QUESTIONS AND ANSWERS — A CATEGORY ARISING IN LINEAR LOGIC, COMPLEXITY THEORY, AND SET THEORY

Andreas Blass

Abstract. A category used by de Paiva to model linear logic also occurs in Vojtáš's analysis of cardinal characteristics of the continuum. Its morphisms have been used in describing reductions between search problems in complexity theory. We describe this category and how it arises in these various contexts. We also show how these contexts suggest certain new multiplicative connectives for linear logic. Perhaps the most interesting of these is a sequential composition suggested by the set-theoretic application.

Introduction

The purpose of this paper is to discuss a category that has appeared explicitly in work of de Paiva [18] on linear logic and in work of Vojtáš [25, 26] on cardinal characteristics of the continuum. We call this category $\mathcal{PV}$ in honor of de Paiva and Vojtáš (or, more informally, in honor of Peter and Valeria). The same category is implicit in a concept of many-one reduction of search problems in complexity theory [15, 23].

The objects of $\mathcal{PV}$ are binary relations between sets; more precisely they are triples $\mathbf{A} = (A_-, A_+, A)$, where A_- and A_+ are sets and $A \subseteq A_- \times A_+$ is a binary relation between them. (We systematically use the notation of boldface capital letters for objects, the corresponding lightface letters for the relation components, and subscripts $-$ and $+$ for the two set components.) A morphism from $\mathbf{A}$ to $\mathbf{B} = (B_-, B_+, B)$ is a pair of functions $f_- : B_- \to A_-$ and $f_+ : A_+ \to B_+$ such that, for all $b \in B_-$ and all $a \in A_+$,

$$A(f_-(b), a) \implies B(b, f_+(a)).$$

(Note that the function with the minus subscript goes backward.) Composition of these morphisms is defined componentwise, with the order reversed

1991 *Mathematics Subject Classification.* 03G30 03F65 03E05 03E75 18B99 68Q25.

Partially supported by NSF grant DMS-9204276.

Typeset by $\mathcal{A}_{\mathcal{M}}\mathcal{S}$-TEX

on the minus components: $(f \circ g)_- = g_- \circ f_-$ and $(f \circ g)_+ = f_+ \circ g_+$. This clearly defines a category $\mathcal{PV}$.

The category $\mathcal{PV}$ is the special case of de Paiva's construction $\mathbf{GC}$ from [18] where $\mathbf{C}$ is the category of sets. It is also the dual of Vojtáš's category GT of generalized Galois-Tukey connections [25, 26].

Intuitively, we think of an object $\mathbf{A}$ of $\mathcal{PV}$ as representing a problem (or a type of problem). The elements of A_- are instances of the problem, i.e., specific questions of this type; the elements of A_+ are possible answers; and the relation A represents correctness, i.e., $A(x, y)$ means that y is a correct answer to the question x.

There are strong but superficial similarities between $\mathcal{PV}$ and a special case of a construction due to Chu and presented in the appendix of [2] and Section 3 of [3]. (Readers unfamiliar with the Chu construction can skip this paragraph, as it will not be mentioned later.) Specifically, Chu's construction, applied to the cartesian closed category of sets and the object 2, yields a $*$-autonomous category in which the objects are the same as those of $\mathcal{PV}$ and the morphisms differ from those of $\mathcal{PV}$ only in that they are required to satisfy $A(f_-(b), a) \iff B(b, f_+(a))$ rather than just an implication from left to right. This apparently minor difference in the definition leads to major differences in other aspects of the category, specifically in the internal hom-functor and the tensor product. But see [19] for a framework that encompasses both $\mathcal{PV}$ and the Chu construction.

In the next few sections, we shall describe how $\mathcal{PV}$ arose in various contexts. Thereafter, we indicate how ideas that arise naturally in these contexts suggest new constructions in linear logic.

I thank Valeria de Paiva for her helpful comments on an earlier version of this paper.

Reductions of Search Problems

Much of the theory of computational complexity (e.g., [10]) deals with decision problems. Such a problem is specified by giving a set of instances together with a subset called the set of positive instances; the problem is to determine, given an arbitrary instance, whether it is positive. In a typical example, the instances might be graphs and the positive instances might be the 3-colorable graphs. In another example, instances might be boolean formulas and positive instances might be the satisfiable ones. A *(many-one) reduction* from one decision problem to another is a map sending instances of the former to instances of the latter in such a way that an instance of the former is positive if and only if its image is positive. Clearly, an algorithm computing such a reduction and an algorithm solving the latter decision problem can be combined to yield an algorithm solving the former.

There are situations in complexity theory where it is useful to consider not only decision problems but also search problems. A search problem is specified by giving a set of instances, a set of witnesses, and a binary relation between them; the problem is to find, given an instance, some witness related to it. For example, the 3-colorability decision problem mentioned above (given a graph, is it 3-colorable?) can be converted into the 3-coloring search problem (given a graph, find a 3-coloring). Here the instances are graphs, the witnesses are 3-valued functions on the vertices of graphs, and the binary relation relates each graph to its (proper) 3-colorings. Similarly, there is a search version of the boolean satisfiability problem, where instances are boolean formulas, witnesses are truth assignments, and the binary relation is the satisfaction relation. Notice that a search problem is just an object $\mathbf{A}$ of $\mathcal{PV}$, the set of instances being A_- and the set of witnesses A_+.

There is a reasonable analog of many-one reducibility in the context of search problems. A reduction of $\mathbf{B}$ to $\mathbf{A}$ should first convert every instance $b \in B_-$ of $\mathbf{B}$ to an instance $a \in A_-$ of $\mathbf{A}$ (just as for decision problems), and then, if a witness w related to a is given, it should allow us, using w and remembering the original instance b, to compute a witness related to b. Again, an algorithm computing such a reduction and an algorithm solving $\mathbf{A}$ can clearly be combined to yield an algorithm solving $\mathbf{B}$. Most known many-one reductions between NP decision problems [10] implicitly involve many-one reductions of the corresponding search problems.

Formally, a *reduction* therefore consists of two functions, $f_- : B_- \to A_-$ and $f_+ : A_+ \times B_- \to B_+$ such that, for all $b \in B_-$ and $w \in A_+$,

$$A(f_-(b), w) \implies B(b, f_+(w, b)).$$

This is nearly, but not quite, the definition of a morphism from $\mathbf{A}$ to $\mathbf{B}$. The difference is that in a morphism f_+ would have only w, not b, as its argument. Thus, morphisms amount to reductions where the final witness (for b) is computed from a witness w for $a = f_-(b)$ without remembering b. This notion of reduction has been used in the literature [15, 23], but I would not argue that it is as natural as the version where one is allowed to remember b.

These observations lead to a suggestion that we record for future reference.

Suggestion 1. *Find a natural place in the theory of $\mathcal{PV}$ for reductions as described above, i.e., pairs of functions that are like morphisms except that f_+ takes an additional argument from B_- and the implication relating f_- and f_+ is amended accordingly.*

A "dual" modification of the notion of morphism, allowing f_- to have an extra argument in A_+, occurred in de Paiva's work [17] on a categorial version of Gödel's Dialectica interpretation, work that preceded the introduction of $\mathcal{PV}$ in [18].

Linear Logic

The search problems (objects of $\mathcal{PV}$) and reductions (morphisms of $\mathcal{PV}$ or generalized morphisms as in Suggestion 1) described in the preceding section are vaguely related to some of the intuitions that underlie Girard's linear logic [11]. Girard has written about linear logic as a logic of questions and answers (or actions and reactions) [11, 12], so it seems reasonable to try to model this idea in terms of $\mathcal{PV}$. Also, the fact that in a many-one reduction of $\mathbf{B}$ to $\mathbf{A}$ a witness for $\mathbf{B}$ is produced from exactly one witness for $\mathbf{A}$ is reminiscent of the central idea of linear logic that a conclusion is obtained by using each hypothesis exactly once. In this section, we attempt to make these vague intuitions precise. Our goal here is to develop de Paiva's interpretation of linear logic (at least the multiplicative and additive parts; the exponentials will be discussed briefly later) in a step by step fashion that emphasizes the naturality or necessity of the definitions used.

We intend to use objects of $\mathcal{PV}$ as the interpretations of the formulas of linear logic. This corresponds to Girard's intuition that for any formula A there are questions and answers of type A. Of course, in addition to questions and answers, objects of $\mathcal{PV}$ also have a correctness relation between them. It is reasonable to expect that one formula linearly implies another, in a particular interpretation, if and only if there is a morphism in $\mathcal{PV}$ from (the object interpreting) the former to (the object interpreting) the latter; we shall see this more precisely later.

To produce an interpretation of linear logic, we must tell how to interpret the connectives, and we must define what it means for a sequent to be true in an interpretation.

Perhaps the easiest part of this task is to interpret the additive connectives, & and $\oplus$. It seems to be universally accepted [21] that a reasonable categorial model of linear logic will interpret these as the product and coproduct of the category. Fortunately, $\mathcal{PV}$ has products and coproducts, so we adopt these as the interpretations of the additive connectives. The result is that "with" is interpreted as

$$(A_-, A_+, A) \& (B_-, B_+, B) = (A_- + B_-, A_+ \times B_+, W),$$

where

$$W(x, (a, b)) \iff \begin{cases} A(x, a), & \text{if } x \in A_- \\ B(x, b), & \text{if } x \in B_-; \end{cases}$$

"plus" is interpreted as

$$(A_-, A_+, A) \oplus (B_-, B_+, B) = (A_- \times B_-, A_+ + B_+, V)$$

where

$$V((a,b),x) \iff \begin{cases} A(a,x), & \text{if } x \in A_+ \\ B(b,x), & \text{if } x \in B_+; \end{cases}$$

and the additive units are

$$\top = (\emptyset, 1, \emptyset) \quad \text{and} \quad 0 = (1, \emptyset, \emptyset),$$

where 1 represents any one-element set.

These definitions correspond reasonably well to the intuitive meanings of the additive connectives in terms of questions and answers or in terms of Girard's "action" description of linear logic [12]. To answer a disjunction $A \oplus B$ is to provide an answer to one of A and B; correctness means that, confronted with questions of both types, we answer one of them correctly (in the sense of A or B). To answer a conjunction $A \& B$ we must give answers for both, but we are confronted with a question of only one type and only our answer to that one needs to be correct. The intuitive discussion of conjunction, in particular the fact that we must give answers of both types even though only one will be relevant to the question, might make better sense if we think of the answer as being given before the question is known. This is a rather strange way of running a dialogue, but it will arise again later in other contexts (and I've seen examples of it in real life).

There is also a natural interpretation of linear negation, since (cf. [11, 12]) questions of type A are answers of type the negation $A^\perp$ of A and vice versa. We define

$$(A_-, A_+, A)^\perp = (A_+, A_-, A^\perp),$$

where

$$A^\perp(x,y) \iff \neg A(y,x).$$

So linear negation interchanges questions with answers and replaces the correctness relation by the complement of its converse. Perhaps a few words should be said about the use of the complement of the converse rather than just the converse. There are several reasons for this, perhaps the most intuitive being that we are, after all, defining a sort of negation. Another way to look at it is to think of a contest between a questioner and an answerer, where success for the questioner is defined to mean failure for the answerer (cf. the discussion of challengers and solvers in [13]). "That's a good question" often means that I have no good answer. For another indication that the given definition of $^\perp$ is appropriate, see the section on set-theoretic applications below.

Mathematically, the strongest reason for defining $^\perp$ as we did is that it gives a contravariant involution of the category $\mathcal{PV}$. That is, the operation $^\perp$ on objects and the operation on morphisms defined by $(f_-, f_+)^\perp = (f_+, f_-)$

constitute a contravariant functor from $\mathcal{PV}$ to itself, whose square is the identity. This corresponds to the equivalences in linear logic between $A \vdash B$ and $B^\perp \vdash A^\perp$ and between $A^{\perp\perp}$ and A. (Here it is important that our underlying universe of sets is classical. In more general situations of the sort considered in [18], $A^{\perp\perp}$ and A need not be equivalent.)

We turn now to a more delicate matter, the interpretation of the multiplicative connectives. We begin with "times." Girard's intuitive explanation of the difference between the multiplicative conjunction $\otimes$ and the additive conjunction & in [12] is that the former represents an ability to perform both actions while the latter represents an ability to do either one of the two actions (chosen externally). Looking back at the interpretation of &, we would expect to modify it by allowing questions of both sorts, rather than just one, and requiring both components of the answer to be correct. This operation on objects of $\mathcal{PV}$ is quite natural, and occurs in both [18] and [25]. De Paiva uses the notation $\otimes$ for it, although it is not the interpretation of Girard's connective $\otimes$ in her interpretation of linear logic in the categories **GC** of [18]. (It is the interpretation of $\otimes$ in her earlier interpretation of intuitionistic linear logic in categories **DC**, and the notation comes from there.) Vojtáš used the notation $\times$ even though it is not the product in the category. We shall use the notation $\overline{\otimes}$ and regard it as a sort of provisional tensor product. Formally, we define

$$(A_-, A_+, A)\overline{\otimes}(B_-, B_+, B) = (A_- \times B_-, A_+ \times B_+, A \times B),$$

where the relation $A \times B$ is defined by

$$(A \times B)((x, y), (a, b)) \iff A(x, a) \text{ and } B(y, b).$$

Of course, since we have already interpreted negation, our provisional $\otimes$ gives rise to a dual connective, the provisional "par":

$$(A_-, A_+, A)\overline{\invamp}(B_-, B_+, B) = (A_- \times B_-, A_+ \times B_+, P)$$

where

$$P((x, y), (a, b)) \iff A(x, a) \text{ or } B(y, b).$$

To see why these interpretations of the multiplicative connectives are only provisional and must be modified, we turn to the question of soundness of the interpretation. This requires, of course, that we define what is meant by a sequent being valid, which presumably depends on a notion of sequents being true in particular interpretations, i.e, with particular objects as values of the atomic formulas. For simplicity, we work with one-sided sequents, as in [11]. So a sequent is a finite list (or multi-set) of formulas, each interpreted as an object of $\mathcal{PV}$. Since a sequent is deductively equivalent in linear logic with the par of its members, we interpret the sequent as the (provisional) par of its members, i.e., as a certain object of $\mathcal{PV}$. So we must specify what we mean by truth of an object of $\mathcal{PV}$, and then we must try to verify the soundness of the axioms and rules of linear logic.

There are two plausible meanings for "truth" of $\mathbf{A} = (A_-, A_+, A)$, both saying intuitively that one can answer all the questions of type $\mathbf{A}$. The difference between the two is in whether the answer can depend on the question.

The first (provisional) interpretation of truth allows the answer to depend on the question, as one would probably expect intuitively.

$$\models_1 (A_-, A_+, A) \iff \forall x \in A_- \, \exists y \in A_+ \, A(x, y).$$

The second, stronger (provisional) interpretation is that one answer must uniformly answer all questions correctly.

$$\models_2 (A_-, A_+, A) \iff \exists y \in A_+ \, \forall x \in A_- \, A(x, y).$$

Before rejecting the second interpretation as unreasonably strong, one should note that the two interpretations are dual to each other in the sense that $\mathbf{A}$ is true in either sense if and only if its negation $\mathbf{A}^\perp$ is not true in the other sense. Furthermore, the second definition fits better with the idea that truth of a sequent $A \vdash B$ should mean the existence of a morphism from A to B. If we specialize to the case where A is the multiplicative unit 1, so that the sequent $A \vdash B$ becomes deductively equivalent (in linear logic) with $\vdash B$, and if we note that the unit for our provisional $\otimes$ is $(1, 1, \mathrm{true})$, then we see that truth of B should be equivalent to existence of a morphism from $(1, 1, \mathrm{true})$ to B. It is easily checked that existence of such a morphism is precisely the second definition of truth above. It might also be mentioned here that the second definition matches the ideas behind de Paiva's definition of the Dialectica categories in [17], for the Dialectica interpretation produces formulas of the form $\exists\forall$, not $\forall\exists$.

Finally, as we shall see in a moment, each definition has its own advantages and disadvantages when one tries to prove the soundness of linear logic, and eventually we shall need to adopt a compromise between them. The remark above about the relationship between $\models_1$, $\models_2$ and negation suggests that either version of $\models$, used alone, might have difficulties with the axioms $\vdash A, A^\perp$ (which say that linear negation is no stronger than it should be) or the cut rule (which says that linear negation is no weaker than it should be). Let us consider what happens if one tries to establish the soundness of the axioms and cut for either version of $\models$.

For the axioms, we wish to show that $\mathbf{A} \overline{\mathbin{\rotatebox[origin=c]{180}{\&}}} \mathbf{A}^\perp$ is true for each object $\mathbf{A}$ of $\mathcal{PV}$. In $\mathbf{A} \overline{\mathbin{\rotatebox[origin=c]{180}{\&}}} \mathbf{A}^\perp$, the questions are pairs (x, y) where $x \in A_-$ and

$y \in (A^\perp)_- = A_+$, and the answers are pairs (a, b) where $a \in A_+$ and $b \in (A^\perp)_+ = A_-$. The answer (a, b) is correct for the question (x, y) if and only if either $A(x, a)$ or $\neg A(b, y)$ (the latter being the definition of $A^\perp(y, b)$). Obviously, any question (x, y) is correctly answered by (y, x). So $\models_1 \mathbf{A}\overline{\parr}\mathbf{A}^\perp$. On the other hand, we do not in general have $\models_2 \mathbf{A}\overline{\parr}\mathbf{A}^\perp$, since an easy calculation shows that this would mean that in $\mathbf{A}$ either some answer is correct for all questions or some question has no correct answer. There are, of course, easy examples of $\mathbf{A}$ where this fails; the simplest is to take $A_- = A_+ = \emptyset$, and if one insists on non-empty sets then the simplest is $A_- = A_+ = \{1, 2\}$ with A being the relation of equality. So, for the soundness of the axioms, $\models_1$ works properly, but $\models_2$ does not.

Now consider the cut rule. We wish to show that, if $\mathbf{B}\overline{\parr}\mathbf{A}$ and $\mathbf{C}\overline{\parr}\mathbf{A}^\perp$ are true, then so is $\mathbf{B}\overline{\parr}\mathbf{C}$. If we interpret truth as $\models_2$, then this is easy. Suppose (b, x) correctly answers all questions in $\mathbf{B}\overline{\parr}\mathbf{A}$ and (c, y) correctly answers all questions in $\mathbf{C}\overline{\parr}\mathbf{A}^\perp$; we claim that (b, c) correctly answers all questions (p, q) in $\mathbf{B}\overline{\parr}\mathbf{C}$. Indeed, if (p, q) were a counterexample, then b is not correct for p and c is not correct for q, yet (b, x) is correct for (p, y) and (c, y) is correct for (q, x) (where the four occurrences of "correct" refer to $\mathbf{B}$, $\mathbf{C}$, $\mathbf{B}\overline{\parr}\mathbf{A}$, and $\mathbf{C}\overline{\parr}\mathbf{A}^\perp$, respectively). But then we must have, by definition of $\overline{\parr}$, that x correctly answers y in $\mathbf{A}$ and that y correctly answers x in $\mathbf{A}^\perp$. That is impossible, by definition of $^\perp$, so the cut rule preserves $\models_2$. Unfortunately, it fails to preserve $\models_1$. The easiest counterexamples occur when both $\mathbf{B}$ and $\mathbf{C}$ have questions with no correct answers (but B_+ and C_+ are non-empty). Then $\mathbf{B}\overline{\parr}\mathbf{C}$ is not true, so the soundness of the cut rule would require that at least one of $\mathbf{B}\overline{\parr}\mathbf{A}$ and $\mathbf{C}\overline{\parr}\mathbf{A}^\perp$ also fail to be true. That means that either $\mathbf{A}$ or its negation must have a question with no correct answer, i.e., in $\mathbf{A}$ either some answer is correct for all questions or some question has no correct answer. Since that is not the case in general, we conclude that the cut rule is unsound for $\models_1$.

Summarizing the preceding discussion, we have

(1) If we define truth allowing answers to depend on questions ($\models_1$), then the axioms of linear logic are sound but the cut rule is not.
(2) If we define truth requiring the answer to be independent of the question ($\models_2$), then the cut rule is sound but the axioms are not.

Fortunately, there is a way out of this dilemma. Consider the dependence of answers on questions that was needed to obtain the soundness of the axioms. At first sight, it is an extremely strong dependence; indeed, the answer (y, x) is, except for the order of components, identical to the question (x, y). But the dependence is special in that each component of the answer depends only on the *other* component of the question.

Rather surprisingly, this sort of cross-dependence also makes the cut rule sound. To see this, suppose that both $\mathbf{B}\overline{⅋}\mathbf{A}$ and $\mathbf{C}\overline{⅋}\mathbf{A}^{\perp}$ are true in this new sense. That is, there are functions $f : B_- \to A_+$ and $g : A_- \to B_+$ such that, for all $b \in B_-$ and $x \in A_-$,

$$B(b, g(x)) \quad \text{or} \quad A(x, f(b)), \tag{1}$$

and similarly there are $f' : C_- \to (A^{\perp})_+ = A_-$ and $g' : (A^{\perp})_- = A_+ \to C_+$ such that, for all $c \in C_-$ and all $y \in A_+$,

$$C(c, g'(y)) \quad \text{or} \quad \neg A(f'(c), y). \tag{2}$$

Then we claim that $g' \circ f : B_- \to C_+$ and $g \circ f' : C_- \to B_+$ satisfy, for all $b \in B_-$ and $c \in C_-$,

$$B(b, g(f'(c))) \quad \text{or} \quad C(c, g'(f(b))),$$

which means that $\mathbf{B}\overline{⅋}\mathbf{C}$ is true in the "cross-dependence" sense. To verify the claim, let such b and c be given. If $A(f'(c), f(b))$, then (2) implies $C(c, g'(f(b)))$. If $\neg A(f'(c), f(b))$, then (1) implies $B(b, g(f'(c)))$. So the claim is true in either case, and we have verified the soundness of the cut rule.

By allowing the answer in one component of a sequent to depend on the questions in the other components but not in the same component, this "cross-dependence" notion of truth makes crucial use of the commas in a sequent, to distinguish the components. But linear logic requires (by the introduction rules for times and especially for par) that the commas in a sequent behave exactly like the connective ⅋. So it seems necessary to build cross-dependence into the interpretation of this connective. This will lead to the correct definition of the multiplicative connectives, introduced by de Paiva [18], replacing the provisional interpretations given earlier.

The par operation on objects of $\mathcal{PV}$ is defined, as in [18], by

$$(A_-, A_+, A) ⅋ (B_-, B_+, B) = (A_- \times B_-, A_+^{B_-} \times B_+^{A_-}, P)$$

where

$$P((x, y), (f, g)) \iff A(x, f(y)) \text{ or } B(y, g(x)).$$

This operation ⅋ is the object part of a functor, the action on morphisms being $(f ⅋ g)_- = f_- \times g_-$ and $(f ⅋ g)_+ = f_+^{g_-} \times g_+^{f_-}$. It is easy to check that ⅋ is associative (up to natural isomorphism). In the par of several objects, questions are tuples consisting of one question from each of the objects, and answers are tuples of functions, each producing an answer in one component when given as inputs questions in all the other components.

We also interpret commas in sequents as the new $\parr$ (rather than $\overline{\parr}$). This change in the interpretation of the commas makes $\models_2$ behave like the cross-dependence notion of truth described earlier. To see this, note that $\models_2$ requires the existence of a single answer correct for all questions at once, but the new $\parr$ allows that answer to consist of functions whereby each component of the answer can depend on the other components of the question. We therefore adopt $\models_2$ as the (non-provisional) definition of truth, and from now on we write it simply as $\models$. The previous discussion shows that the axioms and the cut rule are sound. We sometimes refer to an answer that is correct for all questions in an object $\mathbf{A}$ as a *solution* of the problem $\mathbf{A}$. So truth means having a solution.

Of course, the new interpretation of par gives, by duality, a new interpretation of times.

$$(A_-, A_+, A) \otimes (B_-, B_+, B) = (A_-^{B_+} \times B_-^{A_+}, A_+ \times B_+, T),$$

where

$$T((f,g),(x,y)) \iff A(f(y),x) \text{ and } B(g(x),y).$$

(This connective was called $\oslash$ in [18].) The units for the multiplicative connectives are $1 = (1, 1, \text{true})$ and $\bot = (1, 1, \text{false})$, where true and false represent the obvious relations on a singleton. The linear implication $A \multimap B$ defined as $A^\bot \parr B$ is

$$(A_-, A_+, A) \multimap (B_-, B_+, B) = (A_+ \times B_-, A_-^{B_-} \times B_+^{A_+}, C),$$

where

$$C((x,y),(f,g)) \iff [A(f(y),x) \implies B(y, g(x))].$$

Notice that a solution of $\mathbf{A} \multimap \mathbf{B}$ is precisely a morphism $\mathbf{A} \to \mathbf{B}$ in $\mathcal{PV}$. This indicates that the definitions of the multiplicative connectives and of truth, though not immediately intuitive, are proper in the context of the category $\mathcal{PV}$.

The belief that these definitions are reasonable is reinforced by de Paiva's theorem [18] that the multiplicative and additive fragment of linear logic is sound for this interpretation. (Her theorem actually covers full linear logic, including the exponentials, but we have not yet discussed the interpretation of the exponentials.)

Linear logic is not complete for this interpretation. For one thing, the interpretation validates the mix rule: If $\mathbf{A}$ and $\mathbf{B}$ are both true, then so is $\mathbf{A} \parr \mathbf{B}$. Also, the interpretation satisfies all formulas of the form $A^\bot \parr (A \parr A)$, a special case of weakening. (General weakening, $A^\perp ⅋ (A ⅋ B)$ is not satisfied; for a counterexample, take B_+ to be empty while all of A_+, A_-, B_- are nonempty.)

The interpretations of $\otimes$ and $⅋$ remain, in spite of their success at modeling linear logic, rather unintuitive. This is attested by the fact that de Paiva [18], while using $\otimes$ to interpret the multiplicative conjunction, calls it $\oslash$ and reserves the symbol $\otimes$ for the more intuitive construction that I called $\overline{\otimes}$. Vojtáš [25] also discusses $\overline{\otimes}$, calling it $\times$, but never has any use for $\otimes$. Since $\overline{\otimes}$ seems much more natural than the "correct" $\otimes$, it should have its own place in the logic.

Suggestion 2. *Find a natural place in the theory of $\mathcal{PV}$ for the operation $\overline{\otimes}$.*

Digression on Game Semantics

At the suggestion of the editors, I give in this section a short survey of my previous work [6] on interpreting linear logic by games, and I suggest some connections between that interpretation and de Paiva's interpretation in the category $\mathcal{PV}$. This section can be skipped without loss of continuity.

The problem, which caused such difficulty in the preceding section, whether truth should require one uniform answer for all questions or (possibly) different answers for different questions ($\models_2$ versus $\models_1$) can be regarded as the problem whether the answer should be given before or after the question is known. One could regard an object of $\mathcal{PV}$ as an incomplete specification of a dialogue between the questioner and the answerer; the specification would be completed by saying who is expected to speak first. Once this idea is introduced, it is natural to consider more complicated dependences between what the two speakers say. For example, one can envision a question consisting of two components, the first of which is asked before the answer, the second after. (Cf. the intuitive description of sequential composition in the next section.) Such dialogues of length greater than two arise naturally when one interprets the additive connectives. The dialogue for $\mathbf{A}\&\mathbf{B}$ should begin with the questioner deciding which of $\mathbf{A}$ and $\mathbf{B}$ to discuss, and the dialogue for $\mathbf{A} \oplus \mathbf{B}$ should begin with a similar decision by the answerer. Dialogues for deeply nested iterations of the additive connectives should therefore involve long alternations between the two speakers.

Game semantics in the sense of [6] interprets the formulas of linear logic as games, i.e., as rules for conducting dialogues and for determining which of the speakers (called players in this context) "wins." (In the simple dialogues where each player speaks only once, the answerer wins if his answer is correct for his opponent's question, and the questioner wins otherwise.) Additive connectives are handled as described above, and linear negation simply interchanges the roles of the two players. Multiplicative connectives are more complicated. The dialogue for $\mathbf{A} \otimes \mathbf{B}$ consists of two interleaved dialogues, one for each of the constituents $\mathbf{A}$ and $\mathbf{B}$, with the questioner allowed to decide when to switch from one constituent to the other, and with the answerer required to win both constituents in order to win the $\otimes$ compound. $\mathbf{A} \parr \mathbf{B}$ is similar with the roles of the two players interchanged. The dialogue for $!\mathbf{A}$ consists of a potential infinity of interleaved dialogues for $\mathbf{A}$, with the questioner deciding when to switch from one to another, with the answerer required to win all constituents in order to win $!\mathbf{A}$, and with an additional coherence constraint on the answerer.

It is proved in [6] that affine logic, i.e., linear logic plus the rule of weakening, is sound for this semantics. There is a completeness theorem for the additive fragment (when dialogues are allowed to be infinitely long), but not for the multiplicative fragment. Subsequently, Abramsky and Jagadeesan [1] made substantial modifications in the game semantics to obtain completeness (and more) for multiplicative linear logic plus the mix rule, and further modifications by Hyland and Ong [14] eliminated the need for the mix rule. Unfortunately, these modifications no longer work as well with the additive connectives.

There is another way to connect game semantics and the question-answer approach formalized in $\mathcal{PV}$. Instead of regarding the questions and answers of the latter approach as individual moves of the players in a very short (two moves) game, one can regard the questions and answers as strategies for the two players in a longer game. This point of view may make the cross-dependence in de Paiva's interpretations of the multiplicative connectives more intuitive. If we think of A_+ as a set of strategies for the answerer in game $\mathbf{A}$ and B_- as a set of strategies for the questioner in $\mathbf{B}$, then the elements of $A_+^{B_-}$ are strategies for the answerer in $\mathbf{A}$ that take into account what the questioner is doing in $\mathbf{B}$. They take it into account unrealistically well, making use of the entire strategy of the questioner rather than just the moves already made by the questioner on the basis of this strategy, but at least this line of thought indicates that cross-dependence of the sort used to interpret $\otimes$ and $\parr$ in $\mathcal{PV}$ is not entirely unmotivated from the game semantical point of view.

Cardinal Characteristics of the Continuum

We begin this section by introducing a few (just enough to serve as examples later) of the many cardinal characteristics of the continuum that have been studied by set-theorists, topologists, and others. For more information about these and other characteristics, see [22] and the references cited there. All the cardinal characteristics considered here (and almost all the others) are uncountable cardinals smaller than or equal to the cardinality $\mathfrak{c} = 2^{\aleph_0}$ of the continuum. So they are of little interest if the continuum hypothesis ($\mathfrak{c} = \aleph_1$) holds, but in the absence of the continuum hypothsis there are many interesting connections, usually in the form of inequalities, between various characteristics. (There are also independence results showing that certain inequalities are not provable from the usual ZFC axioms of set theory.) Part of the work of Vojtáš [25, 26] on which this section is based can be viewed as a way to extract from the inequality proofs information which is of interest even if the continuum hypothesis holds.

Definitions. If X and Y are subsets of $\mathbb{N}$, we say that X *splits* Y if both $Y \cap X$ and $Y - X$ are infinite. The *splitting number* $\mathfrak{s}$ is the smallest cardinality of any family $\mathcal{S}$ of subsets of $\mathbb{N}$ such that every infinite subset of $\mathbb{N}$ is split by some element of $\mathcal{S}$. The *refining number* (also called the *unsplitting* or *reaping number*) $\mathfrak{r}$ is the smallest cardinality of any family $\mathcal{R}$ of infinite subsets of $\mathbb{N}$ such that no single set splits all the sets in $\mathcal{R}$. $\mathfrak{r}_\sigma$ is the smallest cardinality of any family $\mathcal{R}$ of infinite subsets of $\mathbb{N}$ such that, for any countably many subsets S_k of $\mathbb{N}$, some set in $\mathcal{R}$ is not split by any S_k.

These cardinals arise naturally in analysis, for example in connection with the Bolzano-Weierstrass theorem, which asserts that a bounded sequence of real numbers has a convergent subsequence. A straightforward diagonal argument extends this to show that, for any countably many bounded sequences of real numbers $\mathbf{x}_k = (x_{kn})_{n\in\mathbb{N}}$, there is a single infinite $A \subseteq \mathbb{N}$ such that the subsequences indexed by A, $(x_{kn})_{n\in A}$, all converge. If one tries to extend this to uncountably many sequences, then the first cardinal for which the analogous result fails is $\mathfrak{s}$. Also, $\mathfrak{r}_\sigma$ is the smallest cardinality of any family $\mathcal{R}$ of infinite subsets of $\mathbb{N}$ such that, for every bounded sequence $(x_n)_{n\in\mathbb{N}}$, there is a convergent subsequence $(x_n)_{n\in A}$ with $A \in \mathcal{R}$. There is an analogous description of $\mathfrak{r}$, where the sequences $(x_n)_{n\in\mathbb{N}}$ are required to have only finitely many distinct terms. For more information about these aspects of the cardinal characteristics, see [24].

Definitions. A function $f : \mathbb{N} \to \mathbb{N}$ *dominates* another such function g if, for all but finitely many $n \in \mathbb{N}$, $f(n) \leq g(n)$. The *dominating number* $\mathfrak{d}$ is the smallest cardinality of any family $\mathcal{D} \subseteq \mathbb{N}^\mathbb{N}$ such that every $g \in \mathbb{N}^\mathbb{N}$ is dominated by some $f \in \mathcal{D}$. The *bounding number* $\mathfrak{b}$ is the smallest cardinality of any family $\mathcal{B} \subseteq \mathbb{N}^\mathbb{N}$ such that no single g dominates all the members of $\mathcal{B}$.

The known inequalities between these cardinals (and $\aleph_1$ and $\mathfrak{c} = 2^{\aleph_0}$) are

$$\aleph_1 \leq \mathfrak{s} \leq \mathfrak{d} \leq \mathfrak{c},$$

$$\aleph_1 \leq \mathfrak{b} \leq \mathfrak{r} \leq \mathfrak{r}_\sigma \leq \mathfrak{c},$$

and

$$\mathfrak{b} \leq \mathfrak{d}.$$

It is known that any further inequalities between these cardinals are independent of ZFC, except that it is still an open problem whether $\mathfrak{r} = \mathfrak{r}_\sigma$ is provable.

The connection between the theory of these cardinals and the category $\mathcal{PV}$ discussed in previous sections becomes visible when one considers the proofs of some of these inequalities, so we shall prove the two non-trivial (but well known) ones, $\mathfrak{s} \leq \mathfrak{d}$ and $\mathfrak{b} \leq \mathfrak{r}$. (In each case, only the first of the two paragraphs in the proof is relevant to $\mathcal{PV}$, so the reader willing to take the first paragraph on faith can skip the justification in the second paragraph.)

Proof of $\mathfrak{s} \leq \mathfrak{d}$. There is a map $\alpha : \mathbb{N}^\mathbb{N} \to \mathcal{P}(\mathbb{N})$ sending every dominating family $\mathcal{D}$ (as in the definition of $\mathfrak{d}$) to a splitting family (as in the definition of $\mathfrak{s}$). In fact, one can associate to each infinite $X \subseteq \mathbb{N}$ a function $\beta(X) = f \in \mathbb{N}^\mathbb{N}$ such that, if g dominates f, then $\alpha(g)$ splits X.

Given g, to define $\alpha(g)$, partition $\mathbb{N}$ into a sequence of intervals $[0, a_1)$, $[a_1, a_2), \ldots$ such that, for each $n \in \mathbb{N}$, $g(n)$ is at most one interval beyond n (it's trivial to define such a_i's by induction), and let $\alpha(g)$ be the union of the even-numbered intervals. Define $\beta(X)$ to send each $n \in \mathbb{N}$ to the next element of X greater than n. If $f = \beta(X)$, if g dominates f, if a_i's are as in the definition of $\alpha(g)$, and if k is large enough, then the element $f(a_k - 1)$ of X lies in the interval $[a_k, a_{k+1})$. So X meets all but finitely many of the intervals $[a_k, a_{k+1})$ and is therefore split by $\alpha(g)$. $\square$

Proof of $\mathfrak{b} \leq \mathfrak{r}$. There is a function $\beta : \mathcal{P}_\infty(\mathbb{N}) \to \mathbb{N}^\mathbb{N}$ sending every unsplittable family $\mathcal{R}$ (as in the definition of $\mathfrak{r}$) to an undominated family (as in the definition of $\mathfrak{b}$). In fact, one can associate to each $g \in \mathbb{N}^\mathbb{N}$ a set $\alpha(g) = Y \in \mathcal{P}(\mathbb{N})$ such that, if Y does not split X then g does not dominate $\beta(X)$.

The same α and β as in the preceding proof will work, as the properties required of them here are logically equivalent to the properties required there. $\square$

In the notation of the preceding sections, the pair (β, α) in the first of these proofs is a morphism in $\mathcal{PV}$ from $(\mathbb{N}^\mathbb{N}, \mathbb{N}^\mathbb{N}, \text{is majorized by})$ to $(\mathcal{P}_\infty(\mathbb{N}), \mathcal{P}(\mathbb{N}), \text{is split by})$. In the second proof, we used the image of this under $^\perp$, namely that (α, β) is a morphism from $(\mathcal{P}(\mathbb{N}), \mathcal{P}_\infty(\mathbb{N}), \text{does not split})$ to $(\mathbb{N}^\mathbb{N}, \mathbb{N}^\mathbb{N}, \text{does not majorize})$. In both cases, the cardinal inequality follows from the following general fact. Define for each object $\mathbf{A}$ of $\mathcal{PV}$ the *norm* $\|\mathbf{A}\|$ as the smallest cardinality of any set $X \subseteq A_+$ of answers sufficient to contain at least one correct answer for every question in A_- (undefined if there is no such set, i.e., if some question has no correct answer, i.e., if $\mathbf{A}^\perp$ is true). Then the existence of a morphism $f : \mathbf{A} \to \mathbf{B}$ implies that $\|\mathbf{A}\| \geq \|\mathbf{B}\|$, because f_+ sends any set of the sort required in the definition of $\|\mathbf{A}\|$ to one as required for $\|\mathbf{B}\|$. (What I called the norm of $\mathbf{A}$ is, in Vojtáš's notation [25, 26] $\mathfrak{d}(A)$; Vojtáš's $\mathfrak{b}(A)$ is $\|\mathbf{A}^\perp\|$.)

It is an empirical fact that proofs of inequalities between cardinal characteristics of the continuum usually proceed by representing the characteristics as norms of objects in $\mathcal{PV}$ and then exhibiting explicit morphisms between those objects. This fact is explicit in Vojtáš's [25, 26] and implicit in [9]. It applies even to trivial inequalities like $\mathfrak{b} \leq \mathfrak{d}$ (where the required morphism from $(\mathbb{N}^\mathbb{N}, \mathbb{N}^\mathbb{N}, \text{is dominated by})$ to $(\mathbb{N}^\mathbb{N}, \mathbb{N}^\mathbb{N}, \text{does not dominate})$ consists of identity maps on both components) as well as to inequalities much deeper than the examples proved above; see for example the presentation in [9] of Bartoszyński's theorem [4] that the smallest number of meager sets whose union is not meager is at least as large as the corresponding number for "measure zero" in place of "meager."

It is tempting to regard the existence of a morphism $\mathbf{A} \to \mathbf{B}$ as a strong formulation of the inequality $\|\mathbf{A}\| \geq \|\mathbf{B}\|$ that is significant even in the presence of the continuum hypothesis (which makes inequalities between cardinal characteristics trivial as these cardinals lie between $\aleph_1$ and $\mathfrak{c}$ inclusive). The situation is, however, not quite so simple. My student, Olga Yiparaki, has shown that, in the presence of the continuum hypothesis (or certain weaker assumptions), there are morphisms in $\mathcal{PV}$ in both directions between any two objects that correspond (as in [25, 26]) to cardinal characteristics of the continuum. Those morphisms, however, are highly non-constructive, whereas those used in the usual proofs of cardinal inequalities are quite explicit. It therefore seems likely that a strengthening of these cardinal inequalities that retains its significance in the presence of the continuum hypothesis is to require not merely the existence of morphisms but the existence of "nice" morphisms, say ones whose components are Borel mappings.

The linear negation defined on $\mathcal{PV}$ gives a precise version of an intuitive "duality" in the theory of cardinal characteristics. In that theory, one often refers to the cardinals $\|\mathbf{A}\|$ and $\|\mathbf{A}^\perp\|$ as being dual to each other; see for example the introduction to [16]. On cardinals, this is not well defined, for two objects can have the same norm while their negations have different norms, but it is the shadow, in the world of cardinals, of the (well defined) linear negation in $\mathcal{PV}$. It may be worth noting in this connection that $(\mathcal{P}(\mathbb{N}), \mathcal{P}_\infty(\mathbb{N}), \text{does not split})$, whose norm is $\mathfrak{r}$, and $(\mathcal{P}(\mathbb{N})^\mathbb{N}, \mathcal{P}_\infty(\mathbb{N}), \text{has no component that splits})$, whose norm is $\mathfrak{r}_\sigma$, have negations both of norm $\mathfrak{s}$.

In addition to inequalities of the sort discussed above, which relate two cardinal characteristics of the continuum, there are a few theorems that relate three (occasionally even four) of them. We consider one relatively easy example here, since it leads to an idea that should connect to linear logic. The example concerns Ramsey's theorem [20], which asserts (in a simple form) that, whenever the set $[\mathbb{N}]^2$ of two-element subsets of $\mathbb{N}$ is partitioned into two pieces, then there is an infinite $H \subseteq \mathbb{N}$ that is homogeneous in the sense that all its two element subsets lie in the same piece of the partition. The cardinal $\mathfrak{hom}$ was defined in [7] as the smallest cardinality of a family $\mathcal{H}$ of infinite subsets of $\mathbb{N}$ such that, for every partition of $[\mathbb{N}]^2$ as in Ramsey's theorem, a homogeneous set can be found in $\mathcal{H}$. It was shown in [7] that this cardinal is bounded below by $\max\{\mathfrak{r}, \mathfrak{d}\}$ and above by $\max\{\mathfrak{r}_\sigma, \mathfrak{d}\}$. The lower bound amounts to two ordinary inequalities, $\mathfrak{hom} \geq \mathfrak{r}$ and $\mathfrak{hom} \geq \mathfrak{d}$, both of which were proved by exhibiting morphisms between the appropriate objects of $\mathcal{PV}$. The upper bound genuinely relates three cardinals, and we wish to make some comments about its proof, so we begin by sketching the proof.

Proof of $\mathfrak{hom} \leq \max\{\mathfrak{r}_\sigma, \mathfrak{d}\}$. Fix a family $\mathcal{R}_0$ of $\mathfrak{r}_\sigma$ subsets of $\mathbb{N}$ such that no countably many sets split all the sets in $\mathcal{R}_0$. Within each set $A \in \mathcal{R}_0$, fix a family $\mathcal{R}_1(A)$ of $\mathfrak{r}$ sets such that no single set splits them all. Also, fix a family $\mathcal{D}$ of functions dominating all functions $\mathbb{N} \to \mathbb{N}$. For each $A \in \mathcal{R}_0$, for each $B \in \mathcal{R}_1(A)$, and for each $f \in \mathcal{D}$, choose a subset $Z = Z(A, B, f)$ of B so thin that, if $x < y$ are in Z then $f(x) < y$. We claim that the family $\mathcal{H}$ of all these Z's, which clearly has cardinality $\max\{\mathfrak{r}_\sigma, \mathfrak{d}\}$ (since $\mathfrak{r} \leq \mathfrak{r}_\sigma$), contains almost homogeneous sets for all partitions of $[\mathbb{N}]^2$ into two parts. "Almost homogeneous" means that the set becomes homogeneous when finitely many of its elements are removed. Since we can close $\mathcal{H}$ under such finite changes without increasing its cardinality, the claim completes the proof.

To prove the claim, let $[\mathbb{N}]^2$ be partitioned into two parts. For each natural number n let C_n consist of those x for which $\{n, x\}$ is in the first part. By choice of $\mathcal{R}_0$, it contains a set A unsplit by any C_n. Let $g(n)$ be so large that all $x \in A$ with $x \geq g(n)$ have $\{n, x\}$ in the same piece of the partition, and let Q be the set of n for which this is the first piece. Choose $B \in \mathcal{R}_1(A)$ unsplit by Q and $f \in \mathcal{D}$ dominating g. It is then easy to check that $Z(A, B, f)$ is almost homogeneous for the given partition. $\square$

To discuss this proof in terms of $\mathcal{PV}$, we introduce the natural objects of $\mathcal{PV}$ whose norms are the cardinals under consideration. For mnemonic purposes, we name each object with the capital letter corresponding to the lower-case letter naming the cardinal.

$$\mathbf{HOM} = (\{p \mid p : [\mathbb{N}]^2 \to 2\}, \mathcal{P}_\infty(\mathbb{N}), AH),$$

where AH is the relation of almost homogeneity: $AH(p, H)$ means that H is almost homogeneous for the partition p.

$$\mathbf{D} = (\mathbb{N}^\mathbb{N}, \mathbb{N}^\mathbb{N}, \text{is dominated by}).$$

$$\mathbf{R} = (\mathcal{P}(\mathbb{N}), \mathcal{P}_\infty(\mathbb{N}), \text{does not split}).$$

$$\mathbf{R}_\sigma = (\mathcal{P}(\mathbb{N})^{\mathbb{N}}, \mathcal{P}_\infty(\mathbb{N}), \text{has no component that splits}).$$

The structure of the preceding proof is then as follows. From a question p in **HOM**, we first produced a question $(C_n)_{n\in\mathbb{N}}$ in $\mathbf{R}_\sigma$. Using an answer A to this question and also using again the original question p, we produced questions g in **D** and Q in **R**. From answers f and B to these questions, along with the previous answer A, we finally produced an answer H to the original question p in **HOM**.

This can be described as a morphism into **HOM** from a suitable combination of **D**, **R**, and $\mathbf{R}_\sigma$, but the relevant combination is a bit different from what we have considered previously. The part of the construction involving **D** and **R** is just the provisional tensor product $\overline{\otimes}$; that is, we had a question (g, Q) in $\mathbf{D}\overline{\otimes}\mathbf{R}$ and we obtained an answer (f, B) for it. (Strictly speaking, we used a version of **R** on A rather than on $\mathbb{N}$, but we shall ignore this detail.) The novelty is in how $\mathbf{D}\overline{\otimes}\mathbf{R}$ is combined with $\mathbf{R}_\sigma$. For what we produced from p was a question in $\mathbf{R}_\sigma$ together with a function converting answers to this question into questions in $\mathbf{D}\overline{\otimes}\mathbf{R}$. This thing that we produced ought to be a question in the object that is being mapped to **HOM**. An answer in that object ought to be what we used in order to get the answer H for **HOM**, namely (A, B, f).

Motivated by these considerations, we define a connective, denoted by a semi-colon (to suggest sequential composition), as follows.

$$(A_-, A_+, A); (B_-, B_+, B) = (A_- \times B_-^{A_+}, A_+ \times B_+, S),$$

where

$$S((x, f), (a, b)) \iff A(x, a) \text{ and } B(f(a), b).$$

Thus, a question of sort $\mathbf{A};\mathbf{B}$ consists of a first question in **A**, followed by a second question in **B** that may depend on the answer to the first. A correct answer consists of correct answers to both of the constituent questions. Thus, sequential composition can be viewed as describing a dialogue in which the questioner first asks a question in **A**, is given an answer, selects on the basis of this answer a question in **B**, and is given an answer to this as well.

The proof of $\mathfrak{hom} \le \max\{\mathfrak{r}_\sigma, \mathfrak{d}\}$ exhibits a morphism from $\mathbf{R}_\sigma; (\mathbf{D}\overline{\otimes}\mathbf{R})$ to **HOM**. The cardinal inequality follows from the existence of such a morphism, since one easily checks that the operations on infinite cardinal norms corresponding to the operations $\overline{\otimes}$ and ; are both simply max.

The sequential composition of objects of $\mathcal{PV}$ occurs repeatedly in the proofs of inequalities relating three cardinal characteristics. A typical example is the proof that the minimum number of meager sets of reals with a non-meager union is the smaller of $\mathfrak{b}$ and the minimum number of meager sets that cover the real line [16, 5, 9]. Vojtáš [25] describes the strategy for proving such three-way relations between cardinals in terms of what he calls a max-min diagram. This diagram amounts exactly to a morphism from the sequential composition of two objects to a third object. In other words, sequential composition is the reification of the max-min diagram as an object of $\mathcal{PV}$.

Sequential composition also seems a natural concept to add to linear logic from the computational point of view. Linear logic is generally viewed as a logic of parallel computation, but even parallel computations often have sequential parts, so it seems reasonable to include in the logic a way to describe sequentiality. These ideas are not yet sufficiently developed to support any claims about sequential composition, as defined in the $\mathcal{PV}$ model, being the (or a) right way to do this. In addition to semantical interpretations, one would certainly want good axioms governing any sequential composition connective that is to be added to linear logic, and one would hope that some of the pleasant proof theory of linear logic would survive the addition. Much remains to be done in this direction.

Suggestion 3. *Find a place for sequential composition (specifically for the connective called ; above) in linear logic and the theory of $\mathcal{PV}$.*

Generalized Multiplicative Connectives

The previous sections have led to three suggestions of natural connectives to add to linear logic. (Actually, Suggestion 1 concerned not a connective but a modified notion of morphism. But such a modification should correspond to a reinterpretation of $\multimap$ and therefore of $\parr$ and $\otimes$ as well.) The suggested new connectives are all analogous to the multiplicatives in that both the set of questions and the set of answers are cartesian products. (For the additive connectives, one of the two sets was a disjoint union.) The factors in these products are either sets of questions or answers from the constituent objects or else sets of functions, from questions to answers or vice versa. With these preliminary comments, it seems natural to describe general multiplicative conjunctions (cousins of $\otimes$) as follows.

A *general multiplicative conjunction* operates on n objects $\mathbf{A}_1 \dots \mathbf{A}_n$ of $\mathcal{PV}$ to produce an object $\mathbf{C}$, where $C_+ = A_{1+} \times \dots A_{n+}$ and where C_- consists of n-tuples (f_i) of functions where f_i maps some product of A_{j+}'s into A_{i-}. Which A_{j+}'s occur in the domain of which f_i's is given by the specification of the particular connective. An answer (a_i) is correct for a question (f_i) if each a_i correctly answers in $\mathbf{A}_i$ the question obtained by evaluating f_i at the relevant a_j's.

For example, $\otimes$ is a generalized multiplicative conjunction, for which

$n = 2$ and each f_i has domain A_j for the j different from i. Similarly, we obtain $\overline{\otimes}$ if the domains of the f_i's are taken to be empty products (i.e., singletons); no j is relevant to any i. Sequential composition is obtained by having f_1 depend on no arguments while f_2 has an argument in A_1.

Dual (via $^\perp$) to generalized multiplicative conjunctions are generalized multiplicative disjunctions. Here the answers are allowed to depend on some questions, rather than vice versa (exactly which dependences are allowed is the specification of a particular connective), and correctness means correctness in at least one component, rather than in all.

To avoid possible confusion, we stress that the generalization of the multiplicative connectives proposed here is quite different from that proposed by Danos and Regnier [8]. The Danos-Regnier multiplicatives can correspond to many different classical connectives, whereas mine correspond only to conjunction and disjunction. One could, of course, consider combining the two generalizations, but we do not attempt this here.

There are non-trivial *unary* conjunction and disjunction connectives. The conjunction is given by

$$\kappa(A_-, A_+, A) = (A_-^{A_+}, A_+, \kappa A),$$

where

$$\kappa A(f, a) \iff A(f(a), a).$$

The dual disjunction is

$$\alpha(A_-, A_+, A) = (A_-, A_+^{A_-}, \alpha A)$$

where

$$\alpha A(a, f) \iff A(a, f(a)).$$

These operations were called T and R in [18].

The modified concept of morphism from $\mathbf{A}$ to $\mathbf{B}$ in Suggestion 1, where f_+ maps $A_+ \times B_-$, rather than just A_+, into B_+, amounts to a morphism (in the standard $\mathcal{PV}$ sense) from $\mathbf{A}$ to $\alpha\mathbf{B}$. This concept of morphism thus gives rise to the Kleisli category of $\mathcal{PV}$ with respect to the monad α. (We have defined α only on objects, but it is routine to define it on morphisms and to describe its monad structure.)

De Paiva's Dialectica category [17] built over the category of sets has as morphisms $\mathbf{A} \to \mathbf{B}$ the $\mathcal{PV}$ morphisms $\kappa\mathbf{A} \to \mathbf{B}$. It is dual (via $^\perp$) to the category in the preceding paragraph and is the co-Kleisli category of the comonad κ (see [18, Prop. 7]).

The connective α also provides a way to reinstate the notion of truth $\models_1$ that was discarded when we replaced the provisional $\otimes$ and $⅋$ with the final versions. Indeed, $\models_1 \mathbf{A}$ holds if and only if $\models \alpha\mathbf{A}$.

Exponentials

Girard has pointed out that the exponential connectives or modalities, ! and ?, unlike the other connectives, are not determined by the axioms of linear logic. More precisely, if one added to linear logic a second pair of modalities, say $!'$ and $?'$, subject to the same rules of inference as the original pair, then one could not deduce that the new modalities are equivalent to the old. Several versions of the exponentials could coexist in one model of linear logic.

$\mathcal{PV}$ provides an example of this phenomenon. De Paiva [18] gave an interpretation of the exponentials in which ! is a combination of the unary conjunction κ defined above and a construction S where multisets m of questions are regarded as questions and a correct answer to m is a single answer that is correct for all the questions in m. (Neither κ nor S alone can serve as an interpretation of !.) Another interpretation of the exponentials in $\mathcal{PV}$, validating the exponential rules of linear logic, is given by

$$!(A_-, A_+, A) = (1, A_+, U)$$

where 1 is a singleton, say $\{*\}$ and

$$U(*, a) \iff \forall x \in A_- \, A(x, a),$$

and its dual

$$?(A_-, A_+, A) = (A_-, 1, E)$$

where

$$E(a, *) \iff \exists x \in A_+ \, A(a, x).$$

Intuitively, a question of type $!\mathbf{A}$ (namely $*$) amounts to all questions of type $\mathbf{A}$; a correct answer in $!\mathbf{A}$ must correctly answer all questions in $\mathbf{A}$ simultaneously.

It is easy to check that Girard's rules of inference for the exponentials are sound for this simple interpretation.

References

1. S. Abramsky and R. Jagadeesan, *Games and full completeness for multiplicative linear logic*, J. Symbolic Logic (to appear).
2. M. Barr, *$*$-Autonomous Categories*, Lecture Notes in Mathematics 752, Springer-Verlag, 1979.
3. M. Barr, *$*$-Autonomous categories and linear logic*, Math. Struct. Comp. Sci. **1** (1991), 159–178.
4. T. Bartoszyński, *Additivity of measure implies additivity of category*, Trans. Amer. Math. Soc. **281** (1984), 209–213.
5. T. Bartoszyński and H. Judah, *Measure and Category — The Asymmetry* (to appear).

6. A. Blass, *A game semantics for linear logic*, Ann. Pure Appl. Logic **56** (1992), 183–220.
7. A. Blass, *Simple cardinal characteristics of the continuum*, Set Theory of the Reals (H. Judah, ed.), Israel Math. Conf. Proc. 6, 1993, pp. 63–90.
8. V. Danos and L. Regnier, *The structure of multiplicatives*, Arch. Math. Logic **28** (1989), 181–203.
9. D. H. Fremlin, *Cichoń's diagram*, Séminaire Initiation à l'Analyse (G. Choquet, M. Rogalski, and J. Saint-Raymond, eds.), Univ. Pierre et Marie Curie, 1983/84, pp. (5-01)–(5-13).
10. M. R. Garey and D. S. Johnson, *Computers and Intractability*, W. H. Freeman and Co., 1979.
11. J.-Y. Girard, *Linear logic*, Theoret. Comp. Sci. **50** (1987), 1–102.
12. J.-Y. Girard, *Toward a geometry of interaction*, Categories in Computer Science and Logic (J. W. Gray and A. Scedrov, eds.), Contemp. Math. 92, Amer. Math. Soc., 1989, pp. 69–108.
13. Y. Gurevich, *The challenger-solver game: variations on the theme of P=?NP*, Bull. Europ. Assoc. Theoret. Comp. Sci. **39** (1989), 112–121.
14. J. M. E. Hyland and C.-H. L. Ong, *Fair games and full completeness for multiplicative linear logic without the MIX-rule*, preprint (1993).
15. R. Impagliazzo and L. Levin, *No better ways to generate hard NP instances than picking uniformly at random*, Symposium on Foundations of Computer Science, IEEE Computer Society Press, 1990, pp. 812–821.
16. A. Miller, *Additivity of measure implies dominating reals*, Proc. Amer. Math. Soc. **91** (1984), 111–117.
17. V. C. V. de Paiva, *The Dialectica categories*, Categories in Computer Science and Logic (J. W. Gray and A. Scedrov, eds.), Contemp. Math. 92, Amer. Math. Soc., 1989, pp. 47–62.
18. V. C. V. de Paiva, *A Dialectica-like model of linear logic*, Category Theory and Computer Science (D. H. Pitt, D. E. Rydeheard, P. Dybjer, A. Pitts, and A. Poigné, eds.), Lecture Notes in Computer Science 389, Springer-Verlag, 1989, pp. 341–356.
19. V. C. V. de Paiva, *Categorical multirelations, linear logic, and Petri nets*, Tech. Report 225, University of Cambridge Computer Laboratory, 1991.
20. F. P. Ramsey, *On a problem of formal logic*, Proc. London Math. Soc. (2) **30** (1930), 264–286.
21. R. A. G. Seely, *Linear logic, $*$-autonomous categories, and cofree algebras*, Categories in Computer Science and Logic (J. W. Gray and A. Scedrov, eds.), Contemp. Math. 92, Amer. Math. Soc., 1989, pp. 371–382.
22. J. Vaughan, *Small uncountable cardinals and topology*, Open Problems in Topology (J. van Mill and G. Reed, eds.), North-Holland, 1990, pp. 195–218.
23. R. Venkatesan and L. Levin, *Random instances of a graph coloring problem are hard*, Symposium on Theory of Computing, Assoc. for Computing Machinery, 1988, pp. 217–222.
24. P. Vojtáš, *Set-theoretic characteristics of summability of sequences and convergence of series*, Comment. Math. Univ. Carolinae **28** (1987), 173–183.
25. P. Vojtáš, *Generalized Galois-Tukey connections between explicit relations on classical objects of real analysis*, preprint (1991).
26. P. Vojtáš, *Topological cardinal invariants and the Galois-Tukey category* (to appear).

Mathematics Dept., University of Michigan, Ann Arbor, MI 48109, U.S.A.
E-mail address: `ablass@umich.edu`