

Syntax and Models of a non-Associative Composition of Programs and Proofs

Guillaume MUNCH-MACCAGNONI



December 10th 2013

Introduction

●○○○
○○○○○
○○○

The “L” approach

○○○○○
○○○○
○○○
○○○

Duploids

○○○○○
○○○○○○○

Applications

○○○
○○○○○
○○○

Perspectives

○○○





Metaphor

Let us consider a question for a human

$$(1+2) \times (3+4)$$

$$= (1+2) \times (3+4)$$

$$= 3 \times (3+4)$$

$$= 3 \times (3+4)$$

$$= 3 \times 7$$

$$= 3 \times 7$$

$$= 21$$





Metaphor

The same question asked to a machine

Reverse polish notation:

$$12+34+x$$

	Expression	Stack	
	1 2 + 3 4 + x		push
▷	2 + 3 4 + x	1	push
▷	+ 3 4 + x	2, 1	add
▷	3 4 + x	3	push
▷	4 + x	3, 3	push
▷	+ x	4, 3, 3	add
▷	x	7, 3	mult
▷		21	



Metaphor

Moral

There may be many equivalent representations but some are better than others.



Metaphor

In this thesis

Language \rightarrow Representation \rightarrow Model

Language of Proofs, of Programs

Model Mathematical (Denotation), Physical (Implementation)

Representation Two formalisms of which we show the unity:
 Sequent calculus (after **Gentzen**), Continuation-based
 models (after **Landin**).

The polarisation hypothesis

Introducing the idea of interaction

We take into account the **context** of a program or proof. Below: the context is modelled with a stack.

	Expression	Stack	
	match $(\lambda x. \iota_1(t)) u$ with $(v w)$	*	push
▷	$(\lambda x. \iota_1(t)) u$	$(v w) \cdot *$	push
▷	$\lambda x. \iota_1(t)$	$u \cdot (v w) \cdot *$	pop
▷	$\iota_1(t[u/x])$	$(v w) \cdot *$	branch
▷	v	$t[u/x] \cdot *$	



The polarisation hypothesis

Introducing the idea of interaction

Interactive aspects of the Curry-Howard correspondence

- In the proof theory of classical logic
- In the continuation-based modelling of evaluation order and control
- Categorical modelling of effects



The polarisation hypothesis

Introducing polarities

Polarities are a legacy of intuitionism:

- Positive connectives: $\exists, \vee \dots$
- Negative connectives: $\forall, \rightarrow \dots$

They become more important when interaction is taken into account.



The polarisation hypothesis

Introducing polarities

1. Classical proofs: polarities determine cut elimination (constructiveness)
2. Continuation-based modelling of programs: polarities determine whether a continuation is meant to be passed or to be applied (evaluation order)
3. Categories: Polarisation amounts to relaxing the associativity of composition (“Blass” phenomenon)



The polarisation hypothesis

Polarisation-related phenomena

- Blass problem in game semantics (*Abramsky, Melliès*)
- Thunks in call-by-value (*Hatcliff and Danvy, Führmann*)
- Value-restriction of polymorphism in CBV (*Harper and Lillibridge*)
- Arithmetical hierarchy in Peano arithmetic
- Focalisation in proof search (*Andreoli*)
- Orthogonality-based logical relations (*Girard, Krivine, Pitts*)

And probably more

Contribution

What

We contribute to the understanding of the nature, role and mechanisms of polarisation with:

1. a positive characterisation of polarisation where composition is not associative: **duploids**;
2. a decomposition of CPS translations in three meaningful steps (for *shift/reset* operators and 4 call-by-name variants);
3. a formulae-as-types correspondence for a Curry-style classical natural deduction that satisfies $\neg\neg A \simeq A$

Contribution

How

Introducing **L** calculi, for:

- modelling programming languages (after **Landin**)
- investigating the structure of proofs (after **Gentzen**)

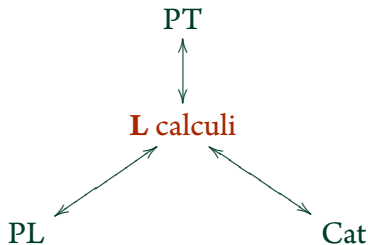
Inspired from:

- the modelling of effects (after **Moggi**)
- the quest for a link between categorical duality and continuations (after **Filinski**)
- the interactive notion of construction (after **Girard, Krivine**)

Contribution

Why

Exhibiting **unifying principles** behind an extension to interaction of the correspondence between proof theory, programming languages and category theory.





Four ideas for modelling interaction

1. Explicit contexts, yielding abstract machines.
2. Solving equations with adjoints, yielding a correspondence with sequent calculus.
3. Introducing the $\tilde{\mu}$ binder, uncovering a symmetry with simplifying virtues.
4. Considering polarisation, yielding extensionality.

Abstract machines

First ingredient: Explicit contexts

The same example as before:

$$\langle \text{match } (\lambda x. \iota_1(t)) u \text{ with } (v|w) \parallel e \rangle$$

$$\triangleright_{\mathbf{R}} \langle (\lambda x. \iota_1(t)) u \parallel (v|w) \cdot e \rangle$$

$$\triangleright_{\mathbf{R}} \langle \lambda x. \iota_1(t) \parallel u \cdot (v|w) \cdot e \rangle$$

$$\triangleright_{\mathbf{R}} \langle \iota_1(t[u/x]) \parallel (v|w) \cdot e \rangle$$

$$\triangleright_{\mathbf{R}} \langle v \parallel t[u/x] \cdot e \rangle$$

Commands $c ::= \langle t \parallel e \rangle$

Contexts $e ::= \star \mid \text{fst} \cdot e \mid \text{snd} \cdot e \mid (u|v) \cdot e \mid u \cdot e$

(for now contexts are stacks)

Abstract machines

Second ingredient: Solving equations with adjoints

Elimination rules:

$$\langle \text{fst}(u) \parallel e \rangle \triangleright_{\mathbf{R}} \langle u \parallel \text{fst} \cdot e \rangle$$

$$\langle \text{snd}(u) \parallel e \rangle \triangleright_{\mathbf{R}} \langle u \parallel \text{snd} \cdot e \rangle$$

$$\langle \text{match } t \text{ with } (u|v) \parallel e \rangle \triangleright_{\mathbf{R}} \langle t \parallel (u|v) \cdot e \rangle$$

$$\langle t \ u \parallel e \rangle \triangleright_{\mathbf{R}} \langle t \parallel u \cdot e \rangle$$

are all “**adjoints**” (in the terminology borrowed by Girard):

$$\langle \tau^*(t) \parallel e \rangle \triangleright_{\mathbf{R}} \langle t \parallel \tau(e) \rangle$$



Abstract machines

Second ingredient: Solving equations with adjoints

We introduce the adjoint construct c^* with the following reduction rule:

$$\langle c^* \parallel e \rangle \triangleright_{\mathbf{R}} c[e/\star]$$

(essentially Parigot's μ)

The adjoint is used to **solve** the reduction rules which are therefore derived:

$$\text{fst}(t) \stackrel{\text{def}}{=} \langle t \parallel \text{fst} \cdot \star \rangle^*$$

$$\text{snd}(t) \stackrel{\text{def}}{=} \langle t \parallel \text{snd} \cdot \star \rangle^*$$

$$\text{match } t \text{ with } (u|v) \stackrel{\text{def}}{=} \langle t \parallel (u|v) \cdot \star \rangle^*$$

$$t u \stackrel{\text{def}}{=} \langle t \parallel u \cdot \star \rangle^*$$

(Classical variant: solving the reduction rules of the operators and C and k_π with μ)





Abstract machines

Reward: Relationship with sequent calculus

$$\frac{\Gamma \vdash t : A \rightarrow B \quad \frac{\Gamma \vdash u : A \quad \overline{\Gamma \mid \star : B \vdash \star : B}}{\Gamma \mid u \cdot \star : A \rightarrow B \vdash \star : B}}{\langle t \parallel u \cdot \star \rangle : (\Gamma \vdash \star : B)}
 }{\Gamma \vdash \langle t \parallel u \cdot \star \rangle^* : B}
 = tu$$

Calculus L

Third ingredient: $\tilde{\mu}$ binder

The $\tilde{\mu}$ binder is symmetric to the adjoint construct:

$$\langle t \parallel \tilde{\mu}x.c \rangle \triangleright_{\mathbf{R}} c[t/x]$$

Application. Solving the equations of branching:

$$\langle l_1(t) \parallel (u|v) \cdot e \rangle \triangleright_{\mathbf{R}} \langle u \parallel t \cdot e \rangle$$

$$\langle l_2(t) \parallel (u|v) \cdot e \rangle \triangleright_{\mathbf{R}} \langle v \parallel t \cdot e \rangle$$

gives:

$$(u|v) \cdot e \stackrel{\text{def}}{=} (e_1|e_2)$$

where:

$$e_1 \stackrel{\text{def}}{=} \tilde{\mu}x. \langle u \parallel x \cdot e \rangle$$

$$e_2 \stackrel{\text{def}}{=} \tilde{\mu}x. \langle v \parallel x \cdot e \rangle$$



Calculus L

Third ingredient: $\tilde{\mu}$ binder

Consequence

The context is no longer necessarily a stack



Calculus L

Benefits of $\tilde{\mu}$

- Highlights choices in the evaluation order (Curien and Herbelin)
- Correspondence with sequent calculus (Curien and Herbelin, Wadler)
- Simpler equational theory: no more commutative cuts and related rules
- Completeness of contexts
- Simplifies orthogonality-based logical relations (M.-M.)

$$\mathbf{L} = \mu + \tilde{\mu}$$

Fourth ingredient: Polarisation

Guillaume Munch-Maccagnoni. **Focalisation and Classical Realisability**. In *Proc. CSL '09*, LNCS. Springer-Verlag, 2009

We define how terms and variables can be of either polarity (in the sense of **Girard and Danos, Joinet and Schellinx**)

Positive polarity The evaluation of a positive term is strict (call by value)

Negative polarity The evaluation of a negative term is delayed (call by name)

Having both is essential for extensionality



Fourth ingredient: Polarisation

Introducing a negative characterisation of polarisation

$$\text{let } x^\ominus = (\text{let } y^+ = t_1 \text{ in } t_2) \text{ in } t_3$$

vs. $\text{let } y^+ = t_1 \text{ in let } x^\ominus = t_2 \text{ in } t_3$

- Above: computes t_3 first
- Below: computes t_1 first

Negative characterisation of polarisation

Rejecting the hypothesis that composition is associative *a priori*.

So-called “Blass problem” (Abramsky, Mellies)



Fourth ingredient: Polarisation

Non-associativity everywhere

$$(h \circ g) \cdot f \neq h \circ (g \cdot f)$$

- : composition in call by value
- : composition in call by name

ML `let y = f x in h (fun () -> g y) ≠ h (fun () -> g (f x))`

Haskell `(\y->h (g y)) $! (f x) ≠ h (g $! (f x))`



Direct models

Various approaches to polarisation

1. Simplifying hypotheses: linearity, call-by-value, call-by-name
2. Indirect models: continuation calculi, call-by-push-value
3. **Here:** we propose to model proofs and programs through a structure where composition is not associative, which characterises polarisation **directly**.



Direct models

A **direct** denotational model offers an exact correspondence between operations of the model and constructions of the language.

Direct models

Comparative Table

Evaluation order	By value	By name	Polarised
Direct model	Thunk (Führmann)	Runnable monad (ex: Selinger)	Duploid
Indirect model	Monad T	Co-monad L	Adjunction $F \dashv G$
Programs	Kleisli maps $P \rightarrow TQ$	co-Kleisli maps $LN \rightarrow M$	Oblique maps $FP \rightarrow N$ $\simeq P \rightarrow GN$
Syntactic data	Values	Stacks	Both
Completion into	Thunkable expressions	Linear evaluation contexts	Both



The duploid construction

Inspired from Girard, Danos-Joinet-Schellinx, Laurent

Let $\uparrow \dashv \downarrow : \mathcal{N} \rightarrow \mathcal{P}$ be an adjunction:

$$\frac{P \rightarrow \downarrow N}{\uparrow P \rightarrow N} (\simeq)$$

P, Q objects of \mathcal{P}

N, M objects of \mathcal{N}

Polarised translation Recipe for defining a notion of morphism $A \rightarrow B$ for any A, B among the objects of \mathcal{P} and \mathcal{N} .

Definition An *oblique morphism* $f : P \rightarrow N$ is either $P \rightarrow \downarrow N$ or $\uparrow P \rightarrow N$

The duploid construction

Inspired from Girard, DJS, Laurent

White composition

$$\frac{\frac{f : P \rightarrow N}{f : P \rightarrow \downarrow N} (\simeq) \quad \frac{g : \downarrow N \rightarrow M}{g : \downarrow N \rightarrow \downarrow M} (\simeq)}{\frac{g \circ f : P \rightarrow \downarrow M}{g \circ f : P \rightarrow M} (\simeq)} (\circ)$$

Black composition

$$\frac{\frac{f : P \rightarrow \uparrow Q}{f : \uparrow P \rightarrow \uparrow Q} (\simeq) \quad \frac{g : Q \rightarrow N}{g : \uparrow Q \rightarrow N} (\simeq)}{\frac{g \bullet f : \uparrow P \rightarrow N}{g \bullet f : P \rightarrow N} (\simeq)} (\circ)$$



The duploid construction

Inspired from Girard, DJS, Laurent

A morphism $A \rightarrow B$ is an oblique morphism:

$$A^+ \rightarrow B^\ominus$$

with

$$A^+ \stackrel{\text{def}}{=} \begin{cases} \downarrow N & \text{if } A = N \\ P & \text{if } A = P \end{cases}$$

$$A^\ominus \stackrel{\text{def}}{=} \begin{cases} N & \text{if } A = N \\ \uparrow P & \text{if } A = P \end{cases}$$

$g \cdot f$ composition in \mathcal{N} of $A \xrightarrow{f} P \xrightarrow{g} B$

$g \circ f$ composition in \mathcal{P} of $A \xrightarrow{f} N \xrightarrow{g} B$



Duploids

Composition in a duploid is written:

$g \circ f$ when the middle object is of any polarity

$g \circ f$ when the middle object is negative

$g \bullet f$ when the middle object is positive

$$(h \bullet g) \bullet f = h \bullet (g \bullet f)$$

$$(h \circ g) \circ f = h \circ (g \circ f)$$

$$(h \bullet g) \circ f = h \bullet (g \circ f)$$

$$(h \circ g) \bullet f \neq h \circ (g \bullet f)$$

See Loday's **duplicial algebras**

Jean-Louis Loday. **Generalized bialgebras and triples of operads.** *arXiv preprint math/0611885*, 2006





Duploids

Linear and thunkable morphisms

Definition

Linear morphism f associates to its right

$$f \circ (g \circ h) = (f \circ g) \circ h$$

Thunkable morphism f associates to its left

$$h \circ (g \circ f) = (h \circ g) \circ f$$

\mathcal{D}_l category of linear morphisms

\mathcal{D}_t category of thunkable morphisms



Duploids

Hom-functor: comparison

Category \mathcal{C} : hom-functor $\mathcal{C}(-, =) : \mathcal{C}^{\text{op}} \times \mathcal{C} \rightarrow \mathbf{Set}$

Duploid \mathcal{D} : hom-functor $\mathcal{D}(-, =) : \mathcal{D}_t^{\text{op}} \times \mathcal{D}_l \rightarrow \mathbf{Set}$



Duploids

Structure of Shifts

\mathcal{N} sub-category of morphisms $N \rightarrow M$.

\mathcal{P} sub-category of morphisms $P \rightarrow Q$.

Characterisation of shifts

- $\uparrow : \mathcal{D}_l \rightarrow \mathcal{N}_l$ equivalence of categories
- $\downarrow : \mathcal{D}_t \rightarrow \mathcal{P}_t$ equivalence of categories

Corollary

1. Adjunction $\boxed{\downarrow \dashv \uparrow}$ with $\downarrow : \mathcal{N} \rightarrow \mathcal{P}$ and $\uparrow : \mathcal{P} \rightarrow \mathcal{N}$.
2. Adjunction $\boxed{\uparrow \dashv \downarrow}$ with $\downarrow : \mathcal{N}_l \rightarrow \mathcal{P}_t$ and $\uparrow : \mathcal{P}_t \rightarrow \mathcal{N}_l$.



Duploids

Functor: comparison

Functor $F : \mathcal{C} \rightarrow \mathcal{C}'$ between categories gives a natural transformation $\mathcal{C}(-, =) \rightarrow \mathcal{C}'(F-, F=)$

Functor $F : \mathcal{D} \rightarrow \mathcal{D}'$ between duploids preserves linearity and thunkability and gives a natural transformation $\mathcal{D}(-, =) \rightarrow \mathcal{D}'(F_t-, F_l=)$



Duploids

Theorem

$$\mathbf{Dupl} \triangleleft \mathbf{Adj}$$

i.e.: the duploid construction (or polarised translation) extends into a functor $j : \mathbf{Adj} \rightarrow \mathbf{Dupl}$ which admits a full and faithful right adjoint $i : \mathbf{Dupl} \rightarrow \mathbf{Adj}$.

\mathbf{Dupl} is the category of duploids and duploid functors.

\mathbf{Adj} is the category of adjunctions and *pseudo-morphisms of adjunctions*.



Results

1. Category *Dupl* of duploids and duploid functors.
2. $Dupl \triangleleft Adj$, in particular:
 - Every duploid comes from an adjunction,
 - Not all adjunctions come from a duploid.
3. Characterisation of adjunctions that arise from duploids.
(Completion of values and stacks + quotient)
4. Internal language / Abstract machines.
5. The duploid is a category if and only if the adjunction is *idempotent*.
6. Kleisli categories correspond to duploids where \downarrow or \uparrow are bijective on objects.



Decomposing CPS translations

Continuation-passing style translations

- provide implementations (Sussman and Steele) to programming languages,
- provide denotational semantics (Strachey and Wadsworth) to programming languages,
- are still an active topic.

The goal is to explain continuation-passing-style translations through **L** calculi. To do so we decompose CPS translations for delimited control calculi through CPS translations for **L** calculi.

Decomposing CPS translations

Delimited control operators

- First-class context delimiters model the fact that captured contexts have a finite extent (**Felleisen**).
- We focused on Danvy and Filinski’s *shift/reset* variant. ($S/\langle \cdot \rangle$)

$$\langle E[S(\lambda x.t)] \rangle \triangleright \langle t[\lambda y.\langle E[y] \rangle/x] \rangle$$

$$\langle V \rangle \triangleright V$$

- We used the decomposition of **Herbelin** and his collaborators with the binder $\mu\hat{t}p$. This binder models an auxiliary stack of contexts:

$$\langle t \parallel e_0 \rangle [e_1, \dots, e_n]$$



Decomposing CPS translations

- The *shift/reset* calculus in call-by-value as well as four call-by-name variants embed, by “*solving equations*”, into a single polarised calculus \mathbf{L} called $\mathbf{L}_{\text{pol},\hat{\text{ip}}^+}$.
- Method: showing that the CPS translations factor through $\mathbf{L}_{\text{pol},\hat{\text{ip}}^+}$, thus refining well-known decompositions of $\lambda\mu$ calculi (**Laurent** and others).
- All the translations preserve equivalences. All the translations starting from \mathbf{L} calculi simulate reductions as well.
- The call-by-name setting is not simply deduced from the call-by-value setting “by duality”.



Involutive negation

- We review the issue of the involution of negation in the proof theory of classical logic
- We give a Curry-style classical natural deduction satisfying $\neg\neg A \simeq A$, which enjoys a formulae-as-types notion of construction.

Motivations

- Explaining an isomorphism of types such as $\neg\neg A \simeq A$ (as in Girard’s **LC**).
- We gave in CSL’09 a syntax for Girard’s **LC**, but defining a notion of quasi-proofs (or programs) was not immediate due to the perfect symmetry. Yet quasi-proofs are essential in classical realisability to obtain models in the sense of model theory.



Involutive negation

Idea 1

We distinguish the type $\sim A$ of captured stacks from the type $A \rightarrow \perp$ of continuations

- From Girard and Danos-Joinet-Schellinx: the rules of negation hide cuts, in particular the following one:

$$\frac{\Gamma, N \vdash \Delta}{\Gamma \vdash \neg N, \Delta}$$

- From Krivine: distinguishing the pseudo-type $\mathcal{X}^- \stackrel{\text{def}}{=} \{k_\pi \mid \pi \in \mathcal{X}\}$ from the type of continuations.
- From Felleisen-Clements: high-level access to captured stacks.



Involutive negation

Idea 2

Delimited control provides a constructive interpretation for \perp in an untyped setting.

(Inspired from Herbelin et al.)



Involutive negation

The $\lambda\ell$ calculus

$$\ell : (A \rightarrow \perp) \rightarrow \sim A$$

$$\text{send} : \sim A \rightarrow A \rightarrow \perp$$

ℓ refines C and j refines k .

$$\langle \ell \| t_{\ominus} \cdot e_{+} \rangle [\pi_{\ominus}, \sigma] \triangleright_{\mathbf{R}} \langle t_{\ominus} \| j_{e_{+}} \cdot \pi_{\ominus} \rangle [\sigma]$$

$$\langle j_{\pi_{+}} \| \pi \rangle [\sigma] \triangleright_{\mathbf{R}} \langle [\pi] \| \pi_{+} \rangle [\sigma]$$

$$\langle \text{send} \| [\pi_{\varepsilon}] \cdot t_{\varepsilon} \cdot \pi'_{\ominus} \rangle [\sigma] \triangleright_{\mathbf{R}} \langle t_{\varepsilon} \| \pi_{\varepsilon} \rangle [\pi'_{\ominus}, \sigma]$$

+ constants that can access the contents of captured stacks $[\pi]$.



Involutive negation

Method

- Extensional reasoning, necessary for proving isomorphisms, is provided on $\lambda\ell$ by the calculus $\mathbf{L}_{\text{pol},\hat{\text{tp}}^\ominus}$.
- The $\lambda\ell$ calculus is then a subset of terms of $\mathbf{L}_{\text{pol},\hat{\text{tp}}^\ominus}$ with no free co-variables which is closed under composition.

Characterisation of linearity and thunkability

A context e is linear if and only if one of the following two equivalent properties holds:

1. for all c, t, q, q' one has $\langle t \parallel \tilde{\mu}q'.\langle \mu q.c \parallel e \rangle \rangle \simeq_{\text{RE}_p} \langle \mu q.\langle t \parallel \tilde{\mu}q'.c \rangle \parallel e \rangle$,
2. for all c, α one has $\langle \mu \alpha.c \parallel e \rangle \simeq_{\text{RE}_p} c[e/\alpha]$.

A term t is thunkable if and only if one of the following two equivalent properties holds:

1. for all c, e, q, q' one has $\langle \mu q'.\langle t \parallel \tilde{\mu}q.c \rangle \parallel e \rangle \simeq_{\text{RE}_p} \langle t \parallel \tilde{\mu}q.\langle \mu q'.c \parallel e \rangle \rangle$,
2. for all c, x , $\langle t \parallel \tilde{\mu}x.c \rangle \simeq_{\text{RE}_p} c[t/x]$.



Characterisation of linearity and thunkability

Semantic notion of stoup

- We have the following stoup-like property:

Let c be a command. The set:

$$\{\alpha^+ \mid \mu\alpha^+.c \text{ is thunkable}\} \cup \{x^\ominus \mid \tilde{\mu}x^\ominus.c \text{ is linear}\}$$

has at most one element.

- With this property we define a semantic notion of stoup that encompasses the stoup of Girard's **LC**, which itself is hard to manipulate formally. By *semantic* we mean that the notion is attached to the properties rather than to the form.

Characterisation of linearity and thunkability

Remarks

- We were not satisfied with the restrictions called *A-normal form* / *monadic normal form* / η restriction. This is because substitution has to be defined for all terms and contexts.
- The notion cannot be reduced to a notion of value as developed through continuation calculi. For instance any term which reduces to a value in a referentially transparent way is thunkable.

Perspectives

Connectives in terms of duploids L calculi only use simple β and η rules. This suggests that connectives should be characterised with a suitable notion of universal property over duploids. (Ongoing work with Fiore and Curien)

Orthogonality-based logical relations Proving syntactic properties of systems, inspired by normalisation-by-evaluation. A conceptual treatment of renaming is expected.

Proof theory of arithmetic It appears that many aspects of the works of Krivine could be simplified with an adequate treatment of polarities.



Perspectives

- Polarised intuitionistic logic** Liang and Miller uncovered an intuitionistic system where the two negations \sim and $\cdot \rightarrow \perp$ coexist. It is similar to our classical natural deduction in which the classical principle $(A \rightarrow \perp) \rightarrow \sim A$ is removed.
- Operational models of programming languages** The **L** calculi do not account yet for linearity-related optimisations described in CPS translations.
- Rewriting theory** The rewriting theory of extensional syntaxes with constructors is a recent and active field of study.



Thank you