

Geometric Logic in Computer Science

Steve Vickers

sjv@doc.ic.ac.uk

Department of Computing, Imperial College
180 Queen's Gate, London SW7 2BZ, United Kingdom

Abstract

We present an introduction to geometric logic and the mathematical structures associated with it, such as categorical logic and toposes. We also describe some of its applications in computer science including its potential as a logic for specification languages.

1 Introduction

What I shall present here is a personal overview of how—and why—I see geometric logic being used in computer science.

Mark Ryan commented on an earlier, rather different, draft of this paper that he understood the title and thought “Oh, good”, but quickly ran into words that made no sense to him. My revised intention therefore is to write a popularization of geometric logic for the benefit of computer scientists. I shan't present any new results, and in fact I shall hardly even present any old ones in any technical details, but I shall try to explain the essential features of the mathematics by explaining them (rather than by leaving the reader to sift them out of a grand formal structure). I shall also try to take stock of the ingredients we have to hand and what roles they should play.

2 Geometric Theories

Although the full mathematical insights come only through category theory, I'm going to start from a very logical point of view because I think in computer science people are more comfortable with that. It is important to realise that the particular properties of geometric logic come as much from its particular definition of *theory* as from its definition of formulas.

Our notion of *vocabulary* is standard in many-sorted logic: it comprises *sorts*, *predicate symbols* and *function symbols* (representing *total* functions). Remember, of course, that functions and predicate symbols have *arities*, specifying the number and sorts of their arguments and (for functions) the sort of the result, and that term formation has to respect the arities. Function and predicates with no arguments are *constants* and *propositions*.

If there are *no* sorts, then function symbols are impossible because there are no possible result sorts. But propositions are possible, so a theory with no sorts is a propositional theory.

Given a vocabulary, *geometric formulas* are constructed in a way that is completely unsurprising except that a peculiar collection of connectives is used: these are

- ♦ conjunction (\wedge for binary conjunction, **true** for nullary conjunction)
- ♦ disjunction (\vee for binary disjunction, **false** for nullary, and moreover *arbitrary infinitary disjunction* is allowed: $\bigvee_{i \in I} \phi_i$)
- ♦ equality (this is sorted, so in $e_1 = e_2$ the two terms e_i must have the same sort)
- ♦ existential quantification \exists

There is also the restriction that a formula may have only finitely many free variables.

Why these connectives? To my mind, the most compelling reason is that they have observational content in a way that the others (\neg , \implies , \forall and \bigwedge) do not. For a detailed discussion, see Vickers [10] or (for the propositional case) [9]; but, briefly, the idea is that the sorts and formulas represent “observational classes” in the real world, each comprising two ingredients:

- ♦ how to “apprehend” elements
- ♦ how to determine that two elements are equal

Moreover, these “how to”s are *positive* (e.g. nothing about determining inequality), *serendipitous* (they merely describe how to know in retrospect when you have succeeded) and *finite* (they don’t call on you to do an infinite amount of work). The geometric connectives can then be interpreted as operations on observational classes.

For the rest of the paper, I shall concentrate more on the mathematical consequences of the particular choice of connectives.

The admission of infinitary disjunction may seem alarming. Since the disjuncts will be indexed by a set, a full formal system for geometric logic must include a formal set theory and that is a heavy overhead. But even the *coherent* fragment, in which disjunctions must be finitary, is interesting, and as we shall see it is desirable to go beyond and bring in formal constructions that are geometric but non-coherent. One important example uses a theory of finite sets and a formalism that includes universal quantification bounded over finite sets.

A *geometric theory* comprises a vocabulary and a set of axioms of the form $\phi \vdash_{\mathbf{x}} \psi$ where

- ♦ \mathbf{x} is a set $\{x_1, \dots, x_n\}$ of sorted variables
- ♦ ϕ and ψ are geometric formulas whose free variables are all taken from \mathbf{x}

Such a set $\{\phi_i \vdash_{\mathbf{x}^i} \psi_i : i \in I\}$ of axioms should be thought of as meaning $\bigwedge_i \forall x_1^i, x_2^i, \dots, x_n^i. (\phi_i \implies \psi_i)$. Now \forall , \implies and infinite \bigwedge are not geometric connectives and don’t appear in ϕ_i or ψ_i ; but just at this one level they are allowed into the geometric theory.

Note that this does not use the standard logical idea that a theory is a vocabulary together with a set of sentences (formulas without free variables). Hence geometric logic may fail to fit assumptions made for a general purpose logic environment. Also, the fact that implication is not internalized in the logic means that Hilbert style presentations of the logic are impossible, and there is no deduction theorem.

Labelled Turnstiles

The label \mathbf{x} on the turnstile is not peculiar to geometric logic. It is seen in Lambek and Scott’s treatment [7] of intuitionistic logic, and its use improves even classical logic. The essential logical point behind it is that the use of the free variables \mathbf{x} is just as much a hypothesis as the premiss ϕ —it hypothesizes the presence of values \mathbf{x} . Some arguments are valid only in the presence of values not explicitly mentioned in ϕ or ψ , and such values must be referred to in the set \mathbf{x} .

In many standard presentations of classical logic, one is allowed to deduce $\forall x.P(x) \vdash \exists x.P(x)$: from $\forall x.P(x)$ we can deduce $P(x)$ by \forall -elimination, and then $\exists x.P(x)$ by \exists -introduction. The deduction looks natural enough, but of course it is invalid in any model with an empty carrier: $\forall x.P(x)$ will then be true, $\exists x.P(x)$ false. The traditional classical response is to reject empty carriers—after all, what is the point of having the language of predicate logic if there is nothing for it to talk about?—but that approach doesn’t work at all well in constructive logic, where “non-emptiness” is a much subtler notion, and even classically it is distinctly problematical in many-sorted logic. The labelled turnstiles allow more careful rules of deduction that are valid even for empty carriers. The example deduction depends not just on the truth of $\forall x.P(x)$, but also on the presence of x . Labelling turnstiles we see that $\forall x.P(x) \vdash_x P(x) \vdash_x \exists x.P(x)$, so we can write $\forall x.P(x) \vdash_x \exists x.P(x)$ meaning that the entailment holds *in the presence of an x* , i.e. if the carrier is inhabited. From that, we should not deduce $\forall x.P(x) \vdash \exists x.P(x)$. This example is not geometric, but similar considerations apply for us. In a sequent calculus formalization, we can reason freely within a context (list of free variables). What we deduce within one context is true also in any bigger context:

$$\frac{\phi \vdash_{\mathbf{x}} \psi}{\phi \vdash_{\mathbf{x},y} \psi}$$

However, contexts can be reduced only in a controlled way using \exists -elimination.

$$\exists E : \frac{\phi \vdash_{\mathbf{x},y} \psi}{\exists y.\phi \vdash_{\mathbf{x}} \psi} \quad y \text{ not in } \mathbf{x}, \text{ nor free in } \psi$$

The effect in natural deduction is that we cannot freely invent new free variables (or constants). Some will be given us as the context for the overall problem (e.g. the free variables in the premisses and conclusion), and they can also be introduced in a controlled manner for \exists -elimination (and, in more standard logics, for \forall -introduction); but that is all.

Examples of Geometric Theories

1. Algebraic theories, possibly many sorted. There are too many different meanings for the word “algebraic”, but what I mean here is “defined by finitary operators and equational laws”. These are geometric theories presented with sorts and functions, but no predicates, and all the axioms are of the form $\vdash_{\mathbf{x}} t_1 = t_2$.

2. Essentially algebraic theories. I mention these explicitly because although they can look difficult to define they share most of the pleasant universal algebraic features of algebraic theories. The big generalization is that the operators can represent operations that are partial, though not in an unrestricted way. The operators are arranged in a well-founded hierarchy, and for each operator its domain of definition is defined by a conjunction of equations involving more primitive operators. The equational laws are interpreted in the sense “if both sides are defined then they are equal”. Putting such a description into the form of a geometric theory is not difficult, though of course a partial operator must be expressed as a predicate symbol, not a function.

A good example is the theory of categories. It has two sorts, for objects and arrows. The most primitive operators (necessarily total) are source, target and identity; and then composition has its domain of definition determined by equality between the target of one arrow and the source of the other.

3. Topological spaces. Let X be a topological space, ΩX its topology (family of open subsets). A corresponding geometric theory can be defined as follows:

- ♦ *no sorts* (it is a propositional theory—hence no functions either)
- ♦ for each open set a , a proposition P_a
- ♦ if $a \subseteq b$ are open sets, then an axiom

$$P_a \vdash P_b$$

- ♦ if S is a family of open sets, then an axiom

$$P_{\cup S} \vdash \bigvee_{a \in S} P_a$$

- ♦ if moreover S is finite, then an axiom

$$\bigwedge_{a \in S} P_a \vdash P_{\cap S}$$

(The converses of the last two axioms follow from the first.)

If $x \in X$, then x gives a model of the theory in which P_a is interpreted as **true** iff $x \in a$. It may be that different points x can give the same model, and that some models do not arise from points in this way. If these pathologies do not happen, and the points are in bijection with the models of the theory, then X is *sober*. If you’re familiar at all with locales, you’ll understand that generators and relations for a frame give propositions and axioms for a propositional geometric theory (see [9]).

Some Logical Manipulations

I shall not attempt to give a full set of proof rules for the logic—you can find them in Makkai and Reyes[8]. There are no surprises, because (apart from the infinite disjunctions) the logic is a restriction of standard logic.

On a short digression, let me show how to simplify the sequents considerably.

Proposition 1 *Every geometric formula $\phi(\mathbf{x})$ is equivalent to one of the form $\bigvee_i (E_i \wedge \exists \mathbf{y}^i \cdot \bigwedge_{j=1}^{n_i} P_{ij})$, where*

1. *The sets \mathbf{x} and \mathbf{y}^i are disjoint.*

2. Each E_i is a conjunction of non-trivial equations among the free variables \mathbf{x} . (A “trivial” equation is one of the form $x = x$, which can obviously be omitted.)
3. Each P_{ij} is either a primitive predicate applied to variables, or is an equation $z = t$ where z is a variable (from \mathbf{x} or \mathbf{y}^i) and t is a primitive function symbol applied to variables.

The description of the possible P_{ij} ’s doesn’t look nice. But it is always possible to present a geometric theory without function symbols, by replacing them by their graphs as predicate symbols, and then each P_{ij} is just a primitive predicate applied to variables.

Proof I shan’t give a detailed proof; after all, I haven’t given the full proof rules. Instead, I shall show how to use equivalences that you might reasonably expect geometric logic to have.

First, formulas of this form are (up to equivalence) closed under the geometric connectives. For disjunctions, this is obvious. Next, consider existential quantification $\exists x$ (x in \mathbf{x}). This should (expected equivalence) distribute over disjunction, and the disjuncts are of the form $\exists x.(E \wedge \exists \mathbf{y}.\bigwedge_{j=1}^n P_j)$. If x appears in E , i.e. there is an equation $x = x'$ (or $x' = x$), then our disjunct is equivalent to $(E' \wedge \exists \mathbf{y}.\bigwedge_{j=1}^n P_j)[x'/x]$, where E' is obtained from E by omitting all equations $x = x'$ or $x' = x$. If x does not appear in E , then the disjunct is equivalent to $E \wedge \exists x.\exists \mathbf{y}.\bigwedge_{j=1}^n P_j$. Finally, consider conjunction. Assuming that conjunction distributes over arbitrary disjunction, we get a disjunction of formulas $(E \wedge \exists \mathbf{y}.\bigwedge_{j=1}^n P_j) \wedge (E' \wedge \exists \mathbf{y}'.\bigwedge_{j=1}^{n'} P'_j)$. By renaming, we can assume that the sets \mathbf{y} and \mathbf{y}' (and \mathbf{x} too, of course) are disjoint, and then the disjunct is equivalent to $(E \wedge E') \wedge \exists \mathbf{y}, \mathbf{y}' . (\bigwedge_{j=1}^n P_j \wedge \bigwedge_{j=1}^{n'} P'_j)$.

It remains to be shown that atomic formulas are of this form. An equation between variables is OK. An equation $f(\dots) = g(\dots)$ is equivalent to $\exists z.(z = f(\dots) \wedge z = g(\dots))$, and the only remaining problem is to arrange for predicates and functions to be applied only to variables. This is done using equivalences such as that between $P(f(\dots), \dots)$ and $\exists u.(u = f(\dots) \wedge P(u, \dots))$. \square

Let us note something rather curious. The equations get divided up between “functional” equations, $z = t(\dots)$, which can be replaced by predicates in a transformed theory presentation and are kept amongst the conjuncts, and “structural” equations $x = x'$, which are ineradicable and pushed outside the existential quantification. The structural equations look like atomic formulas, and you might think their natural place is conjoined with P_{ij} ’s under the existential quantifications. But there is a good sense in which ‘ $E \wedge \exists \mathbf{y}$ ’ is a single unit, a generalized quantifier. E generates an equivalence relation on \mathbf{x} ; let us then choose a representative from each equivalence class and define a new set \mathbf{z} of variables comprising these canonical representatives and also \mathbf{y} . We have an obvious relabelling function $r : \mathbf{x} \longrightarrow \mathbf{z}$, and $E \wedge \exists \mathbf{y}$ is left adjoint to relabelling: $\phi \vdash_{\mathbf{z}} \psi [r(x)/x(x \in \mathbf{x})]$ iff $E \wedge \exists \mathbf{y}.\phi \vdash_{\mathbf{x}} \psi$. (When E is empty, this just reduces to well-known properties of \exists .)

The proposition allows a rather drastic simplification of geometric axioms on the left-hand side, using \vee -elimination and \exists -elimination: for $\bigvee_i (E_i \wedge \exists \mathbf{y}^i.\bigwedge_{j=1}^{n_i} P_{ij}) \vdash_{\mathbf{x}} \psi$ can be replaced by a set (one for each i) of axioms $E \wedge \exists \mathbf{y}.\bigwedge_{j=1}^n P_j \vdash_{\mathbf{x}} \psi$, and then each of these can be replaced (after suitable jug-

gling of the variables) by one of the form $\bigwedge_{j=1}^n P_j \vdash_{\mathbf{z}} \psi$.

3 Models

If a geometric theory is given, then the models for it are understood in an utterly standard way. First, the vocabulary must be interpreted: sorts as sets (the carriers), predicates as relations (subsets of products of carriers) and functions as functions (each from a product of carriers to a carrier). Once that is done, then every geometric formula can be interpreted as a relation in its free variables; then for an axiom $\phi \vdash_{\mathbf{x}} \psi$, both ϕ and ψ can be interpreted as subsets of the product of carriers for the sorts of the variables in \mathbf{x} ; and then the interpretation is a model iff for every such axiom, the subset corresponding to ϕ is included in that for ψ .

Homomorphisms

Suppose a geometric theory is given, and we have two models M and N of it. A homomorphism from M to N is a family of carrier functions $\alpha_\sigma : M_\sigma \rightarrow N_\sigma$, one for each sort σ , that respect the vocabulary ingredients (the axioms don't enter in here). Specifically, if $f(x, y, z, \dots)$ and $P(x, y, z, \dots)$ are function and predicate symbols, then (suppressing sorts)

- ♦ $f_N(\alpha(x), \alpha(y), \alpha(z), \dots) = \alpha(f_M(x, y, z, \dots))$
- ♦ $P_M(x, y, z, \dots) \implies P_N(\alpha(x), \alpha(y), \alpha(z), \dots)$

For algebraic theories, this is exactly the usual definition of homomorphism between algebras. For theories that include predicates, it is worth pointing out the direction ' \implies ' in the second of these conditions. ' \Leftarrow ' would not be compatible with the algebraic homomorphisms, as you can see if you consider replacing an algebraic operator (say multiplication for monoids) by its graph: $P(x, y, z)$ means ' $z = x.y$ '. For an arbitrary monoid homomorphism α we have $z = x.y \implies \alpha(z) = \alpha(x).\alpha(y)$, but \Leftarrow would not be right in general. As an example, for the theory of posets the homomorphisms are the monotone functions.

It is also worth noting that once these two conditions are known for the vocabulary ingredients, they are also true for *all* terms and geometric formulas: this is because of the positivity of geometric logic. Of course, as soon as negation is included, the second condition is destroyed because the implication for $\neg P$ gives us the reverse implication for P . Hence, this notion of homomorphism is one that is not so useful in ordinary logic.

The topological example is interesting. A homomorphism will carry no data, but it still has to obey a non-trivial condition: that for each proposition P_a , if it holds in M then it must hold in N . Let us write $M \sqsubseteq N$ if M and N satisfy this for every P_a , so that there is a (unique) homomorphism from M to N . If M and N correspond to the points x and y , then this says that every open neighbourhood of x also contains y (i.e. meets $\{y\}$), in other words x is in the topological closure of $\{y\}$. This is exactly the specialization preorder on points.

The Category of Models

It is easy to see that under this notion of homomorphism, the models of a theory \mathbb{T} form a category, $\text{Mod}(\mathbb{T})$ —though the topological example shows that $\text{Mod}(\mathbb{T})$ can have extra “topological” structure that is not categorical.

In general, this category has little categorical structure, lacking in general most limits and colimits. However, a very important exception is that it has all filtered colimits. I don’t want to define these in detail (see, e.g., Johnstone [5]), but they are the categorical generalization of directed joins, and I shall describe two particular cases. There are some general points to remember. First, they are constructed set-theoretically: a filtered colimit of models is carried by the set-theoretic filtered colimits of the carriers. Second, the existence of these colimits is intimately bound up with the geometric restrictions that conjunctions must be finite, and only finitely many free variables are allowed in each formula. Third, although the general filtered colimits may seem arcane, even the special cases are crucial in domain theory, both in finding least fixpoints within domains and in solving domain equations.

The first example of filtered colimit is the ω -colimit, of a diagram

$$M_0 \xrightarrow{\alpha_0} M_1 \xrightarrow{\alpha_1} M_2 \xrightarrow{\alpha_2} \dots$$

If $i < j$, let us write α_{ij} for the composite $\alpha_i; \dots; \alpha_{j-1} : M_i \rightarrow M_j$. First, suppose this is a diagram of sets and functions. You can think of progression along the diagram as bringing in more and more elements (insofar as the α_i ’s are not onto), and making more and more equalities among them (insofar as the α_i ’s are not 1-1). The elements in the (co-)limit and the equalities between them are exactly those that appear at some finite stage. (If you want to be more formal, first take the disjoint union of the M_i ’s, then define an equivalence relation by $x \equiv y (x \in M_i, y \in M_j)$ iff $\alpha_{ik}(x) = \alpha_{jk}(y)$ for some k , and then the colimit is the set of equivalence classes.)

Now let us return to the original problem, in which the M_i ’s are models (of \mathbb{T}) and the α_i ’s are homomorphisms. We can apply the set-theoretic construction to the carriers, but more work is needed to show that we still have a model. If f is a function symbol, then to define $f(x, y, z, \dots)$ in the limit we find a finite stage at which x, y, z, \dots all exist (there is such a finite stage *because there are only finitely many variables* x, y, z, \dots) and calculate $f(x, y, z, \dots)$ there. This gives us a well defined result in the limit, because the homomorphisms α_i preserve the results at the finite stages. Predicates are similar; $P(x, y, z, \dots)$ holds in the limit iff it holds at some finite stage (and hence at all subsequent ones, from the homomorphism property). Now consider an axiom $\phi \vdash_{\mathbf{x}} \psi$. If ϕ holds in the limit, then it must hold at some finite stage—the finiteness of conjunctions is used here, as well as the finiteness of the set \mathbf{x} —, so ψ also holds there, so ψ holds in the limit.

I’ll mention one other example, that of splitting idempotents, to dispose of a possible misconception. I said that filtered colimits were the categorical generalization of directed joins, and you know that directed joins are essentially infinite: finite directed joins are trivial. This is not quite the case with filtered colimits. Finding a finite filtered colimit is equivalent to *splitting an idempotent*: that is to say, if we are given a homomorphism $\alpha : M \rightarrow M$ such that $\alpha^2 = \alpha$,

we seek a diagram $M \begin{array}{c} \xrightarrow{p} \\ \xleftarrow{e} \end{array} N$ such that $p;e = \alpha$ and $e;p = \text{Id}_N$. The

argument that such colimits exist is similar to that for ω -colimits.

Let me repeat two points that explain why this doesn't look like conventional logic. First, the above definition of homomorphism works well because geometric logic is positive (no negation or implication), and second, the filtered colimits exist because of the finiteness restrictions.

Varying Set Theory

A point of greater importance than you might expect is that if we vary the notion of set, then we vary the notion of *model*; for instance, we could take “set” to mean “object in some given elementary topos”, and—suitably interpreted, and provided there are enough infinitary colimits to cope with all the disjunctions—the definition of model still makes perfect sense. We have to be careful to reason constructively about these generalized models, for the logic appropriate to an elementary topos is not classical but intuitionistic. But this is not just a concession to the generalizing ambitions of constructivists. It has key significance in at least three ways.

First, the proof rules of geometric logic (as in [8]) are classically incomplete: that is to say, within a given theory, there may be a sequent that holds in all the classical set-theoretic models but is not provable from the axioms using the proof rules. Even in the propositional case, there are theories that have no models at all, but which are still consistent: you can't prove **true** \vdash **false**. (This arises from the infinitary disjunctions. When they are banned—i.e. in *coherent* theories—Deligne's theorem proves completeness.)

However, this incompleteness is the fault *not of the proof rules but of the models*. In the classical category \mathcal{S} of sets, the constraints imposed by having to satisfy excluded middle and choice sometimes make models impossible. But the proof rules are constructive in nature, and hold for models in non-classical set theories, and it turns out that they are complete as long as you allow your set theory to vary (technically, by taking models in other elementary toposes).

Second, although in general we must allow the set theory to vary to get completeness, *for any given theory* there is a “canonical geometric set theory” that contains a *generic* model—any sequent that holds in the generic model is provable. This set theory is really just made by taking the standard sets and freely adjoining a model of the theory. It will be treated in more detail in the next section, where we shall see how it can be used to understand the idea of interpreting one theory in another.

Finally, you'd often like to think of a theory as being concretely embodied in its class of models, but for an arbitrary geometric theory this can't be done naively because you don't know a priori where the models have to be taken from. I shall try to explain how *topos theory* answers this by providing a language that's designed to make it look as though we have a decent category of models.

4 Interpretations

One very restrictive idea of interpretation (of one theory, \mathbb{T} , in another, \mathbb{T}') is a syntactic one: interpret the sorts, predicates and functions of \mathbb{T} as sorts,

predicates and functions of \mathbb{T}' , and prove that axioms of \mathbb{T} become theorems in \mathbb{T}' . We shall be much more liberal, in effect by allowing syntactic interpretations not just in \mathbb{T}' but more generally in theories (or rather theory presentations) equivalent to it.

Equivalence of Presentations

We shall take it that two theories are equivalent if they have the same models (essentially, i.e. up to isomorphism). By incompleteness, it is not enough just to look at classical set-theoretic models; we should give equivalence arguments that are constructively valid and so hold in the more general elementary toposes where the models might be taken. Here are some ways of modifying a theory to get an equivalent one:

- ◆ Add axioms that are consequences of the given ones.
- ◆ Replace axioms by logically equivalent ones as outlined above.
- ◆ For each function symbol $f(x, \dots)$, replace it by a predicate $P(z, x, \dots)$ (which is to represent the graph of the function), and add axioms to say that it is total and single-valued:

$$P(z, x, \dots) = P(z', x, \dots) \quad \begin{array}{l} \vdash_{x, \dots} \quad \exists z. P(z, x, \dots) \\ \vdash_{z, z', x, \dots} \quad z = z' \end{array}$$

Also, eliminate f from formulas by replacing $Q(f(x, \dots))$ by $\exists u. (P(u, x, \dots) \wedge Q(u))$ (u new) and similar manoeuvres.

- ◆ Eliminate the sort structure by replacing sorts σ by unary predicates $S_\sigma(x)$ over a single new sort representing their disjoint union:

$$S_\sigma(x) \wedge S_\tau(x) \quad \begin{array}{l} \vdash_x \quad \mathbf{false} \quad (\sigma \neq \tau) \\ \vdash_x \quad \bigvee_\sigma S_\sigma(x) \end{array}$$

(Note how this works when we start with no sorts at all! The first axiom scheme has no instances, and the second becomes $\vdash_x \mathbf{false}$. The x on the turnstile stops this asserting out-and-out inconsistency; instead, the axiom forces the single sort to represent the empty set.)

The sorting discipline shown by the arities of the symbols must also be taken care of. For instance, if $Q(u)$ had arity σ , then there must be a new axiom $Q(x) \vdash_x S_\sigma(x)$. Functions must be replaced by their graphs, so that if $f(u)$, with arity $\sigma \rightarrow \tau$, is replaced by $P(v, u)$ with arity $\tau \times \sigma$, then we need axioms

$$\begin{array}{l} P(z, x) \vdash_{z, x} \quad S_\tau(z) \wedge S_\sigma(x) \\ S_\sigma(x) \vdash_x \quad \exists z. P(z, x) \quad (\text{modified totality}) \\ P(z, x) \wedge P(z', x) \vdash_{z, z', x} \quad z = z' \end{array}$$

- ◆ Add sorts that can be characterized uniquely up to isomorphism by geometric axioms. For instance, consider products. If σ and τ are existing

sorts, extend the presentation by a new sort π with functions $\text{fst} : \pi \rightarrow \sigma$, $\text{snd} : \pi \rightarrow \tau$, and axioms

$$\text{fst}(z) = \text{fst}(z') \wedge \text{snd}(z) = \text{snd}(z') \quad \begin{array}{l} \vdash_{x:\sigma, y:\tau} \exists z : \pi. (\text{fst}(z) = x \wedge \text{snd}(z) = y) \\ \vdash_{z, z'} z = z' \end{array}$$

In any model, the carrier for π is forced by the axioms to be the product of the carriers for σ and τ : so π is characterized uniquely up to isomorphism by σ and τ . Models for the new theory are essentially the same as those for the old theory, the only difference being that the product $\sigma \times \tau$ is given explicitly instead of implicitly.

Other constructions that can be characterized geometrically include co-products (disjoint unions—even infinitary ones), equalizers and coequalizers (slightly tricky! You need the infinitary disjunctions): in short, all colimits and finite limits.

- ◆ Here is an interesting construction that can be characterized geometrically: finite power sets. The finite power set $\mathbb{F}X$ is just the free semilattice over X , and as it happens free things (for finitary algebraic theories) can always be characterized geometrically. Here is a logical presentation.

Let σ be a given sort. We add a new sort π and functions $\{\cdot\} : \sigma \rightarrow \pi$ (the singleton embedding), $\emptyset : \pi$ and $\cup : \pi \times \pi \rightarrow \pi$; also axioms

$$\begin{array}{l} \vdash_{S, T, U} S \cup (T \cup U) = (S \cup T) \cup U \\ \vdash_S S \cup \emptyset = S \\ \vdash_{S, T} S \cup T = T \cup S \\ \vdash_S S \cup S = S \\ \vdash_S \bigvee_{n=0}^{\infty} \exists x_1, \dots, x_n. S = \{x_1, \dots, x_n\} \\ \{x_1, \dots, x_m\} \subseteq \{y_1, \dots, y_n\} \vdash_{\substack{x_1, \dots, x_m \\ y_1, \dots, y_n}} \bigwedge_{i=1}^m \bigvee_{j=1}^n x_i = y_j \end{array}$$

where

$$\begin{array}{l} \{x_1, \dots, x_n\} \equiv_{\text{def}} \{x_1\} \cup \dots \cup \{x_n\} \\ \{\} \equiv_{\text{def}} \emptyset \\ S \subseteq T \equiv_{\text{def}} S \cup T = T \end{array}$$

Using this construction, we can see that universal quantification is geometric, *provided that it is bounded over finite sets*. If ϕ is a formula with free variables in \mathbf{x} , $x : \sigma$ is one of those variables, and $S : \mathbb{F}\sigma$, then we can define

$$\forall x \in S. \phi \equiv_{\text{def}} \bigvee_{n=0}^{\infty} \exists y_1, \dots, y_n. (S = \{y_1, \dots, y_n\} \wedge \bigwedge_{i=1}^n \phi[y_i/x])$$

The knowledge that this is possible seems to be folklore, but I know of nowhere where the formal details have been set out.

To show how different equivalent presentations may be, an example in [10] has two of which one has infinitely many sorts and functions but no predicates, while the other has one sort, no functions and infinitely many predicates (all unary).

Giraud Frames

Given a theory \mathbb{T} , one fundamental trick of categorical logic is to make a category whose objects are the sorts and predicates, both primitive and derived: a derived sort is one characterizable geometrically, and a derived predicate is just a formula. These are all the things that in models get interpreted as sets, and it is useful to think of them as sets parametrized by the model. For this reason, the category obtained behaves sufficiently like the category of sets to have many nice properties. In particular, by ignoring the parameter, the model, you can generally reason validly as though the objects actually are sets, though of course the reasoning has to be constructive. This category is really the “canonical geometric set theory” for \mathbb{T} referred to earlier; it includes of course the ingredients explicitly presented in \mathbb{T} , and these constitute the “generic model”.

If the theory is \mathbb{T} , then this category is written $\mathcal{S}[\mathbb{T}]$: this means \mathcal{S} (the category of sets) extended by formally adjoining the ingredients of \mathbb{T} (as indeterminate sets) subject to the axioms. (The notation comes from the notation $R[X]$ for a ring of polynomials.) The category is usually called the “classifying topos” of \mathbb{T} , but I shall offer my excuses for not doing so when I discuss toposes. Instead I shall call it the “Giraud frame presented by” \mathbb{T} .

In general, a Giraud frame is a category with all colimits and finite limits, satisfying certain other conditions that (i) make the colimits and limits behave like those in \mathcal{S} , and (ii) ensure that it can be presented by a small (in the set-theoretic sense) theory presentation. The conditions are exactly those set out for Giraud’s theorem in Johnstone [4].

One aspect of Giraud frames being similar to \mathcal{S} is that constructive set theory can be interpreted in them, and we can talk about models of a theory \mathbb{T} in a Giraud frame—in fact, a model of \mathbb{T} in $\mathcal{S}[\mathbb{T}']$ is just an interpretation of \mathbb{T} in \mathbb{T}' . (This is in accordance with what we originally said about interpretations, because $\mathcal{S}[\mathbb{T}']$, by being made from *all* sorts geometrically derivable from \mathbb{T}' , can be seen as including all the theories equivalent to \mathbb{T}' .) Once that is given, then we know—up to isomorphism—how to interpret all the derived types, the objects of $\mathcal{S}[\mathbb{T}]$, and in fact we get a functor from $\mathcal{S}[\mathbb{T}]$ to $\mathcal{S}[\mathbb{T}']$. Moreover, if we have two models and a homomorphism between them, then the homomorphism carrier functions for the primitive sorts extend uniquely to the derived sorts, and categorically we get a natural transformation.

To summarize: if \mathbb{T} is a theory and A is a Giraud frame, then

- ♦ Models of \mathbb{T} in A are functors from $\mathcal{S}[\mathbb{T}]$ to A that preserve the colimits and finite limits. (We shall call such functors *homomorphisms* between Giraud frames. More correctly, they are adjunctions $f^* \dashv f_*$ for which the left adjoint f^* , which anyway preserves colimits, also preserves finite limits.)
- ♦ Homomorphisms between models are natural transformations between the Giraud frame homomorphisms. (Sorry to have the two different kinds of homomorphisms so close together.)

Thus the constructive model theory has been turned into category theory.

Non-geometric Type Constructors

Since the finite power set constructor (which I’m calling \mathbb{F}) is geometric, it’s natural to ask whether the full power set \mathbb{P} is. The answer is no: it cannot be

characterized by geometric axioms. The same also goes for exponentials Y^X (the set of functions from X to Y). The way this is proved is by working in the Giraud frame. As it happens, Giraud frames have \mathbb{P} and exponentials (the categories are elementary toposes, in fact), characterized uniquely but non-geometrically. The functors that correspond to interpretations preserve all the geometric constructions. It is possible with some work to see these categories and functors concretely, and to find examples of interpretation functors that do not preserve \mathbb{P} and exponentials: hence those constructions are not geometric.

There is a parallel here to the way that \forall and \implies come in to a theory at just one level. The predicate and function symbols can be considered to be elements of power type or function type; but this is allowed just at the one level: you can have sorts such as $\mathbb{F}\mathbb{F}\mathbb{F}X$ for finite sets of finite sets of finite subsets of X , but you can't do this with \mathbb{P} .

It is natural to ask whether geometric logic is classical or intuitionistic in nature, though on the face of it the question is meaningless because the prime distinguishing feature—excluded middle—cannot be expressed in the absence of negation. The non-geometric structure of Giraud frames casts some light on this, because using it one can interpret the non-geometric logical connectives and it turns out that excluded middle is not obeyed: hence geometric and intuitionistic logic are intimately associated with one another. On the other hand, the very fact that this extra structure is not preserved by Giraud frame homomorphisms shows a sharp distinction between the two logics.

5 Toposes

I have been somewhat coy so far about the word “topos”. If you're at all familiar with the literature, you will know that it generally means “category somewhat like \mathcal{S} ”. There are elementary toposes, and amongst those there are some—in fact exactly what I have called Giraud frames—that are called Grothendieck toposes; and $\mathcal{S}[\mathbb{T}]$ is normally called the classifying topos of \mathbb{T} .

I'm going to use the word with a quite different meaning. It would be vain to expect to overturn the established usage, but I'd like to try to show you how the word can convey some different intuitions. These are not new insights of my own. Grothendieck invented the word topos as a back-formation from “topology” (so toposes are “those things of which topology is the study”) and said that a topos is a generalized topological space, and topos theorists understand these intuitions perfectly well. However, they have not often expressed them clearly, and I shall try to explain them by enforcing a distinction—between toposes and Giraud frames—that is analogous to that between locales and frames.

Definition 2 *A topos is the space of models (the classifying topos) of a geometric theory. If \mathbb{T} is a geometric theory, then we write $[\mathbb{T}]$ for its classifying topos.*

Definition 3 *Let D and E be toposes. A geometric morphism (or map) from D to E is a continuous transformation of points of D into points of E .*

IMPORTANT! These definitions are mystical. They are intended to convey not the mathematical formalization but the intuitive meaning, and if you try to analyse them compositionally through detailed accounts of the terms “space”, “model”, “geometric theory” and so on, you will get a false formalization.

- ◆ A “space” is more than an unstructured class of points, for we have already seen that the models form a category. But there is also some mysterious topological structure that we haven’t attempted to formalize, so a topos is not just the category of models. One problem in the formalization is how to account for this “topology”.
- ◆ Similarly, “continuous transformation” is mysterious. Actually, even for topological spaces it is quite mysterious.
- ◆ “Models”—where? It is not enough to consider models in \mathcal{S} ; they must be allowed in arbitrary Giraud frames. The formalization must allow for this.
- ◆ “Geometric theory presentations” have sets of sorts, predicates, functions and axioms, and sets of disjuncts in a disjunction: so what geometric theories are possible depends on what your underlying set theory is. Each elementary topos leads to its own theory of Grothendieck toposes. We shall assume a fixed underlying category of “the classical sets” \mathcal{S} .

Implementation 4 *A topos D is equipped with a Giraud frame $\mathcal{S}D$. If D and E are toposes, then a geometric morphism f from D to E is equipped with a Giraud frame homomorphism $\mathcal{S}f$ (or f^*) from $\mathcal{S}E$ to $\mathcal{S}D$ (note the reversal of direction!)*

“Implementation” means that this is just a means to an end, a formalization that gives us a mathematical handle on the prior intuitions. Toposes could equally well be implemented as theory presentations—more easily, in fact, though geometric morphisms are then harder to describe.

“Equipped with” has no technical substance—giving a topos is just the same as giving a Giraud frame. But it is intended to dispel ideas that a topos “is” a Giraud frame and “has” objects and morphisms that are those of the Giraud frame. It decouples the language of toposes from that of Giraud frames, and hides the implementational details behind the “ \mathcal{S} ” prefix. Large parts of the traditional language of toposes are designed to reinforce the “space of models” intuitions, and my own opinion is that this can usefully be taken even further. (See the remarks on notation at the end of this section.)

An example of this is in the apparently perverse direction of geometric morphisms. Suppose $f : D \rightarrow E$ is a geometric morphism, with $D = [\mathbb{T}]$. A point of D is a model of \mathbb{T} , in other words a Giraud frame homomorphism from $\mathcal{S}D$ to your favourite Giraud frame A (where you like to look for models). Composing with $\mathcal{S}f$ turns it into a Giraud frame homomorphism from $\mathcal{S}E$ to A , and so transforms points x of D into points $f(x)$ of E , in accordance with the conceptual definition and without regard to your choice of A . It is easily checked that this extends to homomorphisms, giving a functor, but as for the topological aspects of continuity we are really defining what continuity means by our implementation.

Note also that if f and g are two geometric morphisms from D to E , and $\alpha : f \rightarrow g$ is a natural transformation, then each point x of D gives rise to a homomorphism $\alpha_x : f(x) \rightarrow g(x)$; and that this is natural with respect to x .

Summary: Suppose we have $[\mathbb{T}] \begin{array}{c} \xrightarrow{f} \\ \Downarrow \alpha \\ \xrightarrow{g} \end{array} [\mathbb{T}']$ where \mathbb{T} and \mathbb{T}' are geometric

theories, f and g are geometric morphisms and α is a natural transformation. Then the *intuitions* are

- ♦ $[\mathbb{T}]$ is the space of models for \mathbb{T} . It has category structure using the homomorphisms as arrows (this intuitive category is not \mathcal{S} -like, and is quite different from $\mathcal{S}[\mathbb{T}]$).
- ♦ f is a continuous transformation of models for \mathbb{T} (points of $[\mathbb{T}]$) into models for \mathbb{T}' . It is functorial with respect to homomorphisms, and so can be considered a functor.
- ♦ α is a natural transformation from f to g considered as functors.

As a particularly important example, consider the empty theory $*$ with no vocabulary and no axioms. It has a unique model, and the Giraud frame $\mathcal{S}[*]$ it presents is just \mathcal{S} . $[*]$ can be thought of as a one-point space. A geometric morphism $f : [*] \rightarrow [\mathbb{T}]$ intuitively just picks out a model of \mathbb{T} ; looking at it another way, $\mathcal{S}f$ is a model of \mathbb{T} in \mathcal{S} . A natural transformation between two geometric morphisms f and g is just a homomorphism. Up to equivalence, there is only one geometric morphism from $[\mathbb{T}]$ to $[*]$: intuitively, every point of $[\mathbb{T}]$ has to map to the unique point $*$ of $[*]$.

For another example, consider the theories **MON** and **SET** of monoids and sets. **SET** can be interpreted in **MON** in an obvious way—its only ingredient is a single sort, which is interpreted as the single sort in **MON**—and this corresponds to a geometric morphism $\text{Forget} : [\mathbf{MON}] \rightarrow [\mathbf{SET}]$. On points, it works exactly like the forgetful functor from the category of monoids to that of sets, taking a monoid and returning its carrier, forgetting the multiplicative structure. It is in fact intuitively helpful to think of the classifying toposes $[\mathbf{MON}]$ and $[\mathbf{SET}]$ as the categories of monoids and sets (in these algebraic examples there is no extra “topological” structure), but this is clearly incompatible with any idea that these toposes *are* their Giraud frames. (For instance, the Giraud frame $\mathcal{S}[\mathbf{SET}]$ that implements $[\mathbf{SET}]$ is most definitely not the Giraud frame \mathcal{S} .) That is why I took such care to separate out the notions.

It is also interesting to note that there is a geometric morphism $\text{Free} : [\mathbf{SET}] \rightarrow [\mathbf{MON}]$ that on points constructs the free monoid over a set. Free is left adjoint to Forget just as one would expect from ordinary categories, though the definition of adjunction here has to be the general one that applies within 2-categories.

The 2-category structure—the natural transformations between geometric morphisms—is important, because it gives real support to our intuition that the individual toposes have the structure of categories (and categories with all filtered colimits, because filtered colimits of natural transformations always exist).

A good example of this is part of Johnstone’s discussion [6] of bagtoposes. If D is a topos, then the bagtopos $B_L(D)$ classifies set-indexed families of points of D , and a homomorphism from $(x_i)_{i \in I}$ to $(y_j)_{j \in J}$ is a function $\alpha : I \rightarrow J$ together with, for each i , a homomorphism $f_i : x_i \rightarrow y_{\alpha(i)}$. If you carried out this construction with a category instead of a topos, you’d be constructing the free category-with-all-coproducts over D . Now a category D has an initial object (nullary coproduct) iff the unique functor from D to 1 has a left adjoint, and it has binary coproducts iff the diagonal functor from D to $D \times D$ has a left adjoint. (These are easy to see as direct expressions of the definition of coproduct.) Hence a category that has both these left adjoints has all finite coproducts; and if it has filtered colimits too then it has all coproducts. So we can reasonably define a topos D to “have all coproducts” iff these two left

adjoints both exist, and we just need the 2-categorical structure to be able to define what adjoints are. It then turns out that the two adjoints are equivalent to B_L -algebra structure for D , so for toposes too we can say that $B_L(D)$ is the free topos-with-all-coproducts over D .

Remarks on the Notation

Although the notation $\mathcal{S}[\mathbb{T}]$ is standard, its separation into $[\mathbb{T}]$ and $\mathcal{S}D$ is not. A happy accident of the notation is that \mathcal{S} can stand not only for **Sets**, but equally well for **Sheaves**. Locales can be understood as a particular (*localic*) kind of topos, namely those classifying a propositional theory, and for a locale D , $\mathcal{S}D$ is the category of sheaves over D . This could be made more general. If a topos D is a generalized space, then I propose that the objects of the corresponding Giraud frame $\mathcal{S}D$ should be called *sheaves* over D . Then, just as in [9] a locale has points and opens but not elements, a topos would have points and sheaves but not objects.

6 Three Examples in Computer Science

These are the three examples that I have worked on myself. I should say that applications of elementary toposes—“ \mathcal{S} -like categories”—are more common; it is specifically the application of geometric logic and Grothendieck toposes that is still quite new.

Geometric Theories and Databases

This is my paper [10]. The applications to database theory are very simplistic—as presented, the theory cannot cope with relations between entities, nor with database update to reflect change in the world (as opposed to improved knowledge in the database). However, it sets out some aspects of the move from propositional geometric logic in computer science—principally localic domain theory—to predicate logic.

First, we have an observational account. The propositional case (Abramsky [1], Vickers [9]) corresponds to observing a single atomic world, though the theories may often be constructed (e.g. for product domains) to allow for components. In the predicate case the components are allowed for in the logic itself: a theory then expresses ideas of (i) how we can “apprehend” components of the world—observe their existence and lay hands on them—, and (ii) how to observe equality between apprehended objects.

Second, the theory of the lower powerdomain is generalized to a “bagdomain” that naturally uses toposes instead of locales. (Johnstone [6] took this much further and also showed how to generalize the upper powerdomain.)

Third, it gave an illustration of the use of the well known categorical generalization of algebraicity, replacing ideal completions of posets by ind-completions of categories (i.e. free categories-with-all-filtered-colimits—see [5]).

Topical Categories of Domains

This work is still in preparation, though a preliminary account was given in Vickers [11]. It exploits the fact that theories of “information systems” used to present domains (there are various flavours) are geometric: so there are toposes classifying information systems and hence in a sense classifying domains. Where the sense falls down is in the morphisms—homomorphisms between information systems are not at all the same as the continuous maps between the domains, which must be represented by “approximable mappings”. However, the theory of approximable mappings is also geometric, so there is a classifying topos for them. Putting these together with the appropriate geometric morphisms gives a “topical category”, an internal category in the category of toposes. This means that starting from an ordinary category of domains, its unstructured classes of objects (domains) and arrows (maps) have been made into toposes and hence given categorical and topological structure.

The benefits of doing this are great. First of all, limits needed for finding least fixpoints within domains, and for solving domain equations, exist for free in the filtered colimits. When writing a domain equation $D = F(D)$, it is not possible to express F without it having the necessary continuity and functoriality properties. Because the functoriality is with respect to information system homomorphisms, not continuous maps, the problem of the contravariant argument in the function space construction vanishes—it turns out that for SFP, where function space is expressible, the information systems have so much structure needing to be preserved that homomorphisms correspond to embedding- projection pairs between domains.

It is hoped that this work will lead to an axiomatic account of domain theory.

Geometric Logic as Specification Language

When a software system is specified, an important aspect of it is the way it models the real world. For instance, for a credit account system the computer should be aware that the world contains a class of people (and that people have names, ages, and so on), a subclass comprising creditworthy people, and a subclass of *that* comprising the account holders. In the format of logical theories, we have a sort **person** and functions and predicates (confused with classes)

- ◆ name: **person** \rightarrow **string**
- ◆ age: **person** \rightarrow **num**
- ◆ cw, ah: \mathbb{P} **person**

We also want an axiom $\text{ah}(x) \vdash_x \text{cw}(x)$.

There is on the face of it no special reason why the logic should be geometric, but the observational account gives grounds for believing that the restrictions of geometric logic fit natural restrictions of the real world. For instance, compare cw with ah. There are untold numbers of people in the world, of whom untold numbers are creditworthy. The computer is certainly not expected to have a list of them all; rather it needs to know that if you do come across a person there are various procedures for establishing their creditworthiness—get a bank reference, ask their parents, see if they have an honest face, and so on. On the

other hand, it really does need a complete list of all account holders, so ah is fundamentally different from cw . In geometric logic this would be reflected by making ah not a predicate of arity **person** (“type \mathbb{P} person”), but a constant of sort \mathbb{F} person, and it is only when that is done that useful notions such as the cardinality of ah can be considered. In observational terms, the computer must know how to “apprehend” (i.e. in this case record in its database) individual people and finite sets of them, but not the entire class. The work of Hodges [3] on \mathbb{Z} lends support to this idea that a restricted logic is better for specification.

The equivalences inherent in geometric logic make functions equivalent to their graphs (just as in set theory), and in the observational account functions lose all their computational content and become retrospective checks that the result is correct. This obviously looks like specification as opposed to implementation, and again supports the idea that geometric logic is good for specification (but this time in a negative way, by saying that it is not good for expressing dynamic computational features).

These considerations suggest a program of using geometric logic—or at least a finitary approximation to it, including coherent logic and finite sets and universal quantification bounded over them—as a specification language. The idea is to take existing specificational notation (such as \mathbb{Z}) and give it a semantics in geometric logic and toposes: slogan—schemas are geometric theories. The expected benefits are

- ♦ Some existing specificational constructions will have to be dropped, but if the program is soundly based they should be in some sense unrealistic anyway (e.g. replacing \mathbb{P} by \mathbb{F}).
- ♦ There is a very precise geometricity criterion to decide when a proposed new construction is legitimate.
- ♦ A “categorical specification theory” (or perhaps “categorical schema calculus”) exists in the 2-category of toposes, the category structure being the key to modularization. The reason for this is that the aim of a module is to hide internal workings behind an interface specification of how the module is to relate to all other possible modules, and this is just what a universal property such as that for product or pullback does. It is category structure that makes this possible, by describing through the morphisms how the objects relate to each other. Working in the category of toposes, rather than the opposite category of Giraud frames, aids this through the spatial intuitions: for instance, a product topos is a “space of pairs”.

Acknowledgements

There are things that topos theorists know but do not often write down, and it is some of these mysteries that I have tried to express. But I am uncomfortably aware that I am nowhere near having climbed to the shoulders of the giants. I want to thank Peter Johnstone in particular for much of what understanding I do have. By his careful theoretic working [6] of my bagdomain constructions, not only has he shown me how the work ought to be done technically, but he has also given me invaluable insight into the way the toposes have to be seen in order for the results to make sense.

I must also thank Mark Dawson, Reinhold Heckmann and Mark Ryan for their careful reading of earlier drafts. Their many pertinent comments have

greatly helped me to improve the exposition. In addition, I am very grateful to Mark Dawson for producing the L^AT_EX version of the paper.

My own work mentioned here has been and continues to be supported by the UK Science and Engineering Research Council through the “Foundational Structures in Computer Science” project at Imperial College.

References

- [1] S. Abramsky, Domain theory in logical form, pp. 1-77 in *Annals of Pure and Applied Logic* vol. 51, 1991.
- [2] M.P. Fourman, P.T. Johnstone and A.M. Pitts (eds.), *Applications of Categories in Computer Science*, London Mathematical Society Lecture Note Series vol. 177, Cambridge University Press, 1992.
- [3] Wilfrid Hodges, Another Semantics for Z, unpublished notes.
- [4] Peter Johnstone, *Topos theory*, Academic Press, London, 1977.
- [5] Peter Johnstone, *Stone Spaces*, Cambridge University Press, 1982.
- [6] Peter Johnstone, Partial products, bagdomains and hyperlocal toposes, pp. 315-339 in [2].
- [7] J. Lambek and P.J. Scott, *Introduction to Higher Order Categorical Logic*, Cambridge University Press, 1986.
- [8] Michael Makkai and Gonzalo E. Reyes, *First Order Categorical Logic*, Lecture Notes in Mathematics 611, Springer-Verlag, 1977.
- [9] Steven Vickers, *Topology via Logic*, Cambridge University Press, 1989.
- [10] Steven Vickers, Geometric theories and databases, pp. 288-314 in [2].
- [11] Steven Vickers, Topical categories of domains, pp. 261-274 in Winskel (ed.) Proceedings of the CLICS Workshop 1992, Technical Report DAIMI PB - 397-I, Computer Science Department, Aarhus University, 1992.