) \vdash C}\,\mathbf{ex}_{\Rightarrow} \qquad \frac{\gamma_1 \odot \Gamma_1 \vdash A \quad \gamma_2 \odot \Gamma_2 \vdash B}{(\gamma_1, \gamma_2) \odot (\Gamma_1, \Gamma_2) \vdash A \rhd B}\,\rhd_{\mathbf{I}}$$

$$\frac{\gamma_2 \odot \Gamma_2 \vdash A \rhd B \quad (\gamma_1, r, r, \gamma_3) \odot (\Gamma_1, A, B, \Gamma_3) \vdash C}{(\gamma_1, r * \gamma_2, \gamma_3) \odot (\Gamma_1, \Gamma_2, \Gamma_3) \vdash C}\,\rhd_{\mathbf{E}} \qquad \frac{(\gamma, 1) \odot (\Gamma, A) \vdash B}{\gamma \odot \Gamma \vdash A \rightharpoonup B}\,\rightharpoonup_{\mathbf{I}}$$

$$\frac{\gamma_1 \odot \Gamma_1 \vdash A \rightharpoonup B \quad \gamma_2 \odot \Gamma_2 \vdash A}{(\gamma_1, \gamma_2) \odot (\Gamma_1, \Gamma_2) \vdash B}\,\rightharpoonup_{\mathbf{E}} \qquad \frac{(1, \gamma) \odot (A, \Gamma) \vdash B}{\gamma \odot \Gamma \vdash B \leftharpoonup A}\,\leftharpoonup_{\mathbf{I}}$$

$$\frac{\gamma_1 \odot \Gamma_1 \vdash A \quad \gamma_2 \odot \Gamma_2 \vdash B \leftharpoonup A}{(\gamma_1, \gamma_2) \odot (\Gamma_1, \Gamma_2) \vdash B}\,\leftharpoonup_{\mathbf{E}} \qquad \frac{\gamma \odot \Gamma \vdash A}{r * \gamma \odot \Gamma \vdash \Box_r A}\,\Box_{\mathbf{I}}$$

$$\frac{\gamma_2 \odot \Gamma_2 \vdash \Box_r A \quad (\gamma_1, r, \gamma_3) \odot (\Gamma_1, A, \Gamma_3) \vdash B}{(\gamma_1, \gamma_2, \gamma_3) \odot (\Gamma_1, \Gamma_2, \Gamma_3) \vdash B}\,\Box_{\mathbf{E}} \qquad \frac{(\gamma_1, \gamma_3) \odot (\Gamma_1, \Gamma_3) \vdash B \quad \gamma_2 \odot \Gamma_2 \vdash}{(\gamma_1, 0 * \gamma_2, \gamma_3) \odot (\Gamma_1, \Gamma_2, \Gamma_3) \vdash B}\,\mathbf{weak}$$

$$\frac{(\gamma_1, \gamma, \gamma', \gamma_3) \odot (\Gamma_1, \Gamma, \Gamma, \Gamma_3) \vdash B}{(\gamma_1, \gamma + \gamma', \gamma_3) \odot (\Gamma_1, \Gamma, \Gamma_3) \vdash B}\,\mathbf{contr} \qquad \frac{\gamma_1 \odot \Gamma \vdash A \quad \gamma_1 \leq \gamma_2}{\gamma_2 \odot \Gamma \vdash A}\,\mathbf{approx}$$

Figure 2: Graded Lambek Calculus : Natural Deduction

$$\frac{}{1 \odot x : A \vdash x : A}\ \mathbf{id} \qquad \frac{}{\emptyset \odot \emptyset \vdash \mathsf{unit} : I}\ \mathbf{I_I}$$

$$\frac{(\gamma_1, r, s, \gamma_2) \odot (\Gamma_1, x : A, y : B, \Gamma_2) \vdash t : C \quad \varepsilon(s) \in \{\mathsf{left}, \mathsf{both}\}}{(\gamma_1, s, r, \gamma_2) \odot (\Gamma_1, y : B, x : A, \Gamma_2) \vdash t : C}\ \mathbf{ex}_{\Leftarrow}$$

$$\frac{(\gamma_1, r, s, \gamma_2) \odot (\Gamma_1, x : A, y : B, \Gamma_2) \vdash t : C \quad \varepsilon(r) \in \{\mathsf{right}, \mathsf{both}\}}{(\gamma_1, s, r, \gamma_2) \odot (\Gamma_1, y : B, x : A, \Gamma_2) \vdash t : C}\ \mathbf{ex}_{\Rightarrow}$$

$$\frac{\gamma_1 \odot \Gamma_1 \vdash t_1 : A \quad \gamma_2 \odot \Gamma_2 \vdash t_2 : B}{(\gamma_1, \gamma_2) \odot (\Gamma_1, \Gamma_2) \vdash (t_1, t_2) : A \triangleright B}\ \triangleright_{\mathbf{I}}$$

$$\frac{\gamma_2 \odot \Gamma_2 \vdash t_1 : A \triangleright B \quad (\gamma_1, r, r, \gamma_3) \odot (\Gamma_1, x : A, y : B, \Gamma_3) \vdash t_2 : C}{(\gamma_1, r * \gamma_2, \gamma_3) \odot (\Gamma_1, \Gamma_2, \Gamma_3) \vdash \mathsf{let}\, x \triangleright y = t_1\, \mathsf{in}\, t_2 : C}\ \triangleright_{\mathbf{E}} \qquad \frac{(\gamma, 1) \odot (\Gamma, x : A) \vdash t : B}{\gamma \odot \Gamma \vdash \lambda_r x.t : A \rightharpoonup B}\ \rightharpoonup_{\mathbf{I}}$$

$$\frac{\gamma_1 \odot \Gamma_1 \vdash t_1 : A \rightharpoonup B \quad \gamma_2 \odot \Gamma_2 \vdash t_2 : A}{(\gamma_1, \gamma_2) \odot (\Gamma_1, \Gamma_2) \vdash \mathsf{app}_r\, t_1\, t_2 : B}\ \rightharpoonup_{\mathbf{E}} \qquad \frac{(1, \gamma) \odot (x : A, \Gamma) \vdash t : B}{\gamma \odot \Gamma \vdash \lambda_l x.t : B \leftharpoonup A}\ \leftharpoonup_{\mathbf{I}}$$

$$\frac{\gamma_1 \odot \Gamma_1 \vdash t_1 : A \quad \gamma_2 \odot \Gamma_2 \vdash t_2 : B \leftharpoonup A}{(\gamma_1, \gamma_2) \odot (\Gamma_1, \Gamma_2) \vdash \mathsf{app}_l\, t_1\, t_2 : B}\ \leftharpoonup_{\mathbf{E}} \qquad \frac{\gamma \odot \Gamma \vdash t : A}{r * \gamma \odot \Gamma \vdash \Box t : \Box_r A}\ \Box_{\mathbf{I}}$$

$$\frac{\gamma_2 \odot \Gamma_2 \vdash t_1 : \Box_r A \quad (\gamma_1, r, \gamma_3) \odot (\Gamma_1, x : A, \Gamma_3) \vdash t_2 : B}{(\gamma_1, \gamma_2, \gamma_3) \odot (\Gamma_1, \Gamma_2, \Gamma_3) \vdash \mathsf{let}\, \Box x = t_1\, \mathsf{in}\, t_2 : B}\ \Box_{\mathbf{E}}$$

$$\frac{(\gamma_1, \gamma_3) \odot (\Gamma_1, \Gamma_3) \vdash t : B \quad \gamma_2 \odot \Gamma_2 \vdash}{(\gamma_1, 0 * \gamma_2, \gamma_3) \odot (\Gamma_1, \Gamma_2, \Gamma_3) \vdash t : B}\ \mathbf{weak} \qquad \frac{(\gamma_1, \gamma, \gamma', \gamma_3) \odot (\Gamma_1, \Gamma, \Gamma, \Gamma_3) \vdash t : B}{(\gamma_1, \gamma + \gamma', \gamma_3) \odot (\Gamma_1, \Gamma, \Gamma_3) \vdash t : B}\ \mathbf{contr}$$

$$\frac{\gamma_1 \odot \Gamma \vdash t : A \quad \gamma_1 \leq \gamma_2}{\gamma_2 \odot \Gamma \vdash t : A}\ \mathbf{approx}$$

Figure 3: Graded Lambek Calculus : Term Assignment