1 Introduction

The goal of this research was to probe the semantic structures needed for imperative and logic programming, taking careful account of abstraction and control.

In imperative programming, a crucial problem area has always been sharing. Because of aliasing, and behind-the-scenes dependencies caused by state change, specification and reasoning techniques for realistic languages are excessively complex. The reasons for this complexity run deep, and are not merely down to idiosynchracies in language designs. But models of imperative computation can have elegant properties, as was demonstrated in work on possible world semantics and, particularly, interference (e.g., [5, 4]).

In logic programming, a crucial problem has always been to control the dependencies between branches of a proof encountered in proof-search. This can also be seen as a sharing problem, and is essential to understand in order to give a comprehensive treatment of logic programming languages with substructural features or modules. Also, an account of sharing on the logic programming level could give rise to a principled approach to management of dependencies between different parts of a specification, a serious problem in specification formalisms.

Sharing is just one of the areas of commonality between logic programming and imperative programming. Others are control and abstraction. We first conceived this project with the view that these intuitive connections were suggesting deeper semantic links – between logic and logic programming, imperative programming, and categorical semantics – than is commonly supposed. We believed that cohesive theoretical results, and novel applications, could be obtained by pursuing these connections.

We would suggest that the initial, broad conception of our project has been well borne out by the results we have obtained, as we describe below. Although not all of our results span all of the areas involved, a certain cohesiveness can be seen, for example, in the work on control, with its basis in categorical semantics, and the related use of continuations in work on the proof theory of classical logic. And it can be seen even more strongly in our work on bunched logic, which might be considered as a meeting of the logical perspective of Pym and the imperative viewpoint of O’Hearn, with categorical semantics as mediator; the amount of gelling of the varied constituents exceeded any of our initial expectations.

2 Research Achievements

2.1 Bunched Logic

Our work in bunched logic [OP99, Pym99, O’H99] represents the strongest outcome of this project. The logic $\mathbf{BI}$ of bunched implications was introduced by O’Hearn and Pym in a short article in the Bulletin of Symbolic Logic [OP99]. A systematic presentation, including its semantics and model theory, proof theory, type theory and computational interpretation, is presented in a forthcoming monograph by Pym [Pym00b] (which will appear either in the Studies in Logic and Computation series, published by RSP/Wiley, or in Kluwer’s Applied Logic series).

$\mathbf{BI}$ is a logic that combines intuitionistic logic (with the usual connectives $\wedge, \rightarrow, \vee, \bot$) and a ba-
sic substructural logic (with implication $\rightarrow$ and conjunction $\ast$). Although other logics (particularly linear logic) also combine the two fragments, the way that BI mixes them together is novel, as illustrated by the core theoretical perspectives on it.

- **Algebraic and Categorical Models.** A categorical model of BI is a doubly-closed category, which is a (bi)cartesian closed category together with an additional symmetric monoidal closed structure. A collapsed version of this structure is a BI-algebra, a Heyting algebra with an additional residuated commutative monoid structure. These models make the similarity and difference with linear logic evident. BI uses one category (or algebra) with two closed structures, where linear logic uses two separate categories or algebras (one of which is often a Kleisli category).

- **Bunched Proof Theory.** BI's proof theory is formulated using bunches, which are nested, tree-like contexts built using two forms of combination. One form, "$\therefore$", internalizes the cartesian product (additive conjunction) in a doubly closed category, while the other, "$\therefore$", internalizes the monoidal product (substructural conjunction). The introduction rules for the two implications mimic the two adjunctions in a categorical model.

\[
\frac{\Gamma, \phi \vdash \psi}{\Gamma \vdash \phi \rightarrow \psi} \rightarrow \frac{\Gamma ; \psi \vdash \phi}{\Gamma \vdash \phi \rightarrow \psi},
\]

- **Resource Semantics and Topological Models.** This is a truth semantics, which gives a declarative way of reading BI's formulae. The basic idea is that the truth of a formula $\phi$ is judged "locally", in the sense that it talks about a circumscribed collection of resources. This is formulated in a possible worlds style, where the world $m$ in a forcing judgement $m \models \phi$ is chosen from an ordered commutative monoid of worlds. In example models the resources correspond to areas of computer memory, markings in a Petri net, or processes. The most sophisticated version of the semantics uses topological ideas (Grothendieck topologies) to treat the intuitionistic connectives, with continuity conditions relating to the monoid structure on worlds.

- **The Sharing Interpretation.** This is a way of reading our various semantics, in terms of sharing between parts of a formula or parts of a proof/$\lambda$-term. For instance, when viewed as a function type, $A \rightarrow B$ is the type of functions that don't share resources with their arguments.

Although the basic system is intuitionistic, we have also investigated classical variations; for example, an algebraic model of Boolean BI is a BI-algebra in which the Heyting component is in fact Boolean.

These varied perspectives on BI reinforce one another, and lend support to our contention that it is an elegant and naturally occurring logic. Our technical investigations been driven by them.

- **Basic Proof Theory.** The system has a natural deduction and sequent calculus formulations, with normalization and cut elimination properties [Pym00b].

- **Type-theoretic Interpretation.** A $\lambda$-calculus of proof terms, syntactic results such as subject reduction and normalization, and soundness and completeness results for interpretation in doubly-closed categories. [O’H99, Pym00b].

- **Syntactic Control.** The type system is used to merge syntactic control of interference and Idealized Algol [O’H99, O’H00]. This is the best example of the sharing interpretation of proofs/$\lambda$-terms, and was in fact one of the early inspirations for BI. The resulting system resolves problems with recursion and jumps in the original syntactic control [9], which had befuddled experts for years.

- **Resource Semantics.** Soundness and completeness for a version of the semantics based on Grothendieck topologies. Demonstration of a range of concrete models that illustrate the informal reading, including models based on CCS, Petri nets, and pointers [OPY, Pym00b].

We have also explored BI's place in the spectrum of substructural logics [8], contributing, we believe to the general understanding of these systems. BI is distinct both from the R-like relevant logics, in its semantics and treatment of implication, and from linear logic-like systems, in its treatment of the relationship between intuitionistic and linear implications and distributivity.

These results establish propositional BI as a theoretically cohesive system, with a significant early application. In addition, in work at QMW supported by another EPSRC grant (“Verified Bytecode”), and in work further afield, the pointer model of BI has been utilized as part of an approach to longstanding problems in Hoare logic for pointer programs.
Research in this area, and on applications suggested by the other concrete models, is progressing rapidly.

To sum up our work on bunched logic, following its initial discovery based on proof-theoretic and category-theoretic ideas, the resource interpretations we have been developing have begun to take on more significance; the interpretation in terms of sharing, in particular, fits both imperative and logic programming extremely well. This direction shows good promise as far as further applications go, both those building one existing ones (on pointers and logic programming) and in new directions (such as concurrency and databases).

2.2 Predication and Type Dependency

Many substructural logics have been proposed where structural rules are limited, but where also the treatment of variables in term languages admits all structural rules. When substructural logics are used as type systems they give rise to substructural term languages, and it would seem natural to want to have an account of predication that takes such terms into account. But, although it seems natural, there are many conceptual and technical difficulties that must be overcome to obtain a workable approach; this is why multiplicative quantification or predication has remained a challenging open problem.

The predicate version of BI investigated by Pym [Pym99, Pym00b] has two universal and existential quantifiers, just as propositional BI has two conjunctions and implications. There are additive quantifiers, whose semantics is given just as in intuitionistic logic. There are also multiplicative quantifiers, \( \forall_{\text{new}} x . \phi \) and \( \exists_{\text{new}} x . \phi \). In terms of resource semantics, \( \forall_{\text{new}} x . \phi \) says that \( \phi \) is true for all individuals that access separate resources from the current resource. Good example models are obtained from imperative programming, where \( \forall_{\text{new}} x \) quantifies over values that refer to new or fresh storage. Indeed, the nomenclature for the quantifiers was partially inspired to their relationship to new storage variables in Idealized Algol.

The syntax and proof theory of multiplicative quantification uses trees or bunches of variables, just as propositional BI uses bunches on the propositional level. This gives a link into another, earlier, piece of work, on the substructural logical framework RLF. RLF uses an ordered form of bunched context to formulate its rules.

In the course of the project we obtained the following results.

- The \( \lambda \)-calculus, the type theory underlying RLF, has been formulated. This resolves a difficult open problem concerning the combination of linearity and type dependency. Basic technical results such as normalization and confluence have been established [IP98, 2].
- We showed how the multiplicative function type could be used to represent substructural object logics, with a "uniform encoding" that establishes tight connections between proofs in the object logic and the metalogic. This was a problem area in the original LF.
- Logical frameworks for operational semantics [2, IP98]: an adequate representation of languages such as ML with reference types including, for the first time in logical frameworks, a treatment with close control over garbage cells.
- A class of concrete, possible worlds models of the dependent linear function type [IP99, IP].
- A categorical semantics, model theory, and proof theory of multiplicative quantification in predicate BI [Pym99, Pym00b].
- Proof-search (logic programming) [OP99, Pym99]: the proof theory of predicate BI provides a basis for a logic programming language within which sharing and non-sharing of data by procedures can be modelled directly. This work is continuing, in particular with a study of applications to modules, really exploiting multiplicative predication and quantifiers, in the work of Pym’s PhD student, Pablo Armelin.

These results are but a start. From the theoretical point of view multiplicative quantification and type dependency are delicate and raise many problems. A more substantial and general theory is called for, and is under development (taking cue from these initial results). As for applications, this direction appears to be particularly relevant in situations where substructural features are on view in a type system or object logic, and for this reason substructural frameworks are beginning to attract attention from researchers on proof-carrying code.

2.3 Classical Logic, Proof-search and Categorical Semantics

Pym, in joint work with Eike Ritter, has pursued a substantial project in the semantics and proof theory of classical logic. Motivated by a desire to understand the operational and denotational semantics of
models of computation based on proof-search, e.g., logic programming.

Our work here falls into three main chunks, each of which relies on a type-theoretic analysis of the key classical connective, i.e. disjunction:

- Basic proof theory [PW00a]: Here we develop a proof-theoretic analysis of the embedding of the intuitionistic sequent calculus within the classical sequent calculus. We exploit type-theoretic methods to show how to recover intuitionistic soundness from classical searches. We apply our analysis to uniform proofs, a candidate theoretical foundation for logic programming;

- Application to resolution [PW00b]: We apply our methods to obtain a rational reconstruction of intuitionistic resolution from classical resolution. Again, our techniques are primarily type-theoretic. This work is related to Pym’s work with Harland on resource-distribution in linear logic, in which linearly sound proofs are recovered from classical searches via systems of Boolean constraint equations [1]. Both of these works are contributions to our general approach to a theory of search [GP00, Pym00a];

- Categorical model theory [PR]: We provide a semantic analysis of classical disjunction, revealing how the familiar collapse of a categorical semantics of classical proofs is sensitive to the formulation of disjunction.

Of particular significance is our development of a semantics of classical disjunctive proofs using a category of continuations.

2.4 Control and continuations

Our most substantial work on control was led primarily by our RA, Hayo Thielecke, who arrived in July 1997, having completed a thesis on the categorical structure of continuations. While highly original, this thesis also gave an extremely abstract picture of control. We, and he, were interested to see if its ideas could be leveraged to make more concrete insights about control.

This began with a paper on the expressiveness of the most powerful form of control, the call/cc construct. Thielecke showed that the ability to use a continuation multiple times, by backtracking, led to great expressive power, in terms of distinguishing ability. This work came about as a result of an exploration of technical properties of the categorical models, specifically properties of the centre of a premonoidal category. But the results could be phrased just in terms of programs and operational semantics.

Thielecke’s study of the relationship between the centre and notions of “effect free” morphism has played a leading role in other work. One paper [PT99], by him and John Power, gives a categorical account of higher-order structure in call by value, building on the Power-Robinson notion of premonoidal category. Other work, by Carsten Fuhrmann at Edinburgh, investigates some of the properties identified by Thielecke for effects other than continuations.

Thielecke’s next step was to begin an analysis of the behavioural, jumping properties of exceptions. This was done first by comparing their expressiveness to that of call/cc. In a paper with Jon Riecke from Bell Labs, Thielecke showed that pure exceptions and call/cc (i.e., without state) are incomparable in expressive power: each can break program equivalences that the other can not. This drives a very clear wedge between these two prominent forms of control, one that is fundamental in that it involves program equivalence only, and does not rely on implementation strategies (stacks v heaps), important though these are.

In a further paper the impact of state on the expressive power of continuations and exceptions is analyzed [Thi00], and a final paper makes a further comparison by exploring their typing properties and making connections to logic [Thi].

All told, Thielecke’s work is remarkable for two reasons. First, it links very abstract categorical notions to concrete programming language questions, questions that can be posed without mentioning the abstract techniques used to help resolve them. Second, he has provided the first fundamental study of behavioural properties of exceptions. In ways, the exception work is modest theoretically, but that is part of the point given the dearth of work giving formal analyses of exceptions. This begins to correct an imbalance in the theory of control. While call/cc is extremely powerful only one language, Scheme, has it as part of its standard. But while there have been dozens of papers on continuations, there has only been a handful of papers analyzing the much more widely used exceptions.

2.5 Abstract structures, logical relations and realizability

Our work on BI helps address notions of shared and private state. However we have also made progress
on reasoning about abstraction properties. The major work here was done by Robinson in collaboration with John Power [PR00a, PR00b], but building on earlier pieces of work by Robinson alone and O’Hearn. The achievement here was to make precise a standard notion of (observational) equivalence for abstract structures, and to prove that a certain generalised form of logical relation was sound and complete for proving equivalence. In order to allow for higher types one has to get beyond the simplest set-based semantics, and this makes the simple form of logical relations inappropriate. What Robinson did was take a form of relation introduced by Jung and Tiuryn to study definability, and called by them “Kripke logical relations of varying arity”, show how these could be regarded as unary relations, and then adapt a binary variant to prove completeness. In [PR00a] the original presentation and proof is cleaned up considerably, and in [PR00b] the result is extended from cartesian closed to monoidal closed categories, thus extending the class of languages to which it applies.

Two pieces of further work present themselves. The first is to attempt to extend the result to closed premonoidal categories, increasing the language class to include many more forms of control. The second is to relate this work to alternative approaches which use a relaxation of the logical relations condition. Some material on this had to be cut from the accepted version of [PR00b] for reasons of space. However Power and Robinson have established a link in which the relaxed version is a kind of global section of theirs.

The second thread was again in conjunction with Power [PR99]. This work is a generalisation of Moggi’s notion of computational monads to allow data to be kept in the context. The point of this can be seen by considering the notion of state. In Moggi’s original treatment the State monad is

$$X \mapsto [S \mapsto [S \times X]]$$

so that a term of type $B$ with a variable from $A$ is represented by a morphism

$$A \mapsto [S \mapsto [S \times X]]$$

A more natural and equivalent representation is

$$S \times A \mapsto S \times B$$

which would allow a more optimal treatment of the state, and this is what is allowed by Power and Robinson’s generalisation. Specifically, this builds on previous work by Power and Rosolini, establishing one possible generalisation. Robinson observed that this could be made symmetric between input and output, allowing both a more natural mathematical structure, and potentially a wider range of examples. The work provides a mathematical treatment up to the point where it would be relatively simple to write a Haskell library implementing dyads, converting computational monads to dyads, and allowing composition of dyads. This is essentially the functionality of the standard library for computational monads.

A particularly tantalizing account of abstraction is provided by realizability semantics. This provides a way of tying an extensional structure to an intensional one, so that one gets structures with extensional behaviour, but in which every operation is denotable by a term. This intuition is made much more precise and general than previously in some work in progress [MRR]. This work is still unfinished, but is concerned with the categorical construction of realizability models, and a a draft paper containing precise results should be ready for submission within the next month.

3 Research Impact

Most of the work presented here has had some impact inside the Semantics and Theoretical Computer Science communities. However, some of it is starting to have an impact outside of these communities. Our work on $\text{BI}$ is still very new, and challenges some long-held assumptions. Some leading figures, such as Girard, have been very supportive, while others have yet to see its advantages. $\text{BI}$ has been used in recent advances on program logic for pointers, a longstanding open problem [3]. The basic idea is to use $\text{BI}$’s resource semantics to make local statements about areas of memory. We have submitted a grant proposal based on these ideas, but they have also been taken up elsewhere in the community, for example by John Reynolds at Carnegie-Mellon [10], and by Uday Reddy’s group in Birmingham [11].

Thielecke results have been recognised well outside the conventional semantics community, particularly by the compiler design research groups at Carnegie-Mellon and Cornell. This is particularly because of his use of concrete runnable examples to illustrate his semantic distinctions.

Our work on classical logic has solved longstanding problems in the theory of proof-search, as well as in the semantics of classical proofs, and has generated problems contributing to two new MATHFIT grant proposals.
4 Training and Personal Outcomes

We employed two RA's on the grant. During the period he worked for us Hayo Thielecke cemented his position as the UK's leading expert on continuations, and extended his reputation internationally as one of the leaders in an admittedly rather specialised, but still significant, field. His intention on joining us had been to find a permanent position in a strong research group, and while with us he was offered and accepted a permanent lectureship at Birmingham, where he is now.

The second RA, Cristiano Calcagno is a research student whose supervision is being shared by Moggi at Genova and O'Hearn at Queen Mary. Calcagno has started to publish, and has a single-author paper accepted at the prestigious Principles of Programming Languages conference (POPL01). This is a significant achievement for a student. Calcagno's work is part of a stream which has led to a new research grant application by O'Hearn and Pym, bringing in Bornat.

We also used the grant to give some travel support to an EPSRC student Paul Levy, see publications [Lev99]. Levy is now in Boston, having been head-hunted as an RA by Harry Mairson. Samin Ishtiaq and Pablo Armelin, both EPSRC-funded students, have been closely involved with the project. Ishtiaq has now graduated and has taken up a post in formal methods with the Cambridge-based chip-design company, ARM. Armelin's thesis is now nearing completion.

The investigators have also advanced their careers thanks in part to this grant. Both O'Hearn and Pym have been promoted to full professor, with the research conducted on the grant putting the final seal on their cases.

Finally, departmentally, this grant has been significant. It has been a major source of funding — leading to several current grant applications by O'Hearn and Pym, by Hyland, Pym and Robinson, and by Pym and Ritter — during the development of the Logic and Foundations of Programming research group.

5 Explanation of Expenditure

Staff: The major part of our expenditure was on our designated RA, Hayo Thielecke. Hayo left a little early, in order to fit in a lengthy visit to Tennent in Canada before taking up his post at Birmingham (this visit was funded by Tennent). This left us with a small amount spare, which enabled us to fund Calcagno for a period. Calcagno's funding should be viewed more as training than advanced research, and although his time on the project did not result in any direct publications, the training he has received has enabled him to produce work of the kind mentioned above.

Travel: The next largest item is travel. We have travelled extensively during the period, and presented a good deal of work at conferences, as can be seen from the attached project bibliography. Presentation of most of the conference papers there was supported out of this grant.

Equipment: There were two major areas of equipment expenditure. The first was the purchase of a number of personal computers at the start of the grant. This helped significantly to kit out the burgeoning research group. One of them is being used to prepare this report. The other area was contributions to the department's normal rolling program of equipment replacement. Notable expenditure here included an early contribution to a workgroup printer, and a late contribution when the department replaced its file server.

Consumables: A large item is a fixed charge levied by the department covering stationery, printing and photocopying costs, postage, phone bills and the like. The remainder is the usual small pieces of equipment and software.

6 Further Research

We have mentioned earlier the continuing nature of much of this work. In particular, we should single out again the continuing work on pointer logic and BL, for which we have submitted a new proposal to EPSRC. The present work has laid the foundations for an approach to proof-search, for which support is again being sought.

7 Web site and documentation

This report, together with links to participants pages and papers at http://www.dcs.qmw.ac.uk/~edmundry/L56439/

The bibliography for this report is split into two sections. The list of supported publications contains those works which have been at least partially supported by this grant.
Supported Publications


Other References


Nicola Gambino. I describe the main results obtained during the EPSRC postdoctoral fellowship that I held at the University of Cambridge. The fellowship focused on the interplay between category theory and mathematical logic. 1. Wellfounded trees W-types in categories. Types of wellfounded trees, or W-types, are one of the most important components of Martin-Löf's dependent type theories. They allow us to define a wide class of total imperative programs correspond, via the game semantics, to proofs in WS1. Moreover, because WS1 is more expressive than the typing system for a typical programming language, it can express finer behavioural properties of strategies. In particular, we can embed rst-order intuitionistic logic with equality, Polarized Linear Logic, and a nitary imperative language with ground store, coroutines and some infinite data structures. We also take rst steps towards answering some of the questions posed by the logic and its semantics: Are there any formulas which only have imperative proofs, but the foundations of a new "categorical" combinatory logic are laid down. Compilation and evaluation techniques are investigated. A simple abstract machine, called the Categorical Abstract Machine, is presented: it has served as the core of the implementation of the language CAML, of the ML family, developed at INRIA-Roquencourt and Ecole Normale Superieure, and first released in 1987. A mathematical semantics of sequentiality is proposed, in which "sequential algorithms" rather than functions are used to interpret procedures. The theoretical investigation has led to the development of a programming language, CDSO, in which basic and functional types are not differentiated. The evaluation framework is a demand-driven data flow network.