Entailed Ranking Arguments

Alan Prince

Department of Linguistics *and*
Center for Cognitive Science
Rutgers University/New Brunswick
• Draft of February 23, 2002 •
ROA-500, http://roa.rutgers.edu

ABSTRACT

An 'elementary ranking condition' (ERC) embodies the kind of restrictions imposed by a comparison between a desired optimum and a single competitor. All entailments between elementary ranking conditions can be ascertained through three simple formal rules; one of them introduces a method of argument combination, *fusion*, shown to have the same sense as in relevance logic. Fusion is also central to detecting inconsistency in a set of ERCs; inconsistency and entailment are closely related here, much as in ordinary logic. Fusion therefore plays a key role in the definition of Recursive Constraint Demotion (RCD: Tesar & Smolensky 1994, 1998). When ERCs are hierarchized by the ranking of the constraints that crucially evaluate them, their entailment and fusional relations are seen to correlate with aspects of ranking structure. RCD and the Minimal Stratified Hierarchy it produces also figure prominently in an efficient procedure for calculating entailments. Harmonic bounding, both simple and collective, leads to the existence of entailment relations, and removal of entailment dependencies from a set of ERCs eliminates harmonic bounding in its underlying candidate set. The logic of entailment in OT is seen to be the implication-negation fragment of RM (Sobociński 1952, Parks 1972) and the logic of OT in general is shown by a semantical argument to be precisely RM itself. When the logic is extended from ERCs to constraints, it allows for a direct representation of the notion of a strict domination hierarchy using only the connectives of the logic; various ranking restrictions are shown to follow when logical relations exist between constraints.

Acknowledgments

Thanks to Vieri Samek-Lodovici, Seth Cable, Paul Smolensky, Bruce Tesar, and Graham Horwood for valuable comments on previous partial versions of these remarks; to Matthew Stone for a helpful pointer to the logic literature, to S.-J. Leslie for useful P.C. discussion; to Jane Grimshaw and John McCarthy for discussion of fundamental background issues. This work was supported at various points by a fellowship from the John Simon Guggenheim Memorial Foundation, by NSF Grant BCS-0083101, and by the generosity of the Faculty of Arts and Sciences, Rutgers University/New Brunswick. General thanks to all those whose enthusiasm for this project has buoyed my own; special thanks to Dick Foley, most undeanly dean of deans; and deepest gratitude to Bridget Gellert Lyons, *il coltello più affilato nel cassetto*. The author retains responsibility for errors and the like.

Entailed Ranking Arguments

Contents

Prefatory

> Little given to abstract ideas, we accept things as they are and we attempt to the maximum of our ability to protect ourselves against delusions about realities. – Metternich

Herein, some brief remarks on the usefulness of logic; the structure of OT; and the stylistic and typographical conventions employed in this paper.

Utilities

From the questions pursued here, and the answers arrived at, one can extract tools for the construction, correction, compacting, expansion, and analysis of optimality-theoretic arguments. By means of these, harmonically-bounded candidates can be discerned and eliminated, crucial rankings identified, redundancies eliminated, consequences determined, constraint systems rapidly analysed for consistency with data, and even (at the margins of applicability) related constraints ranked by their intrinsic logical properties. It may be that the necessary aspects of these functions are second nature to many; but perhaps not to all.

Beyond such immediate applications, it is to be hoped that the connections found here between Optimality Theory and the relevance logic RM will lead to a better understanding of OT, and even perhaps of RM.

In presenting the results, I have felt constrained to depart from the familiar canons and tropes of linguistic exposition, moving toward a more euclidean mode. Although this leads in places to a certain dessication of the narrative, the alternatives are harder to live with. For the most part, there is simply no way to find out whether some thesis holds, or even what it is, really, without attempting definition and proof.

A primer of Optimality Theory

Given a 'candidate set' of alternatives and a set of constraints bearing on their desirability, Optimality Theory defines the sense in which certain of the alternatives *best satisfy* a prioritization of the constraint set. The basic idea is that a relation of ranking, or 'strict domination', holds between constraints, written $C_i >> C_j$. For any total ranking order on the constraint set, or 'constraint hierarchy', an alternative ω is *optimal* iff for every competitor *x*, the highest-ranked constraint distinguishing ω from *x* decides in favor of ω.

Following Samek-Lodovici & Prince 1999 [SLP], we conceive of a constraint as a function which takes a set as an argument and a returns a subset thereof. Each constraint is associated with an order on candidate sets: it returns the maximal elements in that order.

Each such order is presupposed to be a *stratified partial order*, where by 'stratified' we mean that if elements x,y are noncomparable in the order, they have the same order relations to other elements. A set so ordered is partitioned into 'strata'; elements in the same stratum are not ordered with respect to each other, but are each ordered identically with respect to members of other strata.

We will write $x \approx y$ for 'x and y belong the same order stratum'. Concisely, a partial order is stratified iff noncomparability is an equivalence relation.

A constraint hierarchy is a composition of constraint functions. The ranking order on the constraint set corresponds to the order of composition of the constraint functions, with highest-ranked being first applied. This formalizes the intuitive sense that a constraint hierarchy progressively winnows the candidate set, each constraint eliminating suboptimal elements from the set it receives, and then passing on the reduced set to its next-lower-ranked neighbor. In SLP:44ff., §4.1, it is observed that constraints and constraint hierarchies so defined have certain properties – for example, 'downward inheritance of optimality': if ω is optimal in a set K, then it is also optimal in any $K' \subseteq K$ to which it belongs. Composition of constraint functions is shown to *induce* a stratified partial order on the candidate set (SLP:53). This suggests that we might reconstruct the notion of constraint as a kind of function that has certain properties, and derive its associated order as a concomitant.[1] The following conditions will achieve this goal:

Let U be the set of all candidates; a constraint C is a function meeting the following restrictions:

Def. Constraint function. A *constraint* is a function $C:\wp(U) \rightarrow \wp(U)$ that satisfies

(i) Choice. $C(X) \subseteq X$.

(ii) Forced Choice. $C(X)=\emptyset \Rightarrow X=\emptyset$

(iii) Contextual Independence of Choice. If $Y \cap C(X) \neq \emptyset$, then $C(Y \cap X)=Y \cap C(X)$.

Condition (i) means that the optimum is among the set of alternative candidates; condition (ii) requires that a choice be made. Condition (iii) ensures 'downward inheritance of optimality', from this implicit subclause:

(iiia) $Y \cap C(X) \subseteq CY$.

In addition, it ensures what we might call 'upward inheritance of equivalence'

(iiib) $C(Y) \subseteq Y \cap CX$ if $Y \cap C(X) \neq \emptyset$

By (iiib), if $Y \subseteq X$ contains any members of C(X), then all of C(Y) must be in C(X). This correlates with the notion that members of a stratum share order relations.

Following and extending SLP:32, ex. (63), which is aimed at whole hierarchies rather than their constituent constraints, we define an order relation as follows, given such a C:

Def. Order associated with a constraint. $x >_C y$ iff $C(\{x,y\}) = \{x\}$

It is demonstrated in Appendix 1 below, p.102, that when a function $C:\wp(U) \rightarrow \wp(U)$ meets conditions (i)-(iii), the induced relation '$>_C$' yields a stratified partial order on U; and further that C returns exactly the maximal elements in its argument. It is additionally shown that any composition of functions satisfying (i)-(iii) also satisfies (i)-(iii). This means that a constraint hierarchy is also a constraint (SLP:38).

---

[1] Besnard, Fanselow, and Schaub (2001) independently re-work the SLP *constraint* in a similar fashion (thanks to Jane Grimshaw for bringing this work to my attention).

Given a hierarchy of constraints, *i.e.* a functionally-composed sequence of constraint functions, an optimal element in a candidate set K is therefore a maximal element in the order induced on K by the hierarchy.

Because constraint functions output sets, it also makes sense to speak of their 'intersection':

Def. Intersection of Constraints. $f \cap g(X) =_{df} f(X) \cap g(X)$

Constraint conflict occurs when different orders of composition yield different results: when $f \circ g(X) \neq g \circ f(X)$. From SLP:43, §4.1, the Favoring Intersection Lemma, we know that when constraints do not conflict, their composed output (in any order) is equal to their intersection.

• If $f \circ g(X) = g \circ f(X)$, then $f \circ g(X) = f \cap g(X)$.

This allows us to broaden the definition of an OT grammar to include composition *and* intersection of constraint functions. Under this broader definition, the grammar is itself a stratified hierarchy of constraints, where each stratum is an interection of its constraints, and the strata are composed to yield the hierarchy. If we allow *free* intersection, then the grammar is no longer guaranteed to be a 'constraint' in the sense defined above, since it will fail Forced Choice whenever $f \cap g(X) = \emptyset$ for nonnull X. Intersection of conflicting constraints yields null output, quite possibly a useful state of affairs. Standardly, however, and throughout this paper, constraints in the same stratum are assumed to be nonconflicting, i.e. to have nonnull intersections.

A Boolean constraint system, in which the chosen output must satisfy all constraints, is then just a monostratal grammar, where intersection is the only means of constraint combination. Introducing composition is the equivalent of introducing constraint ranking. Of particular importance is the stratified hierarchy that is closest to the Boolean model, the *Minimal Stratified Hierarchy*, which has as few constraints as possible in the scope of other constraints. Precise definitions will be offered below from a couple of different points of view (§4, §8).

Our focus of interest is almost entirely order-theoretic, and therefore we will be primarily concerned the decisions made by constraint functions as they apply to sets consisting of just two candidates. For any constraint function f, there are just three possible outcomes on pair input:

$f(\{x,y\}) = \{x\}$	$x >_C y$
$f(\{x,y\}) = \{y\}$	$y >_C x$
$f(\{x,y\}) = \{x,y\}$	$x \approx_C y$

Taking these relations as atoms of description leads to a three-valued logic that forms the basis of our investigation. It will be explicitly examined in §7 and §8.

In linguistic applications, Optimality Theory provides a way of relating two sets of linguistic representations, X and Y, by defining the sense in which, given $x \in X$, some element $y \in Y$ is matched to it by virtue of best satisfying the set of constraints bearing on the $\{x\} \times Y$ relation. Each single candidate, then, is an ordered pair (x,y), which we may notate as $x \rightarrowtail y$. The issue at hand, for each constraint and for the hierarchy as a whole, is whether the linguistic relation or 'mapping' $x \rightarrowtail y$ is better than, worse than, or indistinguishable from every other possible $x \rightarrowtail z$, for $z \in Y$. If there is no better alternative, then $x \rightarrowtail y$ is optimal.

The issue of optimality can be resolved by examining *pairs* of candidates, in each of which the desired optimum $x \mapsto y$ is compared against a single alternative of the form $x \mapsto z$; optimality for $x \mapsto y$ requires that it never lose in any of these competitions. We will write [a~b] for the results of such a pairwise face-off over the constraint set, where a,b each denote a entity of the form $x_i \mapsto y_j$. Each such [a~b] collects the individual responses of each constraint to the comparison of a with b. It therefore provides an 'elementary ranking condition' (ERC) that determines which rankings of the entire constraint set must hold in order for candidate a to survive comparison with candidate b. For much of our discussion, we will treat such [a~b] as atomic and we will focus on relations between them; they and the ERCs they denote will form the units over which the logic of OT develops.

Sylistica & Typographica

Virtually all sets in sight are finite and this fact will not be noted case-by-case. All logical systems involved are complete, and no effort will be made to maintain notational separation between formal deducibility and validity. Among assertions, I distinguish between remarks, lemmas, propositions, and corollaries. The intention is that the meatiest theses find their way into propositions, to which lemmas lead and from which corollaries follow. Propositions are numbered sequentially within the major section where they occur: thus, Proposition 8.3 is the third proposition in §8. The others, along with definitions, are only given example numbers, though corollaries are linked by title to their governing propositions.

The following acronyms are used:

ERC	elementary ranking condition
OT	Optimality Theory
MSH	minimal stratified hierarchy
PC	propositional or predicate calculus
S	the logic of Sobociński 1952
VS	the vectorized version of S
RM	R-mingle
OP	ordered polyvaluation
AB	Anderson & Belnap 1975
ABD	Anderson, Bellnap, & Dunn
SLP	Samek-Lodovici and Prince 1999
FW	*Finnegans Wake*

I have not been careful about the use/mention distinction, though mentioned items are enclosed in single quotes from time to time, in the interests of clarity. In the text below, the sign '∘' is used only in its relevance-logic sense of 'fusion' and never indicates function-composition. The end of a proof is marked by the symbol □.

0. Preliminaries

Summary: Elementary ranking conditions can be represented as vectors whose coordinate entries take one of three values: W, L, *e*. Directly tied to the logic of the theory, such vectors support the key analytic computations on ranking arguments.

A *ranking argument* over a constraint set involves in the simplest case a comparison between a desired optimum q and a single suboptimal competitor *z*.[2] Any such comparison partitions a constraint set Σ into three disjoint subsets:

(1) **Win, lose, or draw**

W Those constraints *preferring* q to *z*, i.e. those on which q does better than *z*:
- all $C \in \Sigma$ such that $q >_C z$.

L Those constraints preferring *z* to q, i.e those on which z does better than q:
- all $C \in \Sigma$ such that $z >_C q$.

E Those constraints on which q and *z* are noncomparable, *i.e.* do equally well:
- all $C \in \Sigma$ such that $q \approx_C z$ (*i.e.* neither $q >_C z$ nor $z >_C q$).

We tabulate the comparison as in Prince 2000, marking by 'W' any constraint preferring the desired optimum, by 'L' any constraint preferring the desired suboptimum, and using blankness to indicate lack of preference.

(2) **Comparative Tableau**. $C_1 \in$E, $C_2 \in$W, $C_3 \in$L. $q >_{C2} z$. $z >_{C3} q$.

	C_1	C_2	C_3	...
[q ~ *z*]		W	L	...

More concisely: if we settle on a fixed but arbitrary listing of the constraint set under consideration, we can write out the tableau row as a vector with coordinate entries W,L,*e*, with the last standing for 'blank'. Writing Â for the vector derived from (2), we have Â=(*e*,W,L,...).

Each row-vector $\hat{A}_i$ gives rise to an *elementary ranking condition* (ERC), here denoted A_i , which depends only on the contents of W_i and L_i. The condition always takes the same general form: every constraint assessing L – perversely preferring the desired *sub*optimum – must be dominated by some constraint assessing W. (Cf. the Cancellation/Domination Lemma of Prince & Smolensky 1993:148.)

(3) **Elementary Ranking Condition** (∀∃ form). With W,L as in (1):

Every constraint $C \in$L is dominated by some constraint $D \in$W.

$$\forall C \in \Sigma \;\; \exists D \in \Sigma \;\; [C \in \mathrm{L} \supset (D \in \mathrm{W} \wedge D \gg C)]$$

[2] Known as a 'mark-data pair' in Tesar 1995 *et seq.* The notion will be generalized in §2 below.

Under the standard assumption that '>>' is a total order, this is equivalent to following:[3]

(4) **Elementary Ranking Condition** (∃∀ form).

Some constraint D∈W dominates every constraint C∈L.

$\exists D\in\Sigma \quad \forall C\in\Sigma \ [C\in L \supset (D\in W \wedge D>>C)]$

In what follows, we will work with ∃∀ form, and by the term 'elementary ranking condition' or 'ranking argument' we will specifically mean a logical expression of type (4). By 'ERC vector' we will mean a representation such as (e,W,L). Concisely, a ranking argument A_i over a constraint set Σ can be notated as $A_i = \langle W_i,L_i\rangle$,[4] which fixes the free parameters in the expression (4), where $W_i,L_i\subseteq\Sigma$ and $W_i\cap L_i=\emptyset$. We assume that the constraint set Σ is always nonempty, though of course either or both of W_i and L_i may be empty.

It is useful to note the properties of certain special configurations. If L_i is empty, A_i is vacuously *true* under any ranking, because for any assignment to the variable C the antecedent in (4) is false. In this case the tableau-row vector $\hat{A}_i$ consists entirely of W's and/or *e*'s; it imposes no ranking conditions. Let us denote the set of all such vectors by $\mathcal{W}^*$. If L_i is nonempty, but W_i is empty, then A_i is *false*, and no ranking of the constraint set can satisfy it, or satisfy any set of arguments containing it. Let us denote by $\mathcal{L}^+$ the set of vectors containing at least one L and no W's. (The '*' are '+' are chosen for mnemonic relation to the Kleene regular-language operators of the same names.)

If $A_i = \langle W_i,L_i\rangle$ is such that either W_i or L_i is empty, we will say that A_i is ***trivial***. Otherwise, we will say that A_i is ***nontrivial***. If both W_i and L_i are empty, we will say that A_i is ***degenerate*** and we denote it by δ. The degenerate ERC is always true.

Since constraint sets are finite, it might be thought excessively florid to *quantify* over them in stating the logical form of a ranking argument. Universal quantification over a finite set is merely conjunction, existential quantification merely disjunction. In the case of nontrivial ERCs, it is true that the given definitions boil down to a conjunction of disjunctions (∀∃) and disjunction of conjunctions (∃∀). Thus, the definitions (3) and (4) become, for $W_i\in W$ and $L_j\in L$:

(5) **Elementary ranking conditions** using disjunction and conjunction

a. ∀∃: $\wedge_j\vee_i (W_i>>L_j)$ *i.e.* $[W_1>>L_1 \vee W_2>>L_1 \vee ...] \wedge [W_1>>L_2 \vee W_2>>L_2 ...] \wedge ...$

b. ∃∀: $\vee_i\wedge_j (W_i>>L_j)$ *i.e.* $[W_1>>L_1 \wedge W_1>>L_2 \vee ...] \vee [W_2>>L_1 \wedge W_2>>L_2 ...] \vee ...$

But not every comparative vector yields this kind of expression – trivial vectors do not. An element of $\mathcal{L}^+$ offers no constraint to place in dominant position in a term $C_i>>C_j$ and an element of $\mathcal{W}^*$ offers nothing for the subordinate position. To sustain the isomorphism between comparative row-vectors and ERCs, we therefore insist on the quantifiers.

---

[3] By logic, we have (4)→(3), since $\exists x\forall yP(x,y)\rightarrow\forall y\exists xP(x,y)$. Now assume (3): by total ordering, some member of W dominates all the others, and by virtue of that must also dominate everything in L.

[4] This expression makes it clear that we are near to dealing with the "polarity" of relevance logic semantics (Dunn 1986:189). In §7, we explore the logical status of our procedures and representations.

An analyst or learner deals with a collection of ranking arguments: the goal is to find a ranking of the constraint set Σ that satisfies them all simultaneously, or show that no such ranking exists.[5] We will refer to a (total) ranking of Σ as a *model,* and a total ranking of Σ satisfying a set of arguments $\mathcal{A}$ will be termed 'a model of $\mathcal{A}$'. A nontrivial ERC – one corresponding to a vector containing both W and L – is guaranteed to be true in at least one model; and also to be false in at least one.

For any collection of ranking arguments, it may well be that some proper subset is sufficient to determine the rankings that satisfy the collection. All other arguments are then redundant and eliminable, in that they are logically entailed by one or more arguments belonging to the determining subset. Beyond considerations of efficiency, eliminating redundancies in argument ought to lead to increased insight into the workings of the constraint system.

It is also often the case in practice that the analyst or learner is in possession of a set of ranking arguments from which further requirements follow, but only with the application of (a sometimes tortured) logic. Finding such entailments may be essential for discovering the necessary rankings. For example, given arguments φ:(W,L,W) and ψ:(e,W,L), we must conclude that $C_1 \gg C_2$, despite the fact that φ only commits to the disjunctive necessity of $C_1 \gg C2$ *or* $C_3 \gg C_2$. But reaching the desired conclusion, even in this simple case, uses the following kind of argument:

> From φ we know that *either* $C_1 \gg C_2$ *or* $C_3 \gg C_2$. From ψ, we know that $C_2 \gg C_3$. This is incompatible with the second disjunct of φ. Therefore the first disjunct of φ must hold, and we have $C_1 \gg C_2$.

When more than two arguments are involved, or when each one has disjunctive conditions within it, as is commonly the case, the complexity of the deduction grows rapidly. Consider a small modification of the above example: from φ:(W,L,W) and ψ:(W,W,L) taken jointly, it follows that $C_1 \gg C_2$ and $C_1 \gg C_3$ must both hold. The argument runs like this:

> From φ we know that *either* $C_1 \gg C_2$ *or* $C_3 \gg C_2$. From ψ, we know that *either* $C_1 \gg C_3$ *or* $C_2 \gg C_3$. Conjoining these arguments produces four cases to consider. Of these, we may dismiss one immediately as inconsistent: '$C_3 \gg C_2$ *and* $C_2 \gg C_3$'. The remaining three cases are [1] $C_1 \gg C_2$ *and* $C_1 \gg C_3$, [2] $C_1 \gg C_2$ *and* $C_2 \gg C_3$, [3] $C_3 \gg C_2$ *and* $C_1 \gg C_3$. Observe that by transitivity of '>>', case [2] is equivalent to [2'] $C_1 \gg C_2$ *and* $C_2 \gg C_3$ *and* $C_1 \gg C_3$. Similarly, case [3] is equivalent to [3'] $C_3 \gg C_2$ *and* $C_1 \gg C_3$ *and* $C_1 \gg C_2$. Now observe that [1], [2'], and [3'] all contain the expression [1]. Therefore, since $p \vee (p \wedge q) \vee (p \wedge r)$ is equivalent to p, we have [1]. QED.

A particularly interesting entailment, which may also be imperspicuous, is that a set of arguments is inconsistent. Analysts and learners both need to be able to recognize this state of affairs, for it indicates that some hypothesis must be changed or abandoned.

In §§1-2, we examine the conditions under which such entailments obtain, and find three operations on row vectors – L-retraction, W-extension, and fusion – which are sufficient to give the entailment relation in full. We then determine in §3 the condition ('W-compliance') under which a set of ERCs may be exactly replaced with a single ERC via fusion.

---

[5] More specifically, the learner's task is simply to find a ranking that works, while the analyst (ever mindful of explanation and the possibilities of improving the theory) will want to determine the set of *necessary* ranking relations that every model must respect: those that are entailed by ERC set.

With these notions in hand, the examples just reviewed come out like this:

•First example: (W,L,W) and (e,W,L) fuse to (W,L,L), which gives (W,L,e) by L-retraction.

•Second example: (W,L,W) and (W,W,L) fuse to (W,L,L), an equivalent of the original pair.

In §4 relations are examined between entailment, fusion, and the Minimal Stratified Hierarchy (MSH) produced by Recursive Constraint Demotion (RCD: Tesar 1995, Tesar & Smolensky 1994, 1998, 2000). Fusion, though conjunction-like, is in general not as algebraically well-behaved as conjunction, because a fusion does not entail its fusands the way a conjunction entails its conjuncts. But good order is restored in the MSH. When ERCs are ranked according to the rank of the constraints that satisfy them, it emerges that the *fusion* of an ERC set of always lies at least as high as any individual argument in the set; and further, that any entailed argument lies at least as high as any member of the minimal set of arguments that entails it. This finding lays the groundwork for an efficient procedure, relying on the inconsistency-detecting powers of RCD, that determines which arguments in a set are entailed by others (§5).

In §6, the notion of harmonic bounding is explicated in terms of the behavior of ERC vectors, the hallmark of bounding being the existence of certain trivial vectors, members of $\mathcal{L}^+$ when the bounded item is placed in desired-optimum position. Bounding, simple or collective, occurs when and only there is an ERC set that fuses to $\mathcal{L}^+$. Bounding induces entailment relations, and it is shown that any set of ERCs that is free of entailments among its members (exclusive of $\varphi \vdash \varphi$!) is also free of bounding relations among the candidate set from which the ERCs are drawn.

The broader logical system in which these inferences take place is examined in §7. Supplementing fusion with a form of negation, as in §2, leads to a full logical system analogous to the propositional calculus, with e supplying a third truth-value in addition to T (W) and F (L), and with fusion acting like (a weakened form of) conjunction. This logic is identified as the three-valued logic of Sobociński 1952, which has been shown to be the implication-negation fragment of RM (Anderson & Belnap 1975; Parks 1972), a full-blown logic which includes the familiar logical connectives as well as those of the fusion family. A semantical argument based on the work of Meyer 1975 establishes that RM itself is the logic of optimality theory, in the strong sense that each constraint ranking corresponds to an RM valuation and vice versa. Logical expressions involving ERCs are valid iff they are valid as RM expressions. The notion of an (OT) 'system' is introduced – a set of RM valuations derived from the totality of ranking permutations of a set of constraints – and is shown to behave much like the entire RM model space.

When the same logic is applied to relations between constraints rather than ranking arguments (§8), it is found that its connectives allow for a direct expression of the notion of a strict domination hierarchy. Certain properties of the relation between constraint logic and constraint ranking are then established. The logical discussion leads to a starkly finitistic arithmetic representation of strict domination (§9).

Before we step into the main argument, it is useful to point out the tripartite nature of the entity we are focusing on. An ERC A based on the comparison of a with b is a logical expression that defines the ranking conditions under which a certain ordering relation, $a \succeq b$, holds in the candidate set. The ERC vector Â collects and organizes the constraint-performance data on which the ERC is based. Since each A corresponds biuniquely to an Â and since A holds of a ranking iff the order relation $a \succeq b$ holds of the candidate set, we will not always avoid the temptation to identify these notions, though we will aim for distinctness when it matters.

1. L-retraction and W-extension

Summary: All nontrivial ERCs entailed by any single nontrivial ERC are obtained by adding W's or taking away L's in its vectorial representation.

To begin with, consider two distinct ranking arguments A_1 and A_2 that are alike except that A_1 has additional constraints in its set of optimum-antagonists L_1 , so that $\mathsf{L}_2 \subseteq \mathsf{L}_1$.

	...	C_i	C_j	C_k	C_m
$\hat{A}_1$	...	W	L	L	
$\hat{A}_2$	...	W	L		

In this case, argument A_1 is strictly stronger and more informative than A_2. We can properly say that argument A_1 *entails* A_2, since every model of A_1 is also a model of A_2. For this relation, we will write $A_1 \vdash A_2$.[6] Specifically: A_1 says that some constraint $D \in \mathsf{W}_1 = \mathsf{W}_2$ dominates all of L_1, and this D must also dominate all of L_2, because $\mathsf{L}_2 \subseteq \mathsf{L}_1$. Whenever A_1 holds of some ranking, A_2 holds as well.

Now consider ranking arguments A_1 and A_2 that are alike except that the optimum-*preferring* set W_2 for A_2 has additional constraints in it, so that $\mathsf{W}_1 \subseteq \mathsf{W}_2$.

	...	C_i	C_j	C_k	C_m
$\hat{A}_1$	...	W	L		
$\hat{A}_2$	...	W	L	W	

Once again, it must be that $A_1 \vdash A_2$. When A_1 holds of a ranking, some member D of W_1 dominates all of $\mathsf{L}_1 = \mathsf{L}_2$. Clearly this D also belongs to $\mathsf{W}_2 \supseteq \mathsf{W}_1$ and therefore also guarantees the validity of A_2.

These two observations may be combined and strengthened by dropping the likeness condition in favor of simultaneous subsetting in the appropriate directions. Suppose A_1 and A_2 are such that $\mathsf{L}_2 \subseteq \mathsf{L}_1$ and $\mathsf{W}_1 \subseteq \mathsf{W}_2$. Then $A_1 \vdash A_2$, and A_2 is, in essence, a sub-argument or weakened version of A_1. This means that A_2 can be dropped from (or added to) any set $\mathcal{A}$ of ranking arguments to which A_1 belongs, without affecting the consequences of $\mathcal{A}$.

It is also true that if A_i is nontrivial, in the sense that neither W_i nor L_i is empty, then *all* nontrivial elementary ranking arguments that logically follow from A_i have these subset/superset relations with respect to A_i .

---

[6] Blurring the distinction between formal deducibility ($\vdash$) and semantic entailment ($\vDash$) which are at any rate equivalent for the wffs we are dealing with.

(6) **Proposition 1.1**. **Strength of Individual Arguments.**

Given two elementary ranking conditions A_1, A_2, defined as in (4),

(a) if $L_2 \subseteq L_1$ and $W_1 \subseteq W_2$, then $A_1 \vdash A_2$.

(b) Assume A_1, A_2 nontrivial. If $A_1 \vdash A_2$, then $L_2 \subseteq L_1$ and $W_1 \subseteq W_2$.

Pf. Claim (a). Following the line of argument in the text, from $A_1 = \langle W_1, L_1 \rangle$ we have some $D_1 \in W_1$ that dominates every $C \in L_1$. By assumption, $D_1 \in W_2$ and since $L_2 \subseteq L_1$, also by assumption, it follows that D_1 dominates every $C \in L_2$. This validates $A_2 = \langle W_2, L_2 \rangle$.

Claim (b). Proof of the contrapositive: failure of either consequent conjunct yields failure of the antecedent. We proceed by constructing models (rankings) in which A_1 is true but A_2 is false, demonstrating the failure of $A_1 \vdash A_2$, which requires that all models of A_1 are also models of A_2.

Assume first $\neg(L_2 \subseteq L_1)$. Then there is a $C \in L_2$ with $C \notin L_1$. A_1 does not demand that C be subordinated, so there is a model R of A_1 in which C is top-ranked. (Note that the assumed nontriviality of A_1 means that it has some such model.) But in every model of A_2, this C must be dominated (by some $D \in W_2$). Therefore A_2 is false in R. Because there is a model of A_1 that is not a model of A_2, $A_1 \vdash A_2$ fails.

Now assume $\neg(W_1 \subseteq W_2)$. Then there is some $D \in W_1$ with $D \notin W_2$. But there is a model of A_1 in which this very D is top-ranked. If $D \in L_2$, then A_2 is false in that model, and we are done. If $D \notin L_2$, then there is a model R of A_1 in which D is top-ranked and all of L_2 is ranked in an uninterrupted sequence immediately below it (in any ranking order). But A_2 is false in R, because no element of W_2 dominates all of (indeed, any of) L_2. (Note that L_2 is guaranteed to be nonempty, by the assumption of nontriviality.) Again, we have a model in which A_1 is true, but A_2 is false, showing that it is not the case that $A_1 \vdash A_2$. □

The role of the nontriviality assumption can be seen in the following valid entailments which do not meet the subsetting requirements on L and W.

(7) **A trivial entailment**

$\hat{A}_1$	L	
$\hat{A}_2$	W	L

Here $A_1 \vdash A_2$, because A_1 is always false in any ranking (both rankings) of the two constraints; yet it is not the case that $L_2 \subseteq L_1$.

(8) **Another**

$\hat{A}_2$	W	L
$\hat{A}_3$		W

Here $A_2 \vdash A_3$, because A_3 is always true under any ranking; yet it is not the case that $W_2 \subseteq W_3$. By contrast, *non*trivial ERCs are neither true in every model, nor true in none.

Proposition 1.1 gives rise to two rules of inference that can be used to manipulate a tableau-row vector to produce its entailments. The first we can call 'W-extension' — a blank cell may be filled with *W* to produce an entailed argument. The second we will call 'L-retraction' — a cell containing *L* may be replaced with a blank one.

(9) **W-extension**. $e \rightarrow$W.

If $\hat{A}_1$ and $\hat{A}_2$ are identical, except that the i^{th} coordinate of $\hat{A}_1$ is *e* and the i^{th} coordinate of $\hat{A}_2$ is *W*, then $A_1 \vdash A_2$.

(10) **L-retraction**. L$\rightarrow e$.

If $\hat{A}_1$ and $\hat{A}_2$ are identical, except that the i^{th} coordinate of $\hat{A}_1$ is *L* and the i^{th} coordinate of $\hat{A}_2$ is *e*, then $A_1 \vdash A_2$.

W-extension and L-retraction, when applied to nontrivial ERCs, instantiate familiar properties of propositional logic. W-extension adds a disjunct to an ERC, as licensed by the familiar tautology $p \rightarrow p \vee q$ (known as "disjunction introduction or "or-in" when treated as a rule of inference). L-retraction takes away a conjunct according to the scheme $p \wedge q \rightarrow q$ ("and-out" and 'conjunction elimination' name the cognate rule of inference).

These rules dissect the subsetting conditions of Proposition 1.1. Strikingly, by Proposition 1.1b, *all* nontrivial ERCs that follow from any single nontrival ERC may be generated by repeated application of the two procedures; indeed, by a sequence, possibly null, of L-retractions followed by a sequence, possibly null, of W-extensions. Compiling the results of all possible interactions, we see that the following coordinatewise relations are licensed:

(11)

$[\hat{A}_1]_k$		$[\hat{A}_2]_k$
L	$\rightarrow$	L, *e*, W
e	$\rightarrow$	*e*, W
W	$\rightarrow$	W

It is a natural step to explicitly recognize a special subspecies of entailment as being defined by adherence to just these rules; only the relations between certain trivial ERCs would fall outside its purview. We pursue this point in §7 below.

2. Fusion

Summary: Fusion of ERCs, written $\varphi\circ\psi$, extends direct computation of consistency and entailment to *sets* of arguments . A fusion is entailed by the set of its fusands: $\{\varphi,\psi\}\vdash\varphi\circ\psi$. The converse is not generally true, but fusion still computes all ERC consequences, because a slightly weaker, converse-like result does hold. If $\{\varphi,\psi\}\vdash\omega$, then ω is entailed individually either by φ, by ψ, or by $\varphi\circ\psi$. More generally, any ERC entailed by a set Ψ of ERCs follows from a *single* ERC that is the fusion of some subset of Ψ.

Ranking arguments come in sets, over which joint satisfaction is required: we must characterize not only the consequences of single isolated arguments, but also the further consequences that arise when ranking arguments are *conjoined*. These consequences emerge from the order-theoretic properties of ">>" — that it is asymmetric and transitive. For example, given constraints A,B,C, the conjunction of 'A>>B' and 'B>>C' yields the further elementary ranking condition 'A>>C', which follows from neither in isolation. Similarly, the arguments 'A>>B' and 'B>>A' cannot be satisfied together, although each is individually unproblematic.

A statement like 'A>>B ∧ B>>C ' is not of the $\langle$W,L$\rangle$ form and is therefore *not* an ERC: it cannot be expressed as a single row in a tableau. Since our goal is to determine the redundancies and equivalences within a set of ranking arguments, it is important that our calculations yield results within the space of ERCs. Notice too that the most fundamental questions about ranking structure, which may yet be difficult to dig out of a tangle of ranking arguments. – e.g., does C_j *necessarily* dominate C_k ?– have the ERC form. Here we advance a method for constructing an ERC '$\varphi\circ\psi$' from two given ERCs φ and ψ, such that any consequence of the set $\{\varphi,\psi\}$ follows from one of the formulas φ, ψ, and $\varphi\circ\psi$. (The nontrivial consequences follow through W-extension and L-retraction.) We will call $\varphi\circ\psi$ the *fusion* of φ and ψ, abandoning the term 'summation' of Prince 2000:§6. As seen in §7, 'summation' of OT comparative vectors corresponds to the operation of fusion in the logics RM3 and RM (originally discussed *en passant* in Sobociński 1952:52; *cf.* Parks 1972), and we adopt both the standard terminology and the notation from relevance logic (Anderson & Belnap 1975 [AB] *et seq*.). NB: the sign '∘' does not indicate function-composition in this context.

The intuitive idea is that two rows of a tableau can be combined to produce a third that is entailed jointly by the original pair. The mode of combining entries is as follows:

(12) **Def. Fusion of Entries**. For X = W, L, *e*,

X∘X	= X	*Idempotence*
X∘L	= L∘X = L	*Dominance of L*
X∘*e*	= *e*∘X = X	*e is Identity*

Fusion, like conjunction and disjunction is commutative (built into the definition) and associative.[7] The operation can be extended coordinate-wise to the fusion of vectors. The following defines the

[7] Sketch of proof. We want X∘(Y∘Z)=(X∘Y)∘Z. If any one of X,Y,Z is L, then both rhs and lhs must be L. If all are *e*, then both sides are *e*. If X,Y,Z contain at least one W and no L, then both sides are W.

coordinates of the fused vector in terms of the fusion of the entries at each coordinate of the constituent vectors. We write $[\hat{A}]_k$ for the k^{th} coordinate of $\hat{A}$.

(13) **Def. Fusion of vectors by coordinate**. For any $\hat{A}_1, \hat{A}_2$, the vector $\hat{A}_1 \circ \hat{A}_2$ is given by

$$[\hat{A}_1 \circ \hat{A}_2]_k =_{df} [\hat{A}_1]_k \circ [\hat{A}_2]_k$$

A typical outcome can be seen in the following example:

	C_1	C_2	C_3
$\hat{A}_1$	W	L	
$\hat{A}_2$		W	L
$\hat{A}_1 \circ \hat{A}_2$	W	L	L

In the more concise notation, we have $\hat{A}_1 \circ \hat{A}_2 = (W,L,e) \circ (e,W,L) = (W \circ e, L \circ W, e \circ L) = (W,L,L)$, fusing the vectors coordinate-wise.

Notice that by L-retraction $\hat{A}_1 \circ \hat{A}_2 = (W,L,L)$ yields the argument (W,e,L), i.e. $C_1 \gg C_3$. Thus we have in this case successfully produced the key results of the conjunctive non-ERC '$A_1 \wedge A_2$' without venturing outside the ERC domain.

The fusion of a pair (or more generally, a set) of ranking arguments creates a new ranking argument with valuable logical properties. The familiar sort of ranking argument is based on the comparison of a single optimum with a single suboptimum; fusion allows us to generalize to a simultaneous comparison of an optimum with several competitors, or several optima with several competitors, deriving conditions required to handle all such competitions simultaneously in one grammar.

Fusion is 'truth-functional' at the coordinate level, because the value of a fused coordinate is uniquely predictable from the value – W, L, or *e* – of its components. But at the level of vectors, such simple truth functionality no longer holds. Consider an expression like $(W,e) \circ (e,L) = (W,L)$. The first fusand is true in every model, the second in none, but the truth of the fusion is model-dependent. Thus, within a given model, the fact that A is true and B is false does not suffice to predict whether $A \circ B$ is true or false. The same phenomenon shows up with nontrivial vectors as well: $(W,L,e) \circ (e,W,L) = (W,L,L)$ is true in any model where C_1 is top-ranked, even if $C_3 \gg C_2$, which falsifies the second fusand. These examples illustrate a key divergence between fusion and conjunction: while $A \wedge B \vdash A$ holds absolutely, it is not the case that $A \circ B \vdash A$ under all conditions.[8]

Our overarching goal is to show that fusion, aided by L-retraction and W-extension, allows us to reach every elementary ranking condition implied by a given ERC set. We will find that for any such condition entailed by an ERC set $\mathcal{A}$, there is a subset of $\mathcal{A}$ whose fusion entails that condition (Proposition 2.5 below). In determining the consequences of $\mathcal{A}$, we are able to escape from logical

[8] The conditions on A,B under which $A \circ B \vdash A$ does hold are found in §3, (37), p.19.

conjunction plus tortuous applications of the distributive law and attendant manipulations, replacing them with deduction from a single ERC (the fusion of a subset) using only L-retraction and W-extension as rules of inference. Since, as noted above, the kind of truly atomic ranking conditions that we are ultimately interested in – 'C_i>>C_j' – are also ERCs of a particularly simple form, we have lost nothing by staying within the realm of ERCs. If we can obtain all entailed ERCs, we can find out what we need to know about the rankings consistent with a set of ranking arguments.

We proceed by accumulating results that are relevant to the desired conclusion or to other arguments down the line. First we show that the fusion operation behaves properly in general, producing an entailed condition. Extending the notation, we write '$A_1 \circ A_2$' for the ranking *condition* that is associated with the fused vector $\hat{A}_1 \circ \hat{A}_2$.

(14) **Lemma**. Let A_1, A_2 be elementary ranking conditions. Then $\{A_1, A_2\} \vdash A_1 \circ A_2$.

Pf. From $A_1 = \langle W_1, L_1 \rangle$ and $A_2 = \langle W_2, L_2 \rangle$, we have, by the definition of '$\circ$', the following: $A_1 \circ A_2 = \langle (W_1 - L_2) \cup (W_2 - L_1), L_1 \cup L_2 \rangle$. The first coordinate of the fusion arises because of the W$\circ$L=L rule: every W in A_1 that is matched to an L in A_2 is removed, and vice versa.

Now assume that A_1 and A_2 are nontrivial and that $\{A_1, A_2\}$ is consistent, so that it has a model. In any model of $\{A_1, A_2\}$, some constraint from $\Sigma = W_1 \cup W_2 \cup L_1 \cup L_2$ must be ranked above all the others. It cannot come from L_1, for this would contradict A_1, nor can it come from L_2, contradicting A_2. So it comes from $W_1 \cup W_2 - L_1 \cup L_2 = (W_1 - L_2) \cup (W_2 - L_1)$, and $A_1 \circ A_2$ is true in that model.

If $\{A_1, A_2\}$ is inconsistent, the result follows vacuously. If either of A_1 or A_2 is trivial in the sense of lacking L's, so that $\hat{A}_1$ or $\hat{A}_2 \in \mathcal{W}^*$, then $A_1 \circ A_2$ follows by W-extension (at most). If either of A_1 or A_2 is trivial in the sense of lacking W's, so that $\hat{A}_1$ or $\hat{A}_2 \in \mathcal{L}^+$, then it is false in all models and $\{A_1, A_2\}$ is inconsistent. □

To describe the general result, we need a notation for the fusion of an arbitrary set. We will write $f_i A_i$ for the fusion of arguments A_i, and more concisely fS for the fusion taken over all the elements of a set S. In the case of a unit sets S=$\{\psi\}$, we take $f\{\psi\} = \psi$. For the empty set, it is convenient to take $f\{\} = \delta$. Notice that the fusion of an arbitrary set is well-defined because the operation is associative and commutative. We adopt parallel conventions for writing conjunction over arbitrary sets.

(15) **Proposition 2.1**. Let $\mathcal{A} = \{A_i\}$ be a set of ERCs. $\mathcal{A} \vdash f\mathcal{A}$.

Pf. From repeated application of lemma (14) and familiar properties of entailment. Directly, (i) $\mathcal{A} \vdash A_1 \circ A_2$. But then (ii) $\mathcal{A} \cup \{A_1 \circ A_2\} \vdash A_1 \circ A_2 \circ A_3$ because $\{A_3\} \cup \{A_1 \circ A_2\} \vdash A_1 \circ A_2 \circ A_3$ by the lemma. Because $\Psi \vdash \wedge\Psi \wedge Q$ when $\Psi \vdash Q$, for any wff Q and any finite set of wffs Ψ, we have from (i) $\mathcal{A} \vdash [\wedge\mathcal{A} \wedge A_1 \circ A_2]$ and along with (ii) rephrased as $[\wedge\mathcal{A} \wedge A_1 \circ A_2] \vdash A_1 \circ A_2 \circ A_3$ we have by transitivity of entailment $\mathcal{A} \vdash A_1 \circ A_2 \circ A_3$. Repetition yields the result. □

Fusion thus provides us with a legitimate rule of inference for ranking arguments. If we know that φ and ψ hold of some ranking, then we are assured that $\varphi \circ \psi$ also holds of that ranking.

We can further pin down the logical status of A$\circ$B by observing that $A \circ B \vdash A \vee B$, and more generally that the fusion of an ERC set entails disjunction over that set.

(16) **Proposition 2.2**. Let $\mathcal{A}$ be a set of ERCs. $f\mathcal{A} \vdash \vee\mathcal{A}$.

Pf. Consider that $f\mathcal{A} = A_1 \circ A_2 \circ \ldots \circ A_n$ corresponds to $\langle \cup_i W_i - \cup_i L_i, \cup_i L_i \rangle$. If all the ERCs in $\mathcal{A}$ are degenerate, then both $f\mathcal{A}$ and $\vee\mathcal{A}$ are simply valid. Assume then that $\mathcal{A}$ contains at least one nondegenerate ERC. In any model R validating $f\mathcal{A}$, some $C_k \in \cup_i W_i$ dominates all of $\cup_i L_i$. But $C_k \in W_j$ for some j so A_j is true of R and therefore $\vee\mathcal{A}$ is also true of R. □

The relation between conjunction and fusion leads immediately to a consequence of no little interest. If the fusion operation eliminates all W's, then the fusands are inconsistent as a set and no ranking exists that will model them. Symbolically put, if $\hat{A}_1 \circ \hat{A}_2 \in \mathcal{L}^+$, so that $\hat{A}_1 \circ \hat{A}_2$ contains no W's but does contain at least one L, then $\{A_1, A_2\}$ is inconsistent. This follows because $A_1 \circ A_2$ cannot be satisfied, and so neither can $\{A_1, A_2\}$, which entails it. More generally, if $\mathcal{A} = \{A_i\}$ is a set of ERCs such that $f\mathcal{A} = f_i A_i \in \mathcal{L}^+$, then $\mathcal{A}$ is inconsistent, as is any larger ERC set that contains it.

(17) **Proposition 2.3**. Let $\mathcal{A}$ be a set of ERCs. If $f\mathcal{A} \in \mathcal{L}^+$ then $\mathcal{A}$ is inconsistent.

Pf. From Proposition 2.1, $\mathcal{A} \vdash f_i A_i$. But by assumption $f\mathcal{A}$ has no models, so $\wedge\mathcal{A}$ must be false in all models as well. □

(18) **Corollary to Proposition 2.3**. If $\Psi \subseteq \mathcal{A}$, an ERC set, and $f\Psi \in \mathcal{L}^+$, then $\mathcal{A}$ is inconsistent.

Pf. By Prop. 2.3, Ψ has no models. But any model of $\mathcal{A}$ would be a model of Ψ. □

Conversely, if $\mathcal{A}$ is inconsistent, then $\mathcal{A}$ contains a subset that fuses to $\mathcal{L}^+$. To prove this, we focus on a minimal inconsistent subset: its fusion cannot contain W without contradicting inconsistency.

(19) **Proposition 2.4**.**Fusion/Inconsistency**. $\mathcal{A}$ is inconsistent iff there is a $\Psi \subseteq \mathcal{A}$ such that $f\Psi \in \mathcal{L}^+$.

Pf. The RL direction is the corollary to Proposition 2.3. We show the LR implication.

If $\mathcal{A}$ is inconsistent, then $\mathcal{A}$ contains an inconsistent subset Ψ that is minimal in the sense that no proper subset of Ψ is itself inconsistent. Note that Ø is formally consistent, because there is no model in which any of its members is false. So $\Psi \neq$ Ø, but might be a singleton. Now consider $f\Psi$. Suppose that $f\Psi \notin \mathcal{L}^+$. It cannot be that $f\Psi = \delta$, because then for every $\psi \in \Psi$, we have $\psi = \delta$, and Ψ would hold in *every* model, not none. So $f\Psi$ contains W at some coordinate k. But $[f\Psi]_k$=W means that there is some $\psi \in \Psi$ with $[\psi]_k$=W and in addition that for all $\varphi \in \Psi$, $[\varphi]_k \neq$ L. Let us gather all such ψ in a set $\Theta = \{\psi \mid \psi \in \Psi$ and $[\psi]_k = \text{W}\}$. The set Θ has a model, namely one in which C_k stands at the top of the ranking, and indeed Θ holds in any model with C_k at the top.

Consider now $\Psi - \Theta$. Because Ψ is minimal with respect to inconsistency, the relation $(\Psi - \Theta) \subsetneq \Psi$ implies that $\Psi - \Theta$ has a model. (Subsetting is proper because $\Theta \neq$ Ø. If $\Psi - \Theta =$ Ø, then any model will do.) Let R be any such model. Construct a model R′ which is the same as R if C_k is at the top of R, otherwise let R′ be the same as R in every respect except that C_k stands at the top of the ranking in R′. Now, by construction, the constraint C_k assigns *e* to every $\varphi \in \Psi - \Theta$. Therefore C_k has no effect whatsoever on the evaluation of the ERCs in $\Psi - \Theta$, and R′ satisfies $\Psi - \Theta$ because R does. But R′ also satisfies Θ, because C_k is at the top of R′. So R′ satisfies $(\Psi - \Theta) \cup \Theta = \Psi$. But this is a contradiction, since Ψ was assumed inconsistent. Therefore $f\Psi \in \mathcal{L}^+$. □

As a practical application of these results, consider the problem of determining whether a given candidate is a *possible optimum* for some set of constraints — whether a ranking exists for which it is optimal, a problem studied in SLP. Candidates which are optimal for some ranking are predicted to be linguistically possible, and those which are never optimal are predicted to be universally impossible: this is a discrimination of some interest, which we will return to in §6. To test a candidate, construct a set of ERCs which treat it as the desired optimum. If this set is inconsisent, the candidate can never be optimal.

It is laborious to test every subset of constraints for consistency. It is wiser to pursue a hunt for a minimal inconsistent set using fusion as a guide. As in the proof just given, it is easy to identify a row-vector that *cannot* participate in a fusion that goes to $\mathcal{L}^+$: it will have a W at some coordinate (some tableau column) that is nowhere matched to an L in the same column. Eliminate this row from the tableau: it cannot belong to a minimal inconsistent set. The procedure can now be repeated on what remains of the original ERC set, seeking to determine if *it* contains a minimal inconsistent set. Iterate until no further diminution of the tableau is possible. If all ERC rows are discharged and the tableau is emptied, it is consistent, for it contains no minimal inconsistent set. If not, the residual arguments are inconsistent and will fuse to $\mathcal{L}^+$. The procedure can be carried out rapidly with pencil and paper, as recommended in Prince 2000.

This method of elimination is a stripped-down version of the Recursive Constraint Demotion (RCD) algorithm of Tesar 1995, Tesar & Smolensky 2001, and connects with its order-theoretic characterization in SLP. We return to RCD and its intimate relation to fusion in §4.

Our goal is to characterize entailment in much the same way as Proposition 2.4 characterizes inconsistency: in terms of the behavior of a fused subset. In standard logic, inconsistency is directly linked to entailment: $A \vdash B$ iff the expression $A \wedge \neg B$ is invalid, *i.e.* iff the set $\{A, \neg B\}$ is inconsistent. If we extend the logic of ERCs to include a negation like-operator, we will find a similar relation, one that leads directly from Proposition 2.4 to our desired conclusion.

Let us therefore introduce the *negative* of an ERC vector, arrived at by interchanging W and L values. (For purposes of formulating this definition, we observe the careful notational distinction between the ERC vector and the logical expression that interprets it.)

(20) **Def. Negative of ERC vector.**

[1] For a coordinate $[\hat{A}]_k$, the negative is defined by the following table:

$[\hat{A}]_k$	$-[\hat{A}]_k$
W	L
e	*e*
L	W

[2] For an ERC vector $\hat{A}$, $(-\hat{A})$ is defined coordinatewise as follows: $[-\hat{A}]_k = -[\hat{A}]_k$.

This notion of 'negative' is entirely natural in the context of the interpretation of an ERC vector as representing the comparison [a~b] of a candidate *a* with a rival *b*. The negative "$-[a\sim b]$" is exactly the comparison [$b\sim a$], which swaps W and L, while leaving e alone. Consequently, the negative of an ERC vector corresponds closely to the ordinary negation of its associated ERC, in the sense that $-\hat{A}$ is almost always associated with $\neg A$ by the usual rule of interpretation. Only the degenerate ERC δ is exceptional, since $\delta = -\delta$, there being no W's or L's to exchange. Since no wff of ordinary logic can be equivalent to its negation, the negative of δ cannot correspond to the *negation* of the wff expressing δ.

To see how the negative relates to negation, recall that an expression $\langle X,Y\rangle$ has the following interpretation: there is some constraint in X that dominates all constraints in Y.

(21) $$\langle X,Y\rangle \;=\; \exists D\, \forall C\, [C\in Y \rightarrow (D\in X \wedge D>>C)]$$

For the negative $-\langle W,L\rangle = \langle L,W\rangle$, the interpretation reads this way: there is some constraint in L that dominates all those in W. When W and L are both nonempty, each of $\langle W,L\rangle$ and $\langle L,W\rangle$ clearly implies that the other is false, by the asymmetry of domination order. If, for example, $D\in W$ dominates all $C\in L$, then no $C\in L$ can dominate all $D\in W$. This establishes that '$-$' on these ERC vectors is mirrored by '$\neg$' on the logical expressions that interpret them. Should one or the other or both of W,L be empty, some further delicacy is required

• When $L=\varnothing$ but $W\neq\varnothing$, the wff $\langle W,L\rangle$ is of the class $\mathcal{W}^*$, and it holds in every model. (The antecedent of the conditional in (21) is always false.) Swapping W and L to produce $-\langle W,L\rangle = \langle L,W\rangle$ creates an expression of the class $\mathcal{L}^+$, which holds in none. Here '$-$' converts a universally valid wff into a contradiction, exactly as '$\neg$' does.

• When $W=\varnothing$ but $L\neq\varnothing$, the wff $\langle W,L\rangle$ is of the class $\mathcal{L}^+$ and cannot be true in any model. The negative $-\langle W,L\rangle = \langle L,W\rangle$ is of the class $\mathcal{W}^*$, universally valid. Once again '$-$' corresponds exactly to '$\neg$' in its effects.

• When both L and W are empty, the expressions $\langle W,L\rangle$ and $-\langle W,L\rangle = \langle L,W\rangle$ each have empty first coordinates, leading to false antecedents in the logical expression they designate. Both are of the class $\mathcal{W}^*$, and universally valid. In this case '$-$' and '$\neg$' part company. This establishes that the only exception to the equation of negative and negation is, as claimed, the degenerate ERC δ.

(22) **Remark**. If $\hat{A}$ is a *nondegenerate* ERC vector, then $-\hat{A}$ is associated with a formula equivalent to $\neg A$ by the interpretive scheme (4). The converse also holds.

Pf. Along the lines of the discussion in the text.

With the notion of 'negative' in hand, we can reduce entailment to inconsistency. Since *inconsistency* is detected by fusion, it will follow that entailment reduces to fusion. First, let us establish the link between between the entailment relation and a suitable set involving the negative.

(23) **Lemma**. **Inconsistency.** For nondegenerate φ, and $\mathcal{A}$ a set of ERCs, $\mathcal{A} \vdash \varphi$ iff $\mathcal{A} \cup \{-\varphi\}$ is inconsistent.

Pf. By Remark (22), $-\varphi$ is the same as $\neg\varphi$ except when $\varphi=\delta$, so this is just ordinary logic. *I.e.*, writing 'f' for any contradiction, $\mathcal{A}\cup\{-\varphi\}\vdash f$ iff $\mathcal{A} \vdash \neg\varphi \supset f$ iff $\mathcal{A} \vdash \varphi \vee f$ iff $\mathcal{A} \vdash \varphi$. □

Next, the link between fusion and entailment.

(24) **Lemma**. For α, φ nontrivial, $\alpha \vdash \varphi$ iff $\alpha \circ (-\varphi) \in \mathcal{L}^+$.

Pf. LR: By lemma (23), $\{\alpha, -\varphi\}$ is inconsistent. By Proposition 2.4, some subset fuses to $\mathcal{L}^+$. Since φ is nontrivial, so is $-\varphi$. Since $\alpha, -\varphi$ are each nontrivial, neither belongs to $\mathcal{L}^+$ and the $\mathcal{L}^+$ subset cannot be either $\{\alpha\}$ or $\{-\varphi\}$. So it is $\{\alpha, -\varphi\}$and $\alpha \circ (-\varphi) \in \mathcal{L}^+$, as desired.

RL: Since $\alpha \circ (-\varphi) \in \mathcal{L}^+$, $\{\alpha, -\varphi\}$ is inconsistent. By lemma (23), $\alpha \vdash \varphi$. □

The main result relating entailment to fusion follows immediately.

(25) **Proposition 2.5. Entailment/Fusion.** $\mathcal{A} \vdash \varphi$ iff there is a $\Psi \subseteq \mathcal{A}$ such that $f\Psi \vdash \varphi$.

Pf. RL. Assume there is a $\Psi \subseteq \mathcal{A}$ such that $f\Psi \vdash \varphi$. By Proposition 2.1, we have $\Psi \vdash f\Psi$, and since $\Psi \subseteq \mathcal{A}$, $\mathcal{A} \vdash \Psi$. By transitivity of entailment, we have $\mathcal{A} \vdash \varphi$.

LR. First, let us handle various trivial cases.

(i) If $\mathcal{A}$ is inconsistent, then by Proposition 2.4 there is a $\Psi \subseteq \mathcal{A}$ such that $f\Psi \in \mathcal{L}^+$. Since $f\Psi$ is invalid, we have $f\Psi \vdash \varphi$ for any φ at all

(ii) If $\varphi \in \mathcal{W}^*$, then it is entailed by anything, so any subset of $\mathcal{A}$ will do.

(iii) If $\varphi \in \mathcal{L}^+$, then since $\mathcal{A} \vdash \varphi$, it must be that $\mathcal{A}$ is inconsistent. See (i).

Now assume that φ is nontrivial and $\mathcal{A}$ is consistent. From lemma (23), $\mathcal{A} \cup \{-\varphi\}$ is inconsistent. From Proposition 2.4, we have a $\Theta \subseteq \mathcal{A} \cup \{-\varphi\}$with $f\Theta \in \mathcal{L}^+$.

Let $\Psi = \Theta \cap \mathcal{A}$. Since $\mathcal{A}$ is itself consistent, no subset of $\mathcal{A}$ can be inconsistent, and we have $f\Psi \notin \mathcal{L}^+$. Therefore, since $f\Theta \in \mathcal{L}^+$, it must be that $-\varphi \in \Theta$, and $\Theta = \Psi \cup \{-\varphi\}$, whence

$$f\Theta = f(\Psi \cup \{-\varphi\}) = f\Psi \circ -\varphi$$

All we need now is for $f\Psi$ to be nontrivial. We have $f\Psi \notin \mathcal{L}^+$, from the consistency of $\mathcal{A}$. And since $-\varphi$ is nontrivial, we cannot have $f\Psi \in \mathcal{W}^*$, else $f\Psi \circ -\varphi$ would be nontrivial. By lemma (24), then, $f\Psi \circ -\varphi \in \mathcal{L}^+$ implies $f\Psi \vdash \varphi$. □

It is crucially *not* the case that $\mathcal{A} \vdash \varphi$ will always give us $f\mathcal{A} \vdash \varphi$. For example, {(e,L,W), (W,L,e)} entails (W,L,e), but the fusion (W,L,W) does not. Indeed, this kind of behavior demarcates the line between conjunction and fusion. However, we can easily identify one kind of case where conjunctive entailment guarantees fusional entailment: when the entailing set is minimal, in that none of its subsets is sufficient for entailment.

(26) **Def. Minimal entailing set**. If $\Psi \vdash \varphi$, and it's not the case that $\Theta \vdash \varphi$ for any $\Theta \subsetneq \Psi$, then Ψ is a *minimal entailing set* for φ.

Minimal entailing sets have a particular close relationship to fusion.

(27) **Corollary to Proposition 2.5**. Let Ψ be a minimal entailing set for φ. Then $f\Psi \vdash \varphi$.

Pf. Since $\Psi \vdash \varphi$, there is a $\Theta \subseteq \Psi$ with $f\Theta \vdash \varphi$, by Proposition 2.5. But $\Theta \vdash f\Theta$. So $\Theta \vdash \varphi$. By minimality, $\Theta = \Psi$. □

3. Fusion and Conjunction

Summary. Fusion and Conjunction are exactly equivalent when every W in the fusion vector of a set comes only from the fusion of W's. Similar conditions relating W and L coordinates determine when $A \circ B \vdash B$.

Fusion provides a way to ascertain not-necessarily-obvious entailments of ranking conditions, and also therefore a way of identifying redundancies: if $A' = A_1 \circ A_2$, then there is no need to include all of $\{A', A_1, A_2\}$ in the set of ranking arguments, because, by Proposition 2.1, all consequences of A' are entailed by $\{A_1, A_2\}$.

It is natural to ask, then, under what conditions the converse holds, so that $\{A_1, A_2\}$ can be *replaced* by $A_1 \circ A_2$, or more generally, under what conditions a set $\mathcal{A}$ can be replaced by the single ERC f. Since we always have $\mathcal{A} \vdash f\mathcal{A}$, we are interested in the conditions that ensure the converse relation, $f\mathcal{A} \vdash \bigwedge \mathcal{A}$.

Consider the set of constraints that award the 'polar' values W or L in a ranking argument A. These are the constraints that distinguish between the competing members of the candidate pair that generates the ranking argument. If two arguments involve exactly the *same* set of distinguishing constraints, then their fusion is equivalent to their conjunction. Here's an example:

(28) **Fusion as conjunction**

A	W	W		L
B	W	L		W
$A \circ B \equiv A \wedge B$	W	L		L

Equivalence generalizes to a broader class of ERCs — those which have a common core of constraints assigning just W, but outside that shared core each may also have other relation-assessing constraints, so long as they fuse to L or *e*. The following provides an example.

(29) **When fusion is conjunction**

	← all W's →		← no fusion to W →		
A	W	W		L	
B	W	W	L	W	
$A \circ B \equiv A \wedge B$	W	W	L	L	

The result is that the top two rows are jointly equivalent to the third, $A \circ B \equiv A \wedge B$. In this situation, we may replace rows A and B with their fusion $A \circ B$, without loss of information.

Let us call such ERCs 'W-compliant', since wherever they fuse to W, both coordinates must individually be W. The relation extends directly from pairs to general sets of ERCs. W-compliance over a set will mean that when a given coordinate fuses to W, all individual values at that coordinate must also be W.

(30) **Def. W-compliant** . A set of ERCs $\mathcal{A}$ is W-compliant iff $[f\mathcal{A}]_k$ =W implies that $\forall\hat{A}\in\mathcal{A}$, $[\hat{A}]_k$=W.

The useful result we seek (and will find) is that $A\circ B\equiv A\wedge B$ if A and B are W-compliant, and more generally, $f\mathcal{A}\equiv\wedge\mathcal{A}$ for W-compliant $\mathcal{A}$. In essence, W-compliance removes the taint of disjunction from '∘' and renders it logically equivalent to conjunction.

(31) **Proposition 3.1**. Let $\mathcal{A}$ be a set of ERCs. If $\mathcal{A}$ is W-compliant, then $f\mathcal{A}\equiv\wedge\mathcal{A}$.

Pf. Assume $\mathcal{A}$ to be W-compliant. We need only show $f\mathcal{A}\vdash\wedge\mathcal{A}$, since the converse is guaranteed by Proposition 2.1. There are two cases to consider.

[1] Assume first that $f\mathcal{A}$ has no models. Then $f\mathcal{A}\vdash\wedge\mathcal{A}$ holds vacuously..

[2] Now assume $f\mathcal{A}$ has a model R. Then there is a constraint $D\in(\cup_i W_i)-(\cup_j L_j)$ that dominates all of $\cup_j L_j$ in R. (Some such D must exist, else $\mathcal{A}\in\mathcal{L}^+$ and has no models.) By W-compliance, $D\in W_i$ for every $\hat{A}_i\in\mathcal{A}$. Therefore each $A_i\in\mathcal{A}$ is true in R, and thus $\wedge\mathcal{A}$ is also true in R. □

Does fusion-conjunction equivalence conversely imply W-compliance? The answer will be affirmative for sets of nontrivial ERCs and for many kinds of sets containing nontrivial ERCs. The one case requiring special handling involves members of $\mathcal{W}^*$, which are universally valid. Adding a member of $\mathcal{W}^*$ to a set of ERCs $\mathcal{A}$ does not enlarge the set of entailments at all, though it may weaken the fusion. But even when the fusion remains the same, W-compliance may be disrupted. Suppose for example $\mathcal{A}=\{(W,L),(e,W)\}$. Conjunction and fusion are equivalent, but W-compliance fails in the first coordinate. To achieve a fully general characterization, then, we pull out the $\mathcal{W}^*$ subset of $\mathcal{A}$ and focus on the non-$\mathcal{W}^*$ core $\mathcal{A}'$, assumed nonempty, observing that fusion-conjunction equivalence obtains for such $\mathcal{A}$ iff its core $\mathcal{A}'$ is W-compliant and the fusion of the core ($f\mathcal{A}'$) is the same as the fusion of the whole ($f\mathcal{A}$.), so that the $\mathcal{W}^*$ periphery is fusionally inert, as it were, in $f\mathcal{A}$.

(32) **Proposition 3.2.** Let $\mathcal{A}=\mathcal{A}'\cup\Psi$, where $\Psi\subseteq\mathcal{W}^*$, $\mathcal{A}'\cap\mathcal{W}^*=\emptyset$, and $\mathcal{A}'\neq\emptyset$.

$f\mathcal{A}\equiv\wedge\mathcal{A}$ iff $\mathcal{A}'$ is W-compliant and $f\mathcal{A}=f\mathcal{A}'$.

Pf. Observe that (*) $\wedge\mathcal{A}\equiv\wedge(\mathcal{A}'\cup\Psi)\equiv\wedge\mathcal{A}'$ by the logic of conjunction.

RL. By assumed W-compliance, $f\mathcal{A}'\equiv\wedge\mathcal{A}'$ (Prop. 3.1); from $f\mathcal{A}=f\mathcal{A}'$ (assumed) and $\wedge\mathcal{A}'\equiv\wedge\mathcal{A}$ (*), we have $f\mathcal{A}\equiv\wedge\mathcal{A}$.

LR. By assumption, (**) $f\mathcal{A}\vdash\wedge\mathcal{A}$.

[1] If $\mathcal{A}$ has no models, then $f\mathcal{A}\vdash\wedge\mathcal{A}$ yields $f\mathcal{A}\subseteq\mathcal{L}^+$. Since Ψ contributes no L's to $f\mathcal{A}$, it must be that $f\mathcal{A}=f\mathcal{A}'$, and $f\mathcal{A}'\subseteq\mathcal{L}^+$ as well, so that $\mathcal{A}'$ is vacuously W-compliant.

[2] Now assume $\mathcal{A}$ has a model; it follows that $\mathcal{A}'$ has a model, since $\mathcal{A}'\subseteq\mathcal{A}$. So $f\mathcal{A}'\notin\mathcal{L}^+$ and $f\mathcal{A}'$ contains W since $\mathcal{A}'\cap\mathcal{W}^*=\emptyset$. Let K be the set of coordinates on which $[f\mathcal{A}']_k$=W. We argue for the W-compliance of $\mathcal{A}'$ by reductio. If, contrary to the assertion, $\mathcal{A}'$ is *not* W-compliant, then there is a $\alpha\in\mathcal{A}'$ such that for some $j\in K$, we have $[\alpha]_j=e$. But

since $\alpha \notin \mathcal{W}^*$ (by assumption), there must be some other coordinate p on which $[\alpha]_p$=L. Now consider the model R in which C_j is top-ranked and C_p is ranked immediately beneath it. Since $[f\mathcal{A}']_j$=W, $f\mathcal{A}'$ is true in R. But since $[\alpha]_j=e$ and $[\alpha]_p$=L, it must be that α is false in R. So $\wedge\mathcal{A}$ is false in R. But this cannot be, since $f\mathcal{A} \vdash \wedge\mathcal{A}$ by assumption (**). Therefore, $\mathcal{A}'$ must be W-compliant.

We have the following equivalences:

$f\mathcal{A}' \equiv \wedge\mathcal{A}'$ (W-compliance, Prop. 3.1)

$\wedge\mathcal{A}' \equiv \wedge\mathcal{A}$. (*)

$\wedge\mathcal{A} \equiv f\mathcal{A}$ (**)

So $f\mathcal{A} \equiv f\mathcal{A}'$, i.e. $f\mathcal{A} \vdash f\mathcal{A}'$ and $f\mathcal{A}' \vdash f\mathcal{A}$. The desired stronger conclusion, $f\mathcal{A}=f\mathcal{A}'$, follows because $f\mathcal{A}'$ is nontrivial, which via Prop. 1.1 will force identity in all coordinates. □

Further insight into fusion-conjunction relation can be gained if we ask not just about equivalence but about the conditions under which certain simple entailments proceed. From A∧B, we may legitimately infer A and infer B, for any A and B whatever. Fusion works differently. We cannot safely conclude, given a model of A∘B, that A holds or that B holds: only that one of them does – we don't know which.

For example, from φ=(W,L,L) – "C_1 dominates both C_2 and C_3" – we certainly cannot conclude (e,W,L) – that C_2>>C_3. But φ is the fusion of C_2>>C_3 and C_1>>C_2:

(W,L,L)=(W,L,*e*)∘(*e*,W,L)

Furthermore, which fusand holds may vary from model to model of A∘B. Consider

(W,W,L)=(W,*e*,L)∘(*e*,W,L).

Here the fusion entails *neither* fusand, though whenever A∘B holds, at least one will be true.

Conjunction, then, is conservative in a way that fusion is not. This is fusion's great strength: we may lose some information, but we can gain a direct representation of important facts that are only implicit in a conjunctive expression: for example, that C_1>>C_3, from (C_1>>C_2)∧(C_2>>C_3).

It will be useful, then, to understand not only how fusion and conjunction can come to be equivalent, as in Propositions 3.1 and 3.2, but also how certain basic patterns of conjunctive inference can be replicated in fusion. Let us begin with a very general question. Conjunction always allows 'weakening': if A⊢C then A∧B⊢C, no matter what B is. Suppose α⊢φ. Under what conditions on ψ are we guaranteed that α∘ψ⊢φ?

The case of $\varphi \in \mathcal{W}^*$ can be discarded as trivial: valid φ places no requirements on its antecedents. Elsewhere, fusing ψ with α will spoil the entailment relation just in case ψ introduces W into a coordinate in α∘ψ where there is no W in φ. This amounts to requiring that whenever $[\alpha]_k=e$ and $[\varphi]_k=e$, we do not have $[\psi]_k$=W. It should be intuitively clear why this is so: an extra W is an extra disjunction, and from p∨q⊢r we cannot conclude p⊢r.

We can phrase the condition more positively, indeed contrapositively, in terms of subset relations between coordinates bearing polar values W,L in the various participant ERCs. To this end, let us adopt the following notation. To increase legibility, for W_ψ let us write simply W(ψ), and similarly L(ψ) for L_ψ. For W(ψ)∪L(ψ), the collection of constraints assessing polar values in ψ, let us write P(ψ). We may then state the requirement this way: the constraints assessing W in the interloper ψ must assess polar values in α *or* φ:

$$W(\psi) \subseteq P(\alpha) \cup P(\varphi) = L(\alpha) \cup W(\varphi)$$

The latter equality comes about because $L(\varphi)\subseteq L(\alpha)$ and $W(\alpha)\subseteq W(\varphi)$ by virtue of $\alpha\vdash\varphi$.
Before proving the claim, it is useful to sharpen Proposition 1.1b, which broadly declares that for *nontrivial* α, φ, if $\alpha\vdash\varphi$, then $W(\alpha)\subseteq W(\varphi)$ and $L(\varphi)\subseteq L(\alpha)$. But further distinctions can be drawn among the various trivial cases.

(33) **Remark.** Suppose $\alpha\vdash\varphi$. Then the 'W-condition' $W(\alpha)\subseteq W(\varphi)$ and the 'L-condition' $L(\varphi)\subseteq L(\alpha)$ hold except in the following circumstances.

a. If $\varphi\in\mathcal{W}^{*}$, the W-condition is only guaranteed to succeed when $W(\alpha)=\varnothing$, i.e. $\alpha\in\mathcal{L}^{+}$ or $\alpha=\delta$.

b. If $\alpha\in\mathcal{L}^{+}$, the L-condition is only guaranteed to succeed when $L(\varphi)=\varnothing$, i.e. $\varphi\in\mathcal{W}^{*}$.

Pf. (a). Suppose $\varphi\in\mathcal{W}^{*}$. The L-condition holds vacuously, because $L(\varphi)=\varnothing$. Let us then examine the W-condition, by cases. Suppose first $\alpha\in\mathcal{W}^{*}$ and $\alpha\neq\delta$; consider (W)$\vdash$(e). The W-condition fails, but entailment succeeds. Suppose α is nontrivial. Examine (W,L)$\vdash$(e,W). Again, we have entailment without satisfaction of the W-condition. But if α lacks W, so that either $\alpha=\delta$ or $\alpha\in\mathcal{L}^{+}$, the W-condition succeeds vacuously, since $W(\alpha)=\varnothing$.

(b) Suppose $\alpha\in\mathcal{L}^{+}$. We've already covered the case $\varphi\in\mathcal{W}^{*}$, where both conditions succeed vacuously. Suppose now $\varphi\in\mathcal{L}^{+}$. We have (L,*e*)$\vdash$(*e*,L), but failure of the L-condition. Now suppose φ is nontrivial. We have (L,*e*)$\vdash$(W,L), with failure of the L-condition. □

It is now straightforward to deal with $W(\psi)$ in $\alpha\circ\psi\vdash\varphi$.

(34) **Proposition 3.3. FusionalWeakening**.

Suppose $\alpha\vdash\varphi$, with $\varphi\notin\mathcal{W}^{*}$. Then $\alpha\circ\psi\vdash\varphi$ iff $W(\psi)\subseteq P(\alpha)\cup P(\varphi) = L(\alpha)\cup W(\varphi)$.

Pf. From Remark (33), $W(\alpha)\subseteq W(\varphi)$. Going LR, we also have $W(\alpha\circ\psi)\subseteq W(\varphi)$, i.e.

$$W(\alpha)\cup W(\psi) - L(\alpha)\cup L(\psi) \subseteq W(\varphi).$$

From this expression it is clear that for any $x\in W(\psi)$, either $x\notin L(\alpha)$ and $x\in W(\varphi)$, or $x\in L(\alpha)$.

RL. Suppose $W(\psi)\subseteq P(\alpha)\cup P(\varphi) = L(\alpha)\cup W(\varphi)$.

$$W(\alpha)\cup W(\psi) \subseteq W(\alpha) \cup L(\alpha) \cup W(\varphi)$$

So, with $W(\alpha)\subseteq W(\varphi)$, we have

$$W(\alpha)\cup L(\alpha)\cup W(\varphi) \subseteq L(\alpha) \cup W(\varphi)$$

Whence by transitivity of '$\subseteq$'

$$W(\alpha)\cup W(\psi)\subseteq L(\alpha)\cup W(\varphi)$$

Whence

$$W(\alpha\circ\psi)= [W(\alpha)\cup W(\psi) - L(\alpha)\cup L(\psi)]\subseteq [L(\alpha) \cup W(\varphi)- L(\alpha)\cup L(\psi)]$$
$$= W(\varphi)- L(\alpha)\cup L(\psi) \subseteq W(\varphi)$$

Since $L(\psi) \subseteq L(\alpha\circ\psi)$ by the definition of fusion, Proposition 1.1 yields $\alpha\circ\psi\vdash\varphi$. □

Proposition 3.3 allows an advance in the analysis of the minimal entailing set, and indeed of entailing sets in general. Recall that if Ψ is a minimal entailing set for φ, then $f\Psi\vdash\varphi$ (Corollary to Proposition 2.5). The converse does not hold, because a set $\mathcal{A}$ with $f\mathcal{A}\vdash\varphi$ might contain various extraneous ERCs outside a minimal entailing set $\Psi\subsetneq\mathcal{A}$ that nevertheless do not destroy the fusional entailment. For example, (W,L,*e*) and (*e*,W,L) jointly and minimally entail (W,*e*,L), but we can include *e.g.* (W,W,L) with them and still get the entailment from the fusion of the resulting (nonminimal) set.

We now know exactly what such extranea must look like: they cannot introduce W into the fusion where no W was found. When $[f\Psi]_k=[\varphi]_k=e$, the k^{th} coordinate of would-be co-entailer cannot be W.

This means that in any case where $f\mathcal{A}\vdash\varphi$, the entailing set $\mathcal{A}$ can be analyzed into the disjoint union of an entailing (sub)set Ψ and a periphery Θ, where $f\Theta$ stands in the non-W-inserting relationship to $f\Psi$. We do not demand of *each* $\theta\in\Theta$ that it stand in the appropriate relationship to Ψ, for the fusion can makes amends for individual divergences. In particular, if θ_i contains a misplaced W, θ_j can make up for it with an L in the same position. For example, let $\Psi=\{(W,L,e)\}$, and $\Theta=\{(e,L,W),(e,W,L)\}$, with $\varphi=(W,L,e)$. Then

$$f\Psi\vdash\varphi \qquad \textit{i.e. } (W,L,e)\vdash(W,L,e)$$
$$f(\Psi\cup\Theta)= f\Psi\circ f\Theta\vdash\varphi \quad \textit{i.e. } (W,Le)\circ(e,L,W)\circ(e,W,L)\vdash(W,L,e)$$

But not every member of Θ can be harmlessly added to Ψ:

$$(W,L,e)\circ(e,L,W) \nvdash (W,L,e).$$

Due to the generality of Proposition 3.3, this remark holds of any entailing set, and not just the minimal one.

In the example we made use of the following equivalence:

(35) **Remark**. $f(X\cup Y)=fX\circ fY$

Pf. From the fact that "$\circ$" is associative, commutative, and idempotent. □

We can now characterize the structure of sets whose fusion entails a certain ERC.

(36) **Corollary 1 to Proposition 3.3. Structure of Fusional Entailments**.
Let $\mathcal{A}$ be such that $f\mathcal{A}\vdash\varphi$ for some $\varphi\notin\mathcal{W}^{*}$. Then $\mathcal{A}=\Psi\cup\Theta$, where $\Psi\cap\Theta=\varnothing$, with $f\Psi\vdash\varphi$ and $W(f\Theta)\subseteq P(f\Psi)\cup P(\varphi)= L(f\Psi)\cup W(\alpha) = \bigcup_i L(\psi_i)\cup W(\alpha)= \bigcup_i P(\psi_i)\cup P(\alpha)$.

Pf. $f\mathcal{A}= f(\Psi\cup\Theta)=f\Psi\circ f\Theta$, by the remark, to which Proposition 3.3 applies directly. The last two equalities come about because the polar values of fusands are preserved as polar values in the fusion, and in particular the L-coordinates are preserved as L's . □

The general weakening property of conjunction leads to conjunction eliminability: in the theorem '$A\vdash C \Rightarrow A\wedge B\vdash C$', simply take C=A to obtain $A\wedge B\vdash A$ by modus ponens (using $A\vdash A$). The same pattern of reasoning, applied to Proposition 3.3, also leads immediately to a characterization of the conditions under which fusion parallels conjunction in eliminability by the scheme A *op* B → B.

(37) **Corollary 2 to Proposition 3.3**. **Fusion Elimination.** Suppose $\varphi\notin\mathcal{W}^{*}$. $\psi\circ\varphi\vdash\varphi$ iff $W(\psi)\subseteq P(\varphi)$.

Pf. Identify φ and α in Proposition 3.3 and observe that $\varphi\vdash\varphi$. □

Corollary 2 also gives us a quick alternative proof of the W-compliance conditions in Propositions 3.1 and 3.2. Observe that $\alpha\circ\varphi\vdash\alpha\wedge\varphi$ is equivalent to the assertion that $\alpha\circ\varphi\vdash\varphi$ *and* $\alpha\circ\varphi\vdash\alpha$. From these we have $W(\alpha)\subseteq P(\varphi)$ and $W(\varphi)\subseteq P(\alpha)$, which is exactly W-compliance.

We conclude by noting that the converse relation $\alpha\vdash\alpha\circ\varphi$ is also definable in terms of the L-content of the participating ERCs.

(38) **Proposition 3.4**. (a) If $L(\varphi)\subseteq L(\alpha)$, then $\alpha\vdash\alpha\circ\varphi$.
(b) Assume that $\alpha\notin\mathcal{L}^{+}$ or $\varphi\in\mathcal{W}^{*}$. Then $\alpha\vdash\alpha\circ\varphi$ implies $L(\varphi)\subseteq L(\alpha)$.

Pf. (a). We want $W(\alpha)\subseteq W(\alpha\circ\varphi)$ and $L(\alpha\circ\varphi)\subseteq L(\varphi)$. For the W-condition, observe

$$W(\alpha\circ\varphi)= W(\alpha)\cup W(\varphi)-L(\alpha)\cup L(\varphi)\text{, from the definition of fusion.}$$

But since by assumption $L(\varphi)\subseteq L(\alpha)$, we have $L(\alpha)\cup L(\varphi)=L(\alpha)$, so

$$W(\alpha)\cup W(\varphi)-L(\alpha)\cup L(\varphi) = W(\alpha)\cup W(\varphi) - L(\alpha) = W(\alpha)\cup[W(\varphi) - L(\alpha)]$$

The last step is legitimate because $W(\alpha)\cap L(\alpha)=\emptyset$. Consequently

$$W(\alpha) \subseteq W(\alpha)\cup[W(\varphi) - L(\alpha)] = W(\alpha\circ\varphi).$$

As for the L-condition, we want $L(\alpha\circ\varphi)\subseteq L(\alpha)$, *i.e.* $L(\alpha)\cup L(\varphi) \subseteq L(\alpha)$, *i.e.* $L(\alpha) \subseteq L(\varphi)$, which we are given. Therefore, by Proposition 1.1, $\alpha\vdash\alpha\circ\varphi$.

(b). Here we invoke Remark (33). The L-condition is guaranteed to hold except in the one case $\alpha\in\mathcal{L}^{+}$ and $\varphi\notin\mathcal{W}^{*}$, which is precisely what's excluded. By the L-condition, $\alpha\vdash\alpha\circ\varphi$ yields $L(\alpha\circ\varphi)\subseteq L(\alpha)$, i.e. $L(\alpha)\cup L(\varphi) \subseteq L(\alpha)$, *i.e.* $L(\alpha) \subseteq L(\varphi)$. □

4. Entailment, Ranking, and the Minimal Stratified Hierarchy

Summary. Ranking induces a relation between a fused ERC and its fusands, with the fusion as least upper bound. A similar relation holds between an entailed ERC and its (minimal) entailers in hierarchies that satisfy a set of ERCs, and quite generally in the Minimal Stratified Hierarchy produced by Recursive Constraint Demotion, which itself calls crucially on fusion in its formulation.

Fusion leads to a highly varied set of entailment relations between A,B, and $A\circ B$. We know that $(A\wedge B)\vdash(A\circ B)$ and that $(A\circ B)\vdash(A\vee B)$, but beyond these, we may have it that $A\circ B$ entails one of its fusands but not the other; entails both; or entails neither. (These cases classify exhaustively in terms of W- and L- subsetting relations, by means of Propositions 3.3 and 3.4; Appendix 3 provides a chart of examples.) This means that there will be no simple general relation between A, B, and $A\circ B$ in the set of ERCs as (partially) ordered by entailment ($A\vdash B$ as "$A\leq B$").

Conjunction behaves well with respect to the entailment ordering, because both conjuncts are entailed. Writing [X] for the class of wffs logically equivalent to X, we have $[A\wedge B]\leq[A]$ and $[A\wedge B]\leq[B]$. Indeed, $[A\wedge B]$ is the meet (inf, g.l.b.) of [A] and [B] in the Lindenbaum Algebra, which is based on that ordering. From this we have, for example, $[A]\leq[C] \Rightarrow [A\wedge B] \leq[C]$. But nothing similar obtains for fusion.

By contrast, a far more tractable relation between fusions and fusands, as well as between all entailed and entailing ERCs, obtains in the context of constraint rankings that satisfy an ERC set, and even more generally in the stratified hierarchy which is determined by Recursive Constraint Demotion (RCD: Tesar 1995, Tesar & Smolensky 1994, 1998, 2000) as modified below. This hierarchy has been shown to play an important role in the theory quite aside from its role in learning (see *e.g.* SLP), and we will see more evidence of its centrality below.

By 'stratified hierarchy' is meant a certain kind of partial order, one in which incomparability is an equivalence relation. As in all partial orders we have three choices for the relation between any two entitities: a>b, b>a, $a\|b$, where the last means that the order relation does not determine a>b or b>a, *i.e.* a and b cannot be 'compared'. A stratum consists of a set of incomparable elements, which share all order relations. Thus, if $a\|b$ and z>a, we have z>b, and so on.

It is instructive to decouple the characterization of the formal object from the procedure(s) giving rise to it. (In so doing, we reconstrue some of Tesar & Smolensky's results to emphasize certain aspects of the construction.) From Tesar and Smolensky (1994, 1998, 2000), we know that each stratum in this hierarchy consists of constraints that do not conflict over the set of ranking arguments, and that each constraint is located in as high a stratum as it can be. The number of strata is also minimal, in the sense that no hierarchy with fewer strata can resolve all the ranking arguments. Thus, to define it, we seek an extremal principle which will pick out just this hierarchy from all stratified hierarchies consistent with the set of ranking arguments.

As a preliminary, let us define what it means for an ordinary, totally-ordered ranking of constraints to satisfy a set of ERC vectors. Let's construe a constraint C_i to be the function that projects the i^{th} coordinate of an argument vector, so that $C_i(\alpha)=[\alpha]_i$ This yields the following notion of 'comparative satisfaction' involving a constraint C_i and a vector α:

(39) **Comparative Satisfaction of a constraint**. $C_i\vDash\alpha$ iff $C_i(\alpha)$=W.

To extend this to a hierarchy $H=[C_1>>C_2>>\ldots>>C_n]$, it is useful to define the *fusion* of constraints in essentially the same way we defined the fusion of ranking vectors. A constraint, after all, is a column vector in the comparative tableau, just as an ERC is a row vector, and any vectorial operation can be applied to either. Fusion of constraints is defined coordinatewise, where a coordinate is an ERC. Purely for notational convenience, we will indicate the fusion of constraints by '$\otimes$', just to redundantly emphasize what is being fused.

(40) **Constraint Fusion** wrt to a vector. $(C_i \otimes C_j)(\alpha) = C_i(\alpha) \circ C_j(\alpha)$

Since it uses the same mechanism as the argument vector fusion, the constraint fusion has the same properties.[9] In particular, it is associative, so that $\otimes S$, for a set of constraints, is well-defined. With this in hand, we can deal with satisfaction of a hierarchy of constraints. The usual informal observation is that an argument is satisfied if the first nonblank cell encountered in the left-to-right sweep contains a W, or if all cells are blank.

(41) **Hierarchical satisfaction of argument vector.**
$H \vDash \alpha$ iff , either $\alpha=\delta$ or, for some $m \leq n$, for i ranging from 1 to m, $\otimes_i C_i \vDash \alpha$.

The transition to stratified hierarchies is unproblematic. A stratum is a set of constraints; let us rank the strata instead of the constraints, using the same subscripting device to mark domination order. Then we need only form the union of sequential strata, starting with the first (numbered '0') and proceeding inclusively, before fusing them.

(42) **Stratified hierarchical satisfaction of argument vector**
$H \vDash \alpha$ iff either $\alpha=\delta$ or, for some $m \leq n$, $\otimes_{i=0}^{i=m}(\cup_i S_i) \vDash \alpha$.

From the definitions, it follow that any stratum responsible for the satisfaction of some vector must consist only of constraints assessing {e,W} for that vector, i.e. the constraints in the stratum cannot conflict. Satisfaction of a *set* of arguments is simply satisfaction of each:

(43) **Stratified hierarchical satisfaction of a set of argument vectors**
$H \vDash \mathcal{A}$ iff $\forall \alpha \in \mathcal{A}$, $H \vDash \alpha$

It is typically the case that numerous stratified hierarchies will satisfy a given set of arguments. We are interested in the *minimal* hierarchy, the one which departs least from the Boolean ideal of a single stratum. There are two dimensions of minimality: the number of strata, and the number of constraints in each strata. As we will see below, these cannot trade-off against each other in the present context, and the natural combined measure – summing over stratum-depth × number of constraints in the stratum – gives a unique result.

---

[9] To generalize fully to fusions of constraints over sets of argument vectors, we need only imagine the arguments enumerated and then define $C(\{\alpha_i\})=\langle C(\alpha_i)\rangle$, i.e. the (column) vector whose i^{th} coordinate is $C(\alpha_i)$. Then the fusion goes coordinate-wise, in the usual fashion.

Let the strata be numbered from the top. Let $\text{card}(S_k)$ be the cardinality of stratum S_k . Let $\|Ci\|$ be the stratum depth of the stratum to which constraint C_i belongs, as given by the stratum subscript. We then have a natural *subordination index* σ for any hierarchy $H= \langle S_0, S_1,S_2,...,S_n\rangle$.

(44) **Subordination Index**. $\sigma(H)= \Sigma(k \times \text{card}(S_k)) =\Sigma\|C_i\|$.

We may now define the *minimal stratified hierarchy* (MSH) for a given set of ranking arguments and its corresponding constraints, over which the arguments are stated.

(45) **Minimal Stratified Hierarchy** (MSH) for a consistent set of ERCs.

The minimal hierarchy for a consistent set $\mathcal{A}$ of argument vectors, $\mathcal{H}(\mathcal{A})$, is a stratified hierarchy with (a) $\mathcal{H}\vDash\mathcal{A}$, and (b) $\sigma(\mathcal{H})$ minimal, i.e. $\sigma(\mathcal{H})\leq\sigma(H)$ for all $H\vDash\mathcal{A}$.

Concerns about the uniqueness of the MSH will be resolved after the introduction of RCD.

For present purposes it is useful to modify the definition of the minimal hierarchy to include the case of *inconsistent* argument sets. We want to segregate off the sources of inconsistency and place them in the bottom stratum. The subsets of concern are those that fuse to $\mathcal{L}^+$. Given a set of ERCs $\mathcal{A}$ let us enumerate all the sets $\Phi_i\subseteq\mathcal{A}$ such that $f\Phi_i\in\mathcal{L}^+$. Let $\mathcal{F}= \bigcup_i\Phi_i$ collect them all; $\mathcal{F}$ is of course itself inconsistent and $f\mathcal{F}\in\mathcal{L}^+$. But $\mathcal{A}-\mathcal{F}$ is consistent, and indeed is the maximal consistent subset of $\mathcal{A}$. In case of inconsistency, we want the MSH to validate as many ERCs in $\mathcal{A}$ as possible. We therefore characterize the MSH for any set of ERCs, consistent or inconsistent, as follows:

(46) **Minimal Stratified Hierarchy** for a general set of ERCs.

The minimal hierarchy for $\mathcal{A}$ is the minimal hierarchy as defined in (45) for $\mathcal{A}-\mathcal{F}$, with $\mathcal{F}$ appended as the lowest stratum.

Let us now turn to RCD, an effective procedure for constructing a stratified hierarchy that will be shown to be identical to the Minimal Stratified Hierarchy just defined. The product of RCD is known as the "h-dominant Target Stratified Hierarchy" in Tesar & Smolensky 2000:92, or more concisely as the *target hierarchy* or just *target*. It has also been more descriptively dubbed the *Favoring Hierarchy* (Samek-Lodovici & Prince 1999), because it is arrived at by recursively selecting out the subset of constraints that *favor* (do not assess L for) the desired optima – *favor*, in the sense that the desired optima stand at the top of the candidate set, according to these constraints' ordering of the candidates (whence they assess only W and e).

The construction of a stratified hierarchy by RCD is readily explicated in terms of the notion of fusion explored here.[10] Given a constraint set S and set of ERCs $\mathcal{A}$, proceed as follows:

[10] Thanks to Bruce Tesar for suggesting that fusion could play a role in the formulation of RCD. For the various formulations of RCD which we are re-working, see Tesar 1995, Tesar & Smolensky 1994,1998, 2000, as well as the order-theoretic formulation of Samek-Lodovici & Prince 1999, and the formulation in terms of comparative tableaux of Prince 2000.

• Partition S into two subsets determined by the structure of $f\mathcal{A}$: FAVC($\mathcal{A}$) and DISFAVC($\mathcal{A}$) =S−FAVC($\mathcal{A}$). The subset FAVC consists of the those constraints assessing only W and *e* across the set of ranking arguments $\mathcal{A}$: that is, those $C_k \in S$ for which $[f\mathcal{A}]_k$=W or $[f\mathcal{A}]_k$=*e*. The set DISFAVC($\mathcal{A}$) consists of those $C_k \in S$ for which $[f\mathcal{A}]_k$=L.

• Put the constraints in FAVC($\mathcal{A}$) aside as a stratum in the developing hierarchy.

• The set of ERCs $\mathcal{A}$ is likewise partitioned into two subsets: SOLVEDARGS and UNSOLVEDARGS, where UNSOLVEDARGS=$\mathcal{A}$−SOLVEDARGS. The ERCs in SOLVEDARGS are those for which some constraint in FAVC($\mathcal{A}$) assesses W, i.e. those $\alpha \in \mathcal{A}$ for which FAVC($\mathcal{A}$)$\vDash\alpha$, i.e. those for which [$\otimes$FAVC($\mathcal{A}$)](α)=W. The demands on ranking imposed by the arguments in SOLVED are satisified by the stratum just formed. The remaining constraints go into UNSOLVED and are those $\alpha \in \mathcal{A}$ for which [$\otimes$FAVC($\mathcal{A}$)](α)=*e*. The developing hierachy has, as yet, nothing to say about them.

An example of this first step should make its logic clear.

(47) **Specimen Argument Pattern** – Step 1 of RCD for $\mathcal{A} = \{\varphi,\psi,\chi\}$

	C_1	C_2	C_3	C_4	
φ	W	L			SOLVEDARGS
ψ		W	L		UNSOLVEDARGS
χ			W	L	
$f\mathcal{A} = \varphi\circ\psi\circ\chi$	W	L	L	L	
	FAVC($\mathcal{A}$)	← DISFAVC($\mathcal{A}$) →			

At the first step of RCD, we have

Stratum 0 = FAVC($\mathcal{A}$) = $\{C_1\}$

since $[f\mathcal{A}]_1=[\varphi\circ\psi\circ\chi]_1$=W and this is true for no other constraints.

SOLVEDARGS=$\{\varphi\}$, since [$\otimes$FAVC($\mathcal{A}$)](φ)=[$\otimes\{C_1\}$](φ) = $C_1(\varphi)$ = W.

The residual constraints and arguments are these:

UNSOLVEDARGS=$\mathcal{A}$−SOLVEDARGS = $\{\psi,\chi\}$

DISFAVC($\mathcal{A}$) = S−FAVC($\mathcal{A}$) = $\{C_2,C_3,C_4\}$

At this point, some of the ranking arguments have been handled successfully (those in SOLVEDARGS, here merely φ), and the constraints that handle them have been incorporated into a nascent hierarchy (here just C_1). We are left with a diminished set of constraints and a diminished set of arguments: the constraints in DISFAVC($\mathcal{A}$) and the arguments in UNSOLVEDARGS. But these present a problem whose structure is exactly like that of the larger problem we began with: we must satisfy a set of ranking arguments over a set of constraints. We may therefore repeat the procedure just outlined, in the hopes of further diminishing the problem, and so on until it disappears or is shown to be unsolvable. At each stage, we place the new stratum just below the ones previously created. If the original set of ERCs was consistent, we will be able to continue cycling through the procedure until all arguments are placed in SOLVEDARGS and all constraints are stratified. If the original argument set was inconsistent, we will stratify as much as possible, and (in the version given here) place the irresolvable residue of constraints in the bottommost level of the hierarchy.

The key to success lies in the effects of removing from further consideration the ERCs that end up in SOLVED. Returning to our example, we notice that C_2 is *not stratifiable* at the first step because $[\varphi]_2$=L, causing $[f\mathcal{A}]_2$=L. But as soon as φ is removed from consideration on the grounds that it is solved by the first stratum, the situation is quite favorably transformed.

(48) **Specimen Argument Pattern** - Step 2 of RCD $\mathcal{A}' = \mathcal{A} - \{\varphi\} = \{\psi,\chi\}$

	C_1	C_2	C_3	C_4	
φ	W	L			
ψ		W	L		SOLVEDARGS
χ			W	L	UNSOLVEDARGS
$f\mathcal{A}' = \psi\circ\chi$	W	W	L	L	
		FAVC($\mathcal{A}'$)	← DISFAVC($\mathcal{A}'$) →		

The removal of φ has led to the disappearance of the offending L = $[\varphi]_2$ in the C_2 column. We now have $[f\mathcal{A}']_2$=W, and we may proceed by stratifying C_2 and inserting ψ into SOLVED. The procedure may be carried forward yet again on the remaining constraints and ERCs – here χ and $\{C_3, C_4\}$ only– yielding the hierarchy $\{C_1\}>>\{C_2\}>>\{C_3\}$. Only C_4 remains, because it has never played the favoring role in an ERC, but the set of arguments has been tranferred completely to SOLVEDARGS and no argument is left to be judged by C_4.

To complete the characterization of RCD, then, we need to specify its termination conditions. If the original set $\mathcal{A}$ of ERCs has a model, then RCD will successfully eliminate all of $\mathcal{A}$, emptying UNSOLVED, sufficient reason to quit. When the last ERC is thereby eliminated, let any remaining constraints (like C_4 above) be put in a new, lowest stratum. (Technically, with the convention $f\text{Ø} = \delta$, such constraints will fit right into FAVCONS – they do not assess L in any ERC; and nothing special need be said about them.) In the example, C_4 falls into this class: it only becomes rankable once all ERCs have been eliminated.

When there is a nonempty residue of unsolvable ranking arguments RESARG, with the unfortunate property that $f\text{RESARG}\in\mathcal{L}^+$, a correlated residue of constraints RESCON will be left over. No member of RESCON can be put in FAV, and no member of RESARG can be eliminated by being shifted to SOLVED; also reason to quit, after placing RESCON in the bottom stratum.

A final wrinkle devolves from the presence of degenerate ERCs. Since these have no W, they will never gain entry into SOLVED. Yet nothing is required to solve them! We assume a pre-sort whereby degenerate ERCs are removed.

This discussion is implemented in the following recursive definition of RCD.

(49) **RCD**. Definition of H(S,A). S a set of constraints, A a set of ERCs.

FAVC:={C∈S| C(*f*A)≠ L}

SOLVEDARGS := {φ∈A| ⊗FAVC ⊨φ}

If SOLVEDARGS=Ø then H(S,A):=S

Else H(S,A):= FAVC>>H(S−FAVC, A−SOLVEDARGS).

As promised, bottom strata may consist of conflicting but unrankable constraints. For example, if S={(W,L),(L,W)}, then on the very first round, SOLVEDARGS=Ø and we're done, with one stratum. Such an inconsistent stratum may also emerge in the course of recursion, when no more arguments can be solved. Such inconsistency, which infects the entire hierarchy, arises when not every ERC can be transferred to SOLVEDARGS. An inconsistent hierarchy can be detected by placing a suitable clause in the RCD procedure, one sensitive to the presence of unsolved ERCs, or merely by examining the resulting hierarchy: a dead giveway is the presence of an ERC whose highest non-*e* stratum contains an L.

For concreteness, we illustrate how RCD may be reconfigured to announce inconsistency. We introduce the variable "flag", which is initialized to the value "CONSISTENT".

(50) **RCD** with built-in inconsistency detection.

Definition of H(S,A). S a set of constraints, A a set of ERCs.

If A= Ø then H(S,A):=S

Else

FAVC:={C∈S| C(*f*A)≠ L}

SOLVEDARGS := {φ∈A| ⊗FAVC(φ)=W}

If SOLVEDARGS=Ø then H(S,A):=S and flag:="INCONSISTENT"

Else H(S,A):= FAVC>>H(S−FAVC, A−SOLVEDARGS).

The hierarchy that RCD computes – Tesar & Smolensky's *target* (with any nonnull RESCON tucked in beneath) – is precisely the Minimal Stratified Hierarchy for $\mathcal{A}$. To see this, note that RCD *preserves minimality*. Consider the top stratum S_0 and the arguments A_0 that it satisfies. S_0 is clearly minimal over this argument set – it has a subordination index of 0. Now consider $S_0 \cup S_1$ and $A_0 \cup A_1$. The two level hierarchy $H=\langle S_0,S_1\rangle$ is minimal over $A_0 \cup A_1$ – no constraint can be raised up from S_1, no one-level hierarchy will do, and any three-(or more)-level hierarchy can only fare worse on the subordination index. (To construct a three level hierarchy, one must move a constraint weighted 0 or 1 into a position where it is weighted 2, for a net increase of 2 or 1 in the subordination index.) And so on, down through the strata.

Let us now turn to the relations between entailment, fusion, and ranking. As a first observation, we note that entailed ERCs are completely inert with respect to the organization of the MSH, in the sense that the stratal structure of $\mathcal{H}(\mathcal{A})$ is unchanged when entailed ERCs are added to or removed from $\mathcal{A}$. When $\mathcal{A}\vdash\varphi$, $\mathcal{A}$ is logically equivalent to $\mathcal{A}\cup\{\varphi\}$, and both $\mathcal{A}$ and $\mathcal{A}\cup\{\varphi\}$ are consistent with the same rankings, among them the unique minimal stratified hierarchy.

(51) **Remark.** Let $\mathcal{H}(\mathcal{A})$ be the MSH for $\mathcal{A}$. Suppose $\mathcal{A}\vdash\varphi$. Then $\mathcal{H}(\mathcal{A}\cup\varphi)= \mathcal{H}(\mathcal{A})$.

Any stratified hierarchy can also be used to impose a ranking on ERCs associated with it. Imagine a hierarchy, such as the following, displayed in conventional tableau form, with the strata arranged across the top: then the supporting ERCs are forced into a vertical arrangement that echoes the horizontal stratification of the constraints.

(52) **Specimen tableau**

		I	II		III	IV
		C_1	C_2	C_3	C_4	C_5
I	φ	W		L		
II	ψ		W		L	
	χ		W	W	L	
III	ξ				W	L

Let us define the rank $|C|$ of a constraint C belonging to any hierarchy H as the *stratum* to which it belongs in H. (If the hierarchy is a total order, then each stratum contains just one constraint.) It is natural, then, to define the rank of an ERC φ, written $|\varphi|$, as the *highest* stratum which contains one of its polar-valued constraints.[11] This is the stratum that determines the success – or failure – of φ on H. (For convenience, we stipulate $|\delta|$ to be S_0.) In the above, the rank of ψ is stratum II=$\{C_2,C_3\}$. Note that even under total ordering of constraints, several ERCs may have the same rank. If we impose C_2>>C_3 in the above example, we still have $\psi,\chi \in |C_2|$.

(53) **Def**. **Rank**.

The rank $|C|$ of a constraint C in a hierarchy H is the stratum which contains it.
The rank $|\varphi|$ of an ERC is the highest stratum containing a C such that $C(\varphi)\in\{W, L\}$.

Given a set of ERCs $\mathcal{A}$, we can notate the inter-ERC stratal relations like this:

(54) **Relative rank of ERCs**

$\lvert\varphi\rvert > \lvert\psi\rvert$	the rank of φ dominates the rank of ψ
$\lvert\varphi\rvert = \lvert\psi\rvert$	the rank of φ is the same as that of ψ
$\lvert\varphi\rvert \geq \lvert\psi\rvert$	the rank of φ is the same as or dominates that of ψ

[11] The virtues of this method of defining 'rank' are two in number. (1) It allows us to cross-compare the ranks of ERCs and constraints; (2) it uses '$\geq$' in the natural way that corresponds to domination. Each ERC could also be assigned a numerical 'grade' based on its rank; this leads to the multivalued semantics of the logic RM: see §7.3 below.

A constraint hierarchy assigns a rank for any ERC framed in terms of its underlying constraint set. Because the MSH uniquely determines a hierarchy from an ERC set, it is sensible, given remark (51), to speak of the rank in the MSH of an ERC φ entailed by $\mathcal{A}$ even if it is not a member of $\mathcal{A}$ – it is its rank in $\mathcal{H}(\mathcal{A}\cup\varphi)$, which has exactly the strata of $\mathcal{H}(\mathcal{A})$.

The notion of rank interacts in useful ways with both entailment and fusion.

In any hierarchy at all, a fused ERC shows straightforward good behavior with respect to its fusands, simply because all polar values in the fusion come from polar values in the fusands. An immediate consequence is that the fusion is a least upper bound rankwise for all of its constituents.

(55) **Proposition 4.1**. **Fusion Dominance**. Let Ψ be a set of ERCs over the constraint set $\{C_i\}$ used in some hierarchy H. Then, with respect to H, $|f\Psi|\geq|\psi|$, for all $\psi\in\Psi$. Furthermore, $f\Psi$ has the same rank in H as some $\psi\in\Psi$.

Pf. If $f\Psi\neq\delta$, the rank of $f\Psi$ is determined by some C_k with $[f\Psi]_k\in\{W,L\}$. Now consider all those constraints C_h which are of higher rank than C_k in H. These are the only ones that could determine a higher rank for any $\psi\in\Psi$. For each such C_h, it must be that $[f\Psi]_h=e$, else it would determine the rank of $f\Psi$. But since $[f\Psi]_h=e$, it must be that $[\psi]_h=e$ for all h, so none of the C_h can determine the rank of any $\psi\in\Psi$ either. This establishes that $|f\Psi|\geq|\psi|$, for all $\psi\in\Psi$. Furthermore, since $[f\Psi]_k\in\{W,L\}$, it must be that $[\psi_i]_k\in\{W,L\}$ for some $\psi_i\in\Psi$ and therefore that Ck determines the rank of ψ_i.

If $f\Psi=\delta$, it lies at the top by convention, with all its degenerate fusands. □

This result is entirely independent of whether H satisfies any of the ERCs in Ψ. To draw further conclusions, it is useful to note the following:

(56) **Lemma**. If $[\psi]_k$=W or $[\psi]_k$=L, then $|\psi|\geq|C_k|$.

Pf. By definition, $|\psi|$ is the highest stratum occupied by any constraint polar on ψ. □

W-compliance induces sameness of rank for constraints that satisfy a set of ERCs.

(57) **Proposition 4.2**. Let $\Psi\subseteq\mathcal{A}$ be W-compliant. If H is any hierarchy that satisfies Ψ, then $|\varphi|\approx|\psi|$ for all $\varphi,\psi\in\Psi$. In $\mathcal{H}(\mathcal{A})$, whether $\mathcal{A}$ is consistent or not, if $\delta\notin\Psi$, then $|\varphi|\approx|\psi|$ for all $\varphi,\psi\in\Psi$.

Pf. Suppose $H\vDash\Psi$. Then either $f\Psi=\delta$, or $f\Psi$ contains W. If $f\Psi=\delta$, then $\psi=\delta$ for all $\psi\in\Psi$, so that $f\Psi$ and all of Ψ are given top rank by convention. If $f\Psi$ contains W, then its rank is determined by some C_k with $[f\Psi]_k$=W. But by W-compliance, $[\psi]_k$=W for all $\psi\in\Psi$, and so by the lemma, for all ψ, $|\psi|\geq|C_k|=|f\Psi|$. We already have $|f\Psi|\geq|\psi|$, for all ψ, by Proposition 4.1. Therefore $|f\Psi|=|\psi|$, and all $\psi\in\Psi$ have the same rank.

This result applies directly to $\mathcal{H}(\mathcal{A})$, for consistent $\mathcal{A}$. Even if $\mathcal{A}$ is inconsistent, the argument given for the cases $f\Psi\notin\mathcal{L}^+$ goes through. But with if $\mathcal{A}$ inconsistent, it may be that $f\Psi\in\mathcal{L}^+$. Then, assuming $\delta\notin\Psi$, $\Psi\subseteq$ RESARG, and all of Ψ sits in the bottom stratum. □

The requirement $\delta \notin \Psi$ comes into play only when $f\Psi \in \mathcal{L}^+$. In this case, Ψ is W-compliant, but if $\delta \in \Psi$, then Ψ will be split between the top rank (of δ) and the bottom rank (of $\Psi - \{\delta\}$).

We now establish that argument-entailment and ranking are closely related.

(58) **Lemma.** If nondegenerate φ is derived from α by L-retraction alone, then $|\varphi|=|\alpha|$ in any hierarchy that satisfies α, as well as in $\mathcal{H}(\mathcal{A})$, for any $\mathcal{A}$ containing α.

Pf. Given Proposition 4.2, we need only note that α and φ are W-compliant. □

As before, the nondegeneracy caveat is required for the case where $\varphi \in \mathcal{L}^+$: then repeated L-retraction can produce $\varphi = \delta$, eliminating all barriers in φ to ranking by RCD, so that it pops to the top.

(59) **Lemma**. If φ is derived from α by W-extension alone, then $|\varphi| \geq |\alpha|$ in any hierarchy.

Pf. Suppose the rank of α is determined by C_k, so that $|\alpha|=|C_k|$. Now since φ is derived by W-extension, there is another coordinate p such that $[\varphi]_p$=W while $[\alpha]_p=e$. If $|C_p|>|C_k|$, then $|\varphi| > |\alpha|$. Else $|\varphi|=|\alpha|$. □

(60) **Proposition 4.3. Entailment bounding**. If $\alpha \vdash \varphi$, with $\alpha,\varphi \notin \mathcal{W}^*$, then $|\varphi| \geq |\alpha|$ in any hierarchy H with $H \vDash \alpha$, as well as in $\mathcal{H}(\mathcal{A})$, for any $\mathcal{A}$ containing α.

Pf. Consider any hierarchy H with $H \vDash \alpha$. It cannot be that $\alpha \in \mathcal{L}^+$, so α is nontrivial. From Proposition 1.1, we know that all consequences of a single nontrivial ERC follow by some sequence of L-retractions followed by some sequence of W-extensions. By Lemma (58), L-retraction leaves the rank of an argument the same; by Lemma (59), W-extension cannot lower its rank and may raise it. Thus the thesis holds in any $H \vDash \alpha$. Now consider $\mathcal{H}(\mathcal{A})$, for any $\mathcal{A}$ with $\alpha \in \mathcal{A}$. If $\mathcal{H} \vDash \alpha$, the result just shown will apply. If not, then α lies at the bottom, and $|\varphi| \geq |\alpha|$ for every $\varphi \in \mathcal{A}$ including those where $\alpha \vdash \varphi$. □

Entailment bounding can fail when the antecedent or the consequent is in $\mathcal{W}^*$. Consider the following example:

	C_1	C_2
α	W	L
δ		
β	W	
γ		W

Here C_1>>C_2, yielding two strata.. Among the arguments, α,δ,β belong to the first rank, and γ to the second. But $\alpha \vdash \gamma$, even though $|\alpha|>|\gamma|$, because here the consequent is in $\mathcal{W}^*$. Furthermore, $\delta \vdash \gamma$ and $\beta \vdash \gamma$, although $|\delta|>|\gamma|$ and $|\beta|>|\gamma|$, because here the antecedents (and therefore the consequents) are in $\mathcal{W}^*$. And since valid ERCs entail each other, they would all have to be at the same rank, which they manifestly are not. This example also illustrates how the $\mathcal{W}^*$ condition on Proposition 4.3 could be tweaked: since δ is always at the top, it needn't be excluded from consequent position.

Proposition 4.3 has a somewhat paninian flavor, since rank follows from entailment, though it takes place among ranking-arguments rather than constraint rankings. Similar results hold among the constraints, when the operation of fusion is generalized to them: see §8 below.

We now generalize to the situation where entailment involves a set of ERCs. As usual, fusion provides the key, and we first obtain a useful result from the propositions just established:

(61) **Corollary to Proposition 4.3**. Assume $\mathcal{A} \not\subseteq \mathcal{W}^*$ and $\varphi \notin \mathcal{W}^*$. If $f\mathcal{A} \vdash \varphi$, then $|\varphi| \geq |\alpha_i|$ for all $\alpha_i \in \mathcal{A}$ in any H such that $H \vDash f\mathcal{A}$, and in $\mathcal{H}(\mathcal{A})$.

> *Pf.* Proposition 4.1 gives us $|f\mathcal{A}| \geq |\alpha_i|$ in any H, without restrictions. Proposition 4.3 assures us that $|\varphi| \geq |f\mathcal{A}|$ in any $H \vDash f\mathcal{A}$. Therefore, $|\varphi| \geq |\alpha_i|$ in this case. Prop. 4.3 also gives us $|\varphi| \geq |f\mathcal{A}|$ in $\mathcal{H}(\mathcal{A} \cup \{f\mathcal{A}\})$. But note that $\mathcal{H}(\mathcal{A}) = \mathcal{H}(\mathcal{A} \cup \{f\mathcal{A}\})$ by remark (51), since $\mathcal{A} \vdash f\mathcal{A}$. Therefore, $|\varphi| \geq |\alpha_i|$ in $\mathcal{H}(\mathcal{A})$. □

Entailment bounding can be now extended to the general case, in which a conjunction of ERCs entails another ERC. To get a sharp result, we restrict ourselves to the case where every ERC in the conjunction is required for the entailment, *i.e.* to minimal entailing sets, as defined in (26), p. 14. If Ψ is minimal, then φ proves to be an upper bound, in the MSH, for all of Ψ.

(62) **Proposition 4.4**. If $\Psi \vdash \varphi$, with Ψ a minimal entailing set for φ, then $|\varphi| \geq |\psi|$ for all $\psi \in \Psi$, in any H such that $H \vDash \Psi$, and in $\mathcal{H}(\mathcal{A})$ for any $\mathcal{A} \supseteq \Psi$.

> *Pf.* If $\varphi \in \mathcal{W}^*$, then Ψ is empty and the assertion is vacuously true. Assume that $\varphi \notin \mathcal{W}^*$. Note in addition that a nonempty subset of $\mathcal{W}^*$ cannot be a minimal entailing set, because only valid ERCs would be entailed by it, and valid ERCs are entailed by Ø. So $\Psi \not\subseteq \mathcal{W}^*$ and $f\Psi \notin \mathcal{W}^*$. Because Ψ is a minimal entailing set for φ, we have $f\Psi \vdash \varphi$, by the corollary to Proposition 2.5, (27), p. 14.
>
> (a) MSH case. With $\varphi \notin \mathcal{W}^*$ and $f\Psi \notin \mathcal{W}^*$, the relation $f\Psi \vdash \varphi$ gives us $|\varphi| \geq |f\Psi|$, by Proposition 4.3, in any MSH containing $f\Psi$. Let $\mathcal{A}$.be any ERC set with $\Psi \subseteq \mathcal{A}$. We then have $|\varphi| \geq |f\Psi|$ in $\mathcal{H}(\mathcal{A} \cup \{f\mathcal{A}\})$. Now consider $\mathcal{H}(\mathcal{A})$. Since $\Psi \vdash f\Psi$, we have $\mathcal{H}(\mathcal{A}) = \mathcal{H}(\mathcal{A} \cup \{f\mathcal{A}\})$, by remark (51). Therefore $|\varphi| \geq |f\Psi|$ in $\mathcal{H}(\mathcal{A})$. But $|f\Psi| > |\psi|$ for all $\psi \in \Psi$ in any hierarchy.
>
> (b) Say $H \vDash \Psi$. Then $H \vDash f\Psi$ because $\Psi \vdash f\Psi$. Then $|\varphi| \geq |\psi|$ for all $\psi \in \Psi$, directly from the corollary to Prop. 4.3. □

Unlike the situation with the fusands and the fusion, an ERC entailed by a set Ψ, even a minimal entailing set, can *strictly* outrank every member of Ψ. Consider for example the system

A_1: (W,L,*e*) A_2: (W,W,L) A_3: (*e*,W,L)

In any hierarchy validating $\{A_1,A_2,A_3\}$, there are three constraint ranks $|C_1|>|C_2|>|C_3|$, where $|C_1|=|A_1,A_2|$ and $|C_2|=|A_3|$. We have $A_3 \vdash A_2$ (minimally) but also $|A_2|>|A_3|$, strictly.

5. Finding Entailments

Summary. The question of whether $\mathcal{A} \vdash \varphi$ can be efficiently decided for nontrivial φ by performing RCD on $\mathcal{A} \cup \{-\varphi\}$, checking for inconsistency. Various economies may be implemented, using results obtained in §§3-4.

Given a single nontrivial ERC φ, how might one best determine whether or not φ is entailed by an arbitrary set $\mathcal{A}$? We know from Proposition 2.5 (25) , p.14, that $\mathcal{A} \vdash \varphi$ iff there is a $\Psi \subseteq \mathcal{A}$ such that $f\Psi \vdash \varphi$. Fusion is easy to compute, but if we are condemned to root around among the (possibly very) numerous subsets of $\mathcal{A}$, checking the fusional behavior of each, then the reduction of conjunctive entailment to fusional entailment would be of little practical value. But since an intimate rapport obtains between inconsistency and entailment, the inconsistency-detection powers of RCD provide a rapid and efficient method of entailment-detection as well.

From lemma (23) §2 we know that an entailment relation $\mathcal{A} \vdash \varphi$ is accompanied by the inconsistency of $\mathcal{A} \cup \{-\varphi\}$ when $\varphi \neq \delta$. If we want to discover whether in fact $\mathcal{A} \vdash \varphi$, given arbitrary nontrivial φ, we need merely conduct RCD on $\mathcal{A} \cup \{-\varphi\}$. If a consistent hierarchy results, there is no entailment. Inconsistency signals that the entailment is valid.

The RCD-based procedure for entailment-checking can be sharpened in the light of the results of §4.

[1] **Omit irrelevant ERCs and constraints.** In determining the validity of $\mathcal{A} \vdash \varphi$, any ERC ranked above φ in $\mathcal{H}(\mathcal{A} \cup \{\varphi\})$ is irrelevant. We know that if Ψ is a minimal entailing set for φ, then $|\varphi| \geq |\psi|$ for all $\psi \in \Psi$, by Proposition 4.4 (62). This means that to determine the entailment status of φ, we need only examine those ERCs of rank equal to or lower than $|\varphi|$ in $\mathcal{H}(\mathcal{A} \cup \{\varphi\})$.

Furthermore, we may truncate the set of constraints under consideration: none with rank higher than $|\varphi|$ need be included in the calculation. (If C is any such, and ψ is any ERC with $|\psi| \leq |\varphi|$, then C(ψ)=*e* and C disables none of them from belonging to an entailing set for φ.

It might be thought that a similar economy could be imposed at the lower end of the hierarchy, if there are strata below the *lowest*-ranked polar value carried by φ. How could constraints ranked so low interfere with an entailment relation among higher-ranked constraints? Easily, it turns out. Consider, for example, the following array of constraints: is φ entailed by $\{\alpha,\beta,\gamma\}$? The MSH structure on $\{\alpha,\beta,\gamma\}$ is shown.

(63)

	C_1	C_2	C_3	C_4	C_5
α	W			L	
β		W	L		W
γ				W	L
φ		W	L		

Constraint C_5 lies in a stratum below that of φ's lowest polar evaluation (C_3). If we omit C_5, then β looks exactly like φ, and entailment is assured. But φ is not entailed when C_5 is included, as the reader may easily ascertain. Of course, any constraint in a stratum lower than that of the lowest polar evaluation of φ may be omitted if it contains no W's.

[2] **Halt when the answer is found.** There is no need to run the process to its conclusion in every case. We can assume that $\mathcal{A}$ is itself already known to be consistent (else why are we bothering with it?), so we needn't test to be sure that *no* subset of ERCs fuses to $\mathcal{L}^+$. All that matters is whether subsets containing $-\varphi$ are consistent; *i.e.* whether $-\varphi$ itself can be successfully integrated into the minimal hierarchy. Therefore the RCD process can quit as soon as $-\varphi$ is successfully stratified, and need not run on to check whether the rest of the arguments can be handled.

[3] **Amalgamate W-compliant ERCs**. Since entailed ERCs are inert with respect to the structure of the minimal hierarchy, as observed in remark (51), it follows that if an ERC is alone in a stratum, it *cannot be* entailed. To maximize the number of stratum-isolates, as well as to minimize the number of ERCs in the hierarchy, it is advisable to collapse W-compliant sets to their fusions, as licensed by Prop. 3.1 (31). From Prop. 4.2 (57), we know that all members of a W-compliant set have the same rank. It is therefore reasonable to determine $\mathcal{H}(\mathcal{A})$ prior to fusion-collapse, since this will sharply limit the set of ERCs that must be checked for W-compliance.

[4] **Forget stratal details**. When we speak of using RCD to check for consistency, we refer only to its essential machinery; for example, it is *stratifiability* and not the actual stratified hierarchy itself that is required for the consistency proof. There is no need to keep record of the stratal structure of $\mathcal{H}(\mathcal{A}\cup\{-\varphi\})$, and something like the *tableau reduction* of Prince 2000 provides an appropriately stripped-down version of the process.

The following procedure may therefore be outlined:

(64) **General Entailment Check**

[1] Determine $\mathcal{H}(\mathcal{A})$ by RCD.

[2] Collapse all W-compliant sets by fusion; call the result $\mathcal{H}(\mathcal{A}')$.

[4] To check any $\varphi\in\mathcal{A}'$ to see whether $\mathcal{A}'\vdash\varphi$,

3a] Replace φ by $-\varphi$,

3b] Perform RCD on the subhiererarchy of strata $S_i\in\mathcal{H}(\mathcal{A}')$ with $|\varphi|\geq|S_i|$ and over the set of arguments $\Psi\subseteq\mathcal{A}'$ with $|\varphi|\geq|\psi|$ for $\psi\in\Psi$.

- On any step of RCD where $-\varphi$ is stratified, quit: NOT($\mathcal{A}\vdash\varphi$).
- If RCD yields flag="INCONSISTENT", then: $\mathcal{A}\vdash\varphi$.

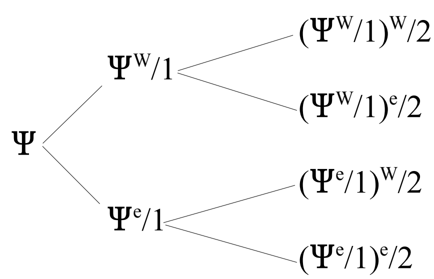
A final procedural question is how the collapse of W-compliant sets should be undertaken. Since members of a W-compliant subset of $\mathcal{A}$ will all have the same rank in $\mathcal{H}(\mathcal{A})$, the matter boils down to examing each rank of arguments to find W-compliant subsets.

A rank consists of a stratum S of constraints and a set Ψ of ERCs. Each $C_k\in S$ bifurcates Ψ into two classes, which we may call Ψ^W/k and Ψ^e/k. Members of Ψ^W/k are those $\psi\in\Psi$ with $[\psi]_k=W$;

the members of Ψ^e/k are those $\psi\in\Psi$ with $[\psi]_k=e$. Any W-compliant subset of Ψ must lie entirely within one or the other of these. The following example illustrates this observation:

		…	C_k	…
	…	…	…	…
Ψ^W/k	ψ_1		W	…
	ψ_2		W	…
Ψ^e/k	ψ_3		*e*	…
	ψ_4		*e*	…
	...	…	…	…

A strategy of bifurcative sorting therefore suggests itself. Placing the constraint $C_k\in S$ in some arbitrary order, as indicated by subscripting, we can first bifurcate C_1 into $\Psi^W/1$ and $\Psi^e/1$. We then bifurcate each of these according to the assignments of C_2, and so on, achieving this kind of result:



When bifurcation yields an ERC set that is null or has a single member, that branch of the search is over. Because the subsets obtained at each level of bifurcation are disjoint, the search will rapidly exhaust Ψ.

The terminal leaves of the tree will contain the candidates for W-compliance. But the search for W-compliant sets is not over when the constraints in the stratum S, correlating with the rank of Ψ, have been gone through. Because we are proceeding by elimination, it may happen – just as in RCD – that the segregation of various ERCs produces new opportunities for coordinates to fuse to W, and therefore for W-compliance to fail. This effect is illustrated in the following 2-stratum example:

	Stratum I		Stratum II		
	C_1	C_2	C_3	C_4	C_5
ψ_1	W	*e*	L	W	L
ψ_2	W	*e*	L	*e*	*e*
ψ_3	*e*	W	L	L	*e*

Here $\Psi^W/1 = \{\psi_1, \psi_2\}$ looks good in the top stratum. But the removal of ψ_3 discloses a failure of W-compliance at C_4. Therefore, the process must continue into the lower strata after the rank of Ψ has been analyzed. What's required is essentially another step of RCD, bring to the fore those constraints that fuse to W over the potentially W-compliant set produced by the first round of bifurcation. Another round of bifurcation may then proceed, and so on.

6. Entailment and Harmonic Bounding

Summary. Harmonic bounding is signaled by a trivial ERC involving the bounded candidate. Harmonic bounding leads to the existence of entailment relations: any set of ERCs that is free of entailments is also free of harmonic bounding within the set of candidates that is involved in the ERCs.

The linguistic forms and relations available to natural languages are a tiny fragment of those that are logically possible. Hard-and-fast principles identifying linguistic primitives and their modes of combination will rule out many conceivable structures. Some conceptions of generative grammar see in explicit 'principles' (deemed 'parameters' when omissible) the primary account of ungrammaticality. Current theory introduces a new source of explanation: many forms that are structurally legitimate in the most basic sense are not admitted by the grammar of any language because they are never *optimal* under any ranking of the constraint set. This shifts the burden from stipulated principles (the constraints themselves) to emergent properties of their interaction.

The most straightforward situation in which a form can be structurally licit but never achieve optimality occurs when it is *harmonically bounded* by some other form (Samek-Lodovici 1992, Prince & Smolensky 1993).

Suppose that candidate *a* incurs no more violations than candidate *b* on any constraint and that on some constraint it incurs fewer: in short, *a* always does at least as well as *b* on any constraint and sometimes does better. Then no ranking can render *b* optimal: in the competition *a* vs. *b*, which corresponds to the ERC we write as [*a*~*b*], any constraint that distinguishes the two (and *a fortiori* the highest-ranked of them) will prefer *a*.

This means that the ERC [*a*~*b*] belongs to the class $\mathcal{W}^{*}$ and its negative [*b*~*a*] to the class $\mathcal{L}^{+}$. Universal validity means success on all possible rankings; universal invalidity means failure on all rankings. We can now see, looking inside the ERC, which has hitherto been treated as atomic, that nondegenerate triviality is co-extensive with the harmonic bounding of one form by another. Notice that we need not even know whether *a* is itself optimal, or even possibly optimal, in order to safely conclude that *b* is hopeless (Prince & Smolensky 1993:95).

Harmonic bounding of one form by another ensures failure of the bounded form not just under the strict domination among constraints assumed by OT, but in virtually any scheme that calculates optimality from violation status on individual constraints, so long as incurring more violations is worse. Simply summing globally across the constraint set, the bounded candidate has more total violations; and, because it is locally never better than its bounder, no constaint-weighting scheme with positive weights (preserving the sense of 'more=worse') can change the relative status of the two competitors. The bounded candidate has literally nothing going for it, no weapon to call on in the struggle with its bounder.

But simple bounding by a single form is not the end of the story. Within Optimality Theory, a *set* of candidates can work collectively in a ranking-sensitive fashion to force perpetual suboptimality on another candidate (SLP 1999). Imagine confronting a ranking with a candidate set {a,b,z}and suppose that under some rankings *a* is optimal with respect to this set, besting *b* and *z*; that under others *b* beats *a* and *z*; and that these two states of affairs exhaust the possibilities of ranking. Then *z* cannot be optimal. (Adding further candidates to the set cannot improve *z*'s performance against *a* and *b*; therefore it is legitimate to draw general conclusions about *z* from this behavior.) The phenomenon of collective bounding is intrinsically tied to the notion of strict domination and is not replicable in full generality within weighted-sum systems. To see how it works in a particular case, consider the following example from alignment theory (McCarthy & Prince 1993).

(65)

	PARSE-σ	ALL-FEET-LEFT
a. (σσ)(σσ)(σσ)	0	6
b. (σσ)(σσ)σσ	2	2
c. σσσσσσ	6	0

The cells under each constraint heading display the number of violations incurred by each form. The constraint PARSE-σ demands that every syllable σ belong to a foot, a constituent indicated here by parentheses, and it assesses one violation for each foot-free σ. ALL-FEET-LEFT demands that each foot appear at the left edge of the form and it assesses one violation for each σ separating a foot from that edge; the violations for each foot in a form are added together to give the violation total for the whole form.

Observe that the sum of violations across the constraint set for the intermediate candidate (b) is *less* than that for the candidates (a) and (c), which are more extreme along each dimension of evaluation. Yet (b) is collectively bounded by (a) and (c) over this constraint set. The ranking PARSE-σ >>ALL-FEET-LEFT selects (a) as the best of the three candidates; the ranking All-Feet-Left>>Parse-σ selects (c).

Portraying the order relations induced by the violation structure diagrammatically, we have:

(66)

PARSE-σ	ALL-FEET-LEFT
a	*c*
|	|
b	*b*
|	|
c	*a*

Here we can see that *b* is bracketed by {*a*,*c*}. If, as in SLP, we think of the constraint hierarchy as a composition of constraint-functions, where each constraint returns the top stratum of the order it imposes on the candidates presented to it, then it is clear that *b* can never be in that top stratum, no matter how the constraints are ranked. With the ranking PARSE-σ >>ALL-FEET-LEFT, for example, PARSE-σ applied to {*a,b,c*} yields {*a*}. ALL-FEET-LEFT is then stationary on {*a*}.

The effect is also clear when presented in ERC form, with *b* placed in the position of the desired optimum:

(67)

$b \sim \{a,c\}$	PARSE-σ	ALL-FEET-LEFT
α: (σσ)(σσ)σσ ~ (σσ)(σσ)(σσ)	L	W
β: (σσ)(σσ)σσ ~ σσσσσσ	W	L
α∘β	L	L

The set of arguments {α,β} is inconsistent precisely because there is no ranking that can make a bounded candidate optimal. For a set of ERCs $\{[z{\sim}a_i], i{\in}I\}$, all based on the same desired optimum, fusion to $\mathcal{L}^+$ is thus the hallmark of bounding: whether it has one member or many, the set $\{a_i, i{\in}I\}$ bounds the candidate *z* and prevents it from being optimal.

It is instructive to examine the kind of sets involved in the bounding relation. As background, we need to specify the order relation induced on candidates by a ranking. SLP:38 defines the relation 'better than' like this, translating into present terms:

(68) **Def**. **'Better than' on a ranking**. For candidates *a*,*b* and the ranking R, *a* is 'better than' or 'beats' *b*, written ($a{\succ}b$;R) iff R⊨[$a{\sim}b$] and [$a{\sim}b$]≠δ.

The codicil [$a{\sim}b$]≠δ prevents *a* and *b* from exhibiting exactly the same profile of violations and therefore doing equally well on the constraint hierarchy. Observe that the definition of 'better than' for rankings extends as well to 'better than' on a single constraint: we need only think of a constraint as a one-member hierarchy. The refined version of the SLP conception proposed above (p. iv) unifies the two notions of 'better than' from the start.

The relation 'better than' is an order: it is obviously asymmetric (and therefore irreflexive) because of the nondegeneracy codicil. To show transitivity, let us work from a slightly more general result.

Let us write #*a* for the number of violations incurred by *a* on a constraint (whose identity will be clear from context). As usual, #*a*<#*b* on C iff C([a~b])=W.

(69) **Lemma**. **Transitivity of** '~'. [$a{\sim}b$], [$b{\sim}c$] ⊢ [$a{\sim}c$].

Pf. From Proposition 2.1, we have {[$a{\sim}b$],[$b{\sim}c$]}⊢[$a{\sim}b$]∘[$b{\sim}c$]. We show that [$a{\sim}b$]∘[$b{\sim}c$]⊢[$a{\sim}c$], because both the W-condition and the L-condition of Proposition 1.1a are satisfied. Suppose that C([$a{\sim}b$]∘[$b{\sim}c$])=W for a given constraint C. Then C([$a{\sim}b$])≠L and C([$b{\sim}c$])≠L. That is, #a≤#b and #b≤#c, and one of the inequalities must be strict for the corresponding ERCs to fuse to W on C. Therefore C([$a{\sim}c$])=W, and the W-condition is satisfied. Suppose now that C([$a{\sim}c$])=L for some C. This excludes the possibility that #a≤#b and #b≤#c simultaneously on C. Therefore either #b>#a or #c>#b, i.e. at least one of C([$a{\sim}b$]) and C([$b{\sim}c$]) is L, so C([$a{\sim}b$]∘[$b{\sim}c$])=L. □

(70)**Corollary**. **Transitivity of 'Better than'**. If $(a\succ b;R)$ and $(b\succ c;R)$, then $(a\succ c;R)$.

Pf. Follows directly from the lemma and the fact that $\{x\succ y;R) \Rightarrow R\vDash[x\sim y]$. □

Now to the characterization of bounding by sets. In the broadest sense, *z* will be bounded so long as we can identify a set of candidates A which is guaranteed to contain, for each ranking, something that beats z. Call such a set a 'covering set' for *z* with respect to Σ.

(71) **Def. Covering Set.** A is a covering set for *z* with respect to a constraint set Σ iff for each ranking R_i of Σ there is some element $a_i \in A$, $a_i \neq z$, with $R_i \vDash [a_i \sim z]$.

The covering set has some useful properties. If A covers *z*, then any superset of A is also covers *z*; similarly, if A doesn't cover *z*, then no subset of A covers *z* either. However, the general 'covering set' notion is too weak to allow a sharper specification of the properties that lead to bounding, because a covering set can contain all kinds of extraneous material which makes no contribution to the bounding effect. Given a covering set A for *z*, we know that the set of ERCs $\{[z\sim a_i]\}$ is inconsistent, but Proposition 2.4 informs us that inconsistency is due to the presence of a *sub*set, possibly proper, that fuses to $\mathcal{L}^+$. It is such subsets that we are really interested in.

SLP:6, §2.1, defines the narrower notion 'bounding set' in terms of two criteria:

(72) **Def. Bounding Set.** A set B⊆K, K a set of candidates, is a *bounding set* B(*z*) for *z*∈K relative to a constraint set Σ, iff B has these properties:

- **Strictness**. Every member of B beats *z* on at least one constraint in Σ.
- **Reciprocity**. If *z* beats some $b_i \in B$ on a certain constraint C∈Σ, then some other $b_k \in B$ beats *z* on the constraint C.

A bounding set is a certain kind of covering set. Of the conditions defining it, 'reciprocity' is the key, as shown in Samek-Lodovici & Prince 2001, where a nonempty set of candidates excluding *z* that meets reciprocity is called a *defeating set* for z.

(73) **Def. Defeating Set**. D⊆K, K a set of candidates, is a defeating set D(*z*) for *z*∈K iff D meets the following three conditions: (i) $D\neq\varnothing$, (ii) $z\notin D$, (iii) $\forall C\in\Sigma\ \forall d\in D, (z\succ d;C) \Rightarrow \exists b\in D\ (b\succ z;C)$.

Translated into present terms, the reciprocity condition (iii) asserts that, given a defeating set $D=\{d_i\}$ for *z*, whenever $C([z\sim d_j])$=W for some C then we also have $C([z\sim d_k])$=L for the same C and some other $d_k \in D$: this ensures that the C column fuses to L even when *z* earns a W by beating some member of the competitor set on C. The other possibilities, which do not invoke the reciprocity condition, are $C([z\sim d_i])=e$ or $C([z\sim d_i])$=L; no matter how these are distributed, any C column containing them will fuse to L, or to *e* when $C([z\sim d_i])=e$ for all i.

Let us introduce some notation to describe the kinds of situations that arise. We are interested in sets of arguments based on one desired optimum: $\{[z\sim d_i]: d_i\in D\}$. For this, let us also write $\{z\sim D\}$. For the fusion over the set, let us write $f i\{[z\sim d_i]: d_i\in D\}$ or $f_i\{[z\sim d_i]\}$, when we wish to emphasize the indexing, and simply $f\{z\sim D\}$when we do not.

We now show that the Samek-Lodovici & Prince 2001 notion of a defeating set, defined in (73), is coextensive with fusion to $\mathcal{L}^+$.

(74) **Remark**. D is a defeating set for z with respect to a set Σ of constraints iff $f_i[z\sim d_i]\in\mathcal{L}^+$, where the fusion is taken over all $d_i\in$D.

Pf. The LR direction is shown in the text. RL: Suppose $f_i[z\sim d_i]\in\mathcal{L}^+$. For any C_k, we have either [i] $C_k([z\sim d_i])=e$ for all i, or [ii] $C_k([z\sim d_j])=$L for *some* j. The first case raises no issues for reciprocity. In the latter case, we may still have $C_k([z\sim d_m])=$W for some m$\neq$j. In this case, z beats d_m; but d_j beats z, satisfying reciprocity. □

This equivalence allows the use of Proposition 2.4 to obtain the fundamental result of bounding theory: that the existence of a defeating set for z ensures that z is never optimal under any ranking, and conversely that any never-optimal candidate has a Defeating Set. Thus, harmonic bounding in both its simple and its collective forms reduces to the existence of the defeating set.

(75) **Proposition 6.1.** A candidate $z\in$K can never be optimal under any ranking of constraints Σ iff there exists for z a defeating set, D(z)$\subseteq$K.

Pf. LR: To say z can never be optimal is to say that $\{z\sim$K$\}$ is inconsistent. By Proposition 2.4, we must have some subset D$\subseteq$K such that f$\{z\sim$D$\}\in\mathcal{L}^+$. Then D forms a defeating set for z, by the remark (74).

RL. If a defeating set D(z)$\subseteq$K exists, then f$\{z\sim$D$\}\in\mathcal{L}^+$ which guarantees the inconsistency of $\{z\sim$K$\}$. □

To see where strictness comes in, consider the notion of a 'minimal defeating set': a defeating set which itself properly contains no defeating set.

A candidate d satisfies strictness against z iff $[z\sim d]$ shows L at some coordinate k, indicating that constraint C_k prefers d to z: $C_k([z\sim d])=$L. If some $d\in$D(z) fails strictness, then d beats z nowhere, so that $[z\sim d]\in\mathcal{W}^*$ and d is harmonically bounded by z. Any such harmonically-bounded candidate d obviously contributes nothing to the bounding effect of D with respect to z. Thus a defeating set consists of a (nonempty) central core satisfying strictness as well as reciprocity – a bounding set – and a periphery, possibly empty, consisting of candidates harmonically bounded by z.

The *minimal* defeating set will not include the $\mathcal{W}^*$ periphery; and therefore the members of the minimal defeating set satisfy strictness. (The converse is not true – satisfying strictness does not guarantee minimality.)

As an illustrative example of the construction, consider the simplest case, the one in which D(z) contains but a single element.

(76)

	C_1	C_2
d	3	5
z	3	7
z~d	*e*	L

The defeating set D(z) is just $\{d\}$. It is nonempty, excludes z, and satisfies reciprocity vacuously, because on no occasion is some member of D(z) beaten by z on a constraint. As expected, $f\{z\text{~}D\}\in\mathcal{L}^+$, albeit somewhat trivially, since $f\{z\text{~}D\}=[z\text{~}d]$.

The constitution of a defeating set D can be viewed from another angle, shifting the focus to the structure induced on the constraint set by D.

Since {z~D} fuses to $\mathcal{L}^+$ over Σ, we know that every individual constraint in Σ fuses over {z~D} either to L or to *e*. Thus a defeating set partitions the entire set of constraints into exactly two subsets: those constraints which treat the candidates in {z}∪D as equivalent, assigning *e* to every comparison in sight; and those on which some $d\in$D beats z.

(77) **Remark. Harmonic Bounding and the constraint set.**

D is a defeating set for z over Σ iff Σ is partitioned into two subsets as follows:

1. $\Sigma|e$: $\{C\in\Sigma\text{: } \forall x\forall y\ \ x,y\in D\cup\{z\} \Rightarrow C([x\text{~}y])=e\}$. Those constraints that do not distinguish, violationwise, the members of $\{z\}\cup$D.
2. Σ|L: $\{C\in\Sigma\text{: } \exists d\in D\ C([z\text{~}d])=L\}$. Those constraints on which *some* member of D beats z. Must be nonempty.

This characterization merely re-states the others: for, given the partition as described, $f\{z\text{~}D\}\in\mathcal{L}^+$.

It is clear, from a straightforward application of the transitivity lemma (69), that bounding of one candidate by another is an order relation, and therefore transitive: if *a* bounds *b*, and *b* bound *c*, then *a* bounds *c*. What of collective bounding? A generalized version of transitivity shows up as well in this case: if a set of candidates A covers *b*, and *b* belongs to a set B covering *z*, substituting A for b in B will produce a covering set for z.

(78) **Remark**. **Generalized transitivity of bounding**. Let A, B be sets of candidates. If A covers b, and b∈B, where B covers z, then (B−{b}) ∪ A covers z.

Pf. Only if *b* belongs to every minimal bounding set D(z)⊆B(z) is the assertion of interest; since otherwise the removal of *b* from B is without consequence for B's covering properties with respect to z. Consider therefore the fate of any minimal bounding set D(z)⊆B(z) with $b\in$D. We need only consider those constraints C on which C([z~b])≠W, i.e. C([b~z])≠L, for these are the only ones that contribute to the bounding effect – the fusion of {z~D} to $\mathcal{L}^+$ – either by providing an L that determines the fusion, or an *e* when all the other $d \in D$ also have $C([z\sim d])=e$. Since A covers b, we are guaranteed a defeating set $A'(b) \subseteq A$ which partitions the constraint set as in Remark (77). For any constraint C, then, either (i) $C([a\sim b])=e$ for all $a \in A'$ or (ii) there is some $g \in A'$ such that $C([g\sim b])=W$. In the first case, any $a \in A'$ will behave exactly like b on C. In the second, $\#g<\#b$ and since $\#b \leq \#z$, we have $\#g<\#z$ and $C([z\sim g])=L$. Thus the fusion to $\mathcal{L}^+$ is maintained. □

Harmonic bounding is notionally parallel to ERC entailment. Just as an entailed ERC is an uninformative accretion on the set of arguments, so a harmonically-bounded candidate adds nothing to the set of candidates that is not already implicit in its bounders. (For example, ranking is completely determined by comparison of possible optima, SLP:5.) The relation between bounding and entailment is more than a pleasant analogy, however, for bounding leads directly to the existence of entailment relations. The chief results are these:

- If A is a covering set for z, then $\{q\sim A\} \vdash [q\sim z]$ for any candidate q.
- If D is a defeating set for z, then $f\{q\sim D\} \vdash [q\sim z]$ for any candidate q.

A striking consequence emerges by contraposition. Suppose the set of ERCs $\mathcal{A}=\{q\sim A\}$ is free of entailments, in the sense that no subset $\Psi \subseteq \mathcal{A}$ entails any member of $\mathcal{A}$ not in Ψ. Let $K=A \cup \{q\}$, the candidate set underlying the ERC set. Then K is also free of harmonic bounding, in the sense that no subset of K is defeating set for any member of K.

Let us now establish these claims.

(79) **Proposition 6.2**. If A is a covering set for z, then $\{q\sim A\} \vdash [q\sim z]$.

Pf. Consider any model $R \vDash \{q\sim A\}$. We know that z is covered by A, but we know nothing about the relation between z and arbitrary q. If $R \vDash [z\sim q]$, then by the transitivity lemma (69), $R \vDash [z\sim a_i]$ for all $a_i \in A$. But since A covers z, $\{z\sim A\}$ contains a subset $D(z)$ with $fD \in \mathcal{L}^+$, so this cannot be true. Contradiction! So it is *not* the case that $R \vDash [z\sim q]$ for any R such that $R \vDash \{q\sim A\}$. Thus, $\{q\sim A\} \cup \{[z\sim q]\} = \{q\sim A\} \cup \{-[q\sim z]\}$ has no models, i.e. is inconsistent. By virtue of Lemma (23), we conclude $\{q\sim A\} \vdash [q\sim z]$. □

To obtain the further conclusion that $f\{q\sim D\} \vdash [q\sim z]$, for a defeating set D, it is helpful to generalize the transitivity property of '~' to include fusions.

(80) **Lemma. Generalized transitivity of '~'**. $[q\sim z], f\{z\sim A\} \vdash f\{q\sim A\}$.

Pf. The same approach works here as in Lemma (69): we show $[q\sim z] \circ f\{z\sim A\} \vdash f\{q\sim A\}$. Suppose that $C([q\sim z] \circ f\{z\sim A\})=W$ for a given constraint C. Then $C([q\sim z]) \neq L$ and $C(f\{z\sim A\}) \neq L$. So, on C, we have $\#q \leq \#z$ and $\#z \leq \#a_i$ for all $a_i \in A$, whence $\#q \leq \#a_i$ so that $C(f\{q\sim A\}) \neq L$. At least one of the inequalities must be strict, so $\#q<\#a_j$ for some a_j, so that in fact $C(f\{q\sim A\})=W$, as desired, giving satisfaction of the W-condition, Proposition 1.1(a). Now assume $C(f\{q\sim A\})=L$ for some C. This excludes the possibility that $\#q \leq \#a_i$ for all $a_i \in A$, so for some j, $\#q>\#a_j$ and $C([q\sim a_j])=L$. Thus we cannot have both $\#q \leq \#z$ and $\#z \leq \#a_j$, so that at least one of $C([q\sim z])$ and $C([z\sim a_j])$ is L. Consequently, $C([q\sim z] \circ f\{z\sim A\})=L$, satisfying the L-condition. □

It is also useful to note that the negative obeys a familiar logical property.

(81) **Lemma. Loose Excluded Middle**. For every ranking R and every ERC α, $R\vDash\alpha$ or $R\vDash-\alpha$.

Pf. If $\alpha=\delta$ then both disjuncts are true in R. Otherwise, the constraint C that is highest ranked constraint in R assigning a polar value to α is also the same for $-\alpha$, but with opposite polarity: C will assign W to one of $\{\alpha,-\alpha\}$. □

We call the Excluded Middle here 'loose' to emphasize the fact that *both* α and $-\alpha$ can hold in a ranking R when $\alpha=(-\alpha)=\delta$.

(82) **Proposition 6.3**. If D is a defeating set for z, then $f\{q\sim D\}\vdash[q\sim z]$ for any candidate q.

Pf. This is trivially true if $[q\sim z]=\delta$, so let us assume $[q\sim z]\neq\delta$. Suppose $R\vDash f\{q\sim D\}$. Assume for purposes of reductio that $R\vDash[z\sim q]$. By the generalized transitivity lemma (80), $R\vDash f\{z\sim D\}$. But this cannot happen, since D defeats z and $f\{z\sim D\}\in\mathcal{L}^+$. Ergo it is not the case that $R\vDash[z\sim q]=-[q\sim z]$. By the lemma of the loose excluded middle, $R\vDash[q\sim z]$. □

An interesting further consequence lurks in the proofs of the transitivity lemmas (69) and (80), which actually establish a stronger result than is stated, because they show that a *fusion* rather than a conjunction entails the desired consequent.. Throughout we have had to deal with the fact that $\mathcal{W}^*$ ERCs are mutually entailing, regardless of the W-condition: thus, $(W,e)\vdash(e,W)$, and vice versa. In the case at hand, however, the W-condition is strictly observed between the consequent and the fusion of the antecedents, even when they are trivial.

Consider a situation in which a harmonically bounds b. We know from Proposition 6.2 that $[q\sim a]\vdash[q\sim b]$. Suppose in addition that q harmonically bounds a i.e. that $[q\sim a]\in\mathcal{W}^*$. Then both antecedent and consequent belong to $\mathcal{W}^*$. But we are still assured that the W-condition holds, so that $C([q\sim a])=W$ implies $C([q\sim b])=W$ for any C. This shows that even for universally valid ERCs, the W-condition can have significant impact. The logic of implication that lies behind this finding will be taken up in §7.

Because bounding in a candidate set A leads to entailment relations in the associated ERC set $\{q\sim A\}$, the *lack* of entailments within such a set will, as noted above, contrapositively lead to the lack of bounding relations in $A\cup\{q\}$. Consequently, if an ERC set of the form $\{q\sim A\}$ is entailment-free – a state easily achievable according to the prescriptions of §5 – then its underlying candidate set $A\cup\{q\}$ is also bounding-free.

(83) **Corollary to Proposition 6.2.** Let $\mathcal{A}$ be a set of ERCs of the form $\{q\sim A\}$ and let $K=A\cup\{q\}$. If for each $a_i\in A$ there is no subset $E\subseteq A$, $a_i\notin E$, such that $\{q\sim E\}\vdash[q\sim a_i]$, then no $z\in A$ is covered by A.

Pf. The existence of a covering set $E(z)\subseteq A$ implies that $\{q\sim E\}\vdash\{q\sim z\}$ by Proposition 6.2, contrary to hypothesis. □

Note that this result includes the case of simple bounding. Suppose $K=\{q,a\}$, where q bounds a. Then $[q\sim q]\vdash[q\sim a]$ because $\vdash[q\sim a]$.

Eliminating entailment therefore eliminates bounding. But it is important to notice a subtlety in the process. To say that there is no bounding in a set K, with K⊆K′ where K′ is a fuller version of the candidate set, is *not* to say that the elements of K are possible optima: they may be bounded in K′.

This observation is relevant to the typical kind of case where one operates on an ERC set to free it of entailments. Suppose we start out with an ERC set $\mathcal{A}'$={q~A′}, with underlying candidate set K′={q}∪A′. We then remove as many ERCs as is needed to arrive at an entailment-free set (the selection process need not be unique), arriving at the ERC set $\mathcal{A}$={q~A}⊆$\mathcal{A}'$, with candidate set K. We know that the K is bounding-free. But K′ might still cover elements of K.

To see this, consider cases where α⊢β and β⊢α simultaneously. If α and β belong to an ERC set $\mathcal{A}'$ that we are stripping of entailments, we have a free choice between removing α and removing β. But in this case we cannot be sure, from W,L-structure alone, that we are correctly removing a candidate that is bounded in K′.

Consider the following example:

(84)

	C_1	C_2
α: q~*a*	W	L
β: q~*b*	W	L

Among the violation profiles compatible with these ERC vectors are the following:

(85)

	VP-I. {*a*,*b*} independent		VP-II. *a* bounds *b*		VP-III. *b* bounds *a*	
	C_1	C_2	C_1	C_2	C_1	C_2
q	0	2	0	2	0	2
a	1	1	1	0	2	1
b	2	0	2	1	1	0

In VP-I, candidates *a* and *b* are independent, neither bounding the other; but in establishing q as optimal, they both supply the same ERC, so only one is needed.

In VP-II, candidate *a* bounds candidate *b* and in VP-III the opposite obtains. If we remove α from the set of ERCs under VP-I, then we are left with *b* in the candidate set, where *b* is not a possible optimum (i.e. over $\{C_1,C_2\}$– but the point is of general applicability).

Similar examples can be constructed with collective bounding. Consider the following case:

(86)

	C_1	C_2
q	0	2
b	2	1
x	3	0
z	1	2

Constructing the ERC set based on q, we find:
(87)

	C_1	C_2
q ~ *b*	W	L
q ~ *x*	W	L
q ~ *z*	W	*e*

Candidate *z* is bounded by q, and [q~*z*] can be removed from the ERC set without prejudice. But [q~*b*] and [q~*x*] tell exactly the same story – each entails the other – and so one of them can be removed. If [q~*x*] is removed, then only [q~*b*] remains, leaving a candidate set {q,*b*}. But over the *original* candidate set {q,*b,x,z*}, *b* is defeated by {x,z}, as can be seen from the following:

(88)

	C_1	C_2
b~ *x*	W	L
b ~ *z*	L	W
[*b*~ *x*]∘[*b* ~ *z*]	L	L

The significant structural relations in the candidate set can be easily discerned in the following diagram, which indicates the orders imposed by the violation patterns. Observe how *b* is bracketed by *x* and *z*.

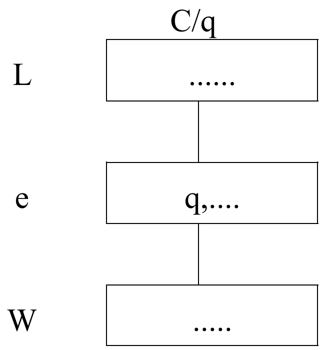
(89)

C_1	C_2
q	***x***
|	|
z	*b*
|	|
b	*z*, q
|	
x	

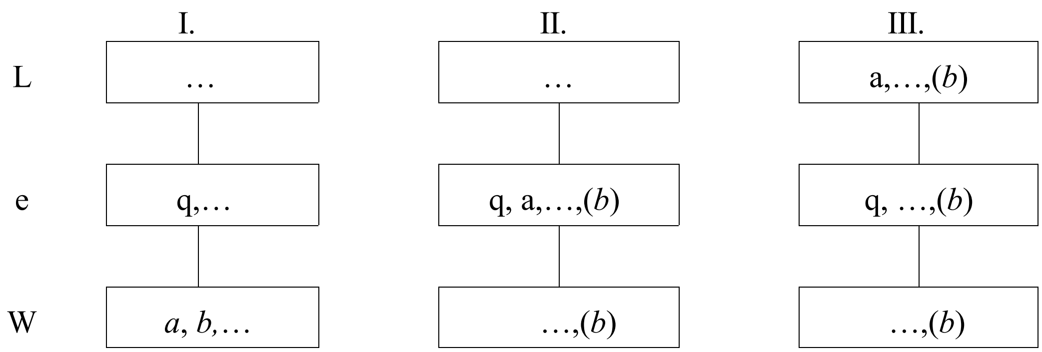
The appropriate response to this observation will depend on the use to which the candidate set K′ is being put. For many purposes, it will be far more important to eliminate entailments – uninformative ERCs – than to eliminate bounded candidates, which as we have seen can still provide informative ERCs. But when bounded candidates are the focus of interest, the phenomenon of mutual entailment can serve as an indicator that further scrutiny is required.

It is instructive to look at the relation between entailment and the kinds of orders imposed by constraints. Given a desired optimum q∈K, the set K is divided into 3 parts with respect to any constraint C: those candidates better than q, those with the same status, and those worse than q. These are the *x*∈K such that C assigns to [q~*x*] respectively L, e, and W. The result looks like this:

(90)



For [q~*a*]⊢[q~*b*], for [q~*a*] nontrivial, only three situations arise:
(91)



The following remarks may be made:

• Case I: C([q~*a*]) = W. By the W-condition, we must have C([q~b])=W.

• Case II: C([q~*a*]) = *e*. Here C([q~b])=*e* or C([q~b])=W.

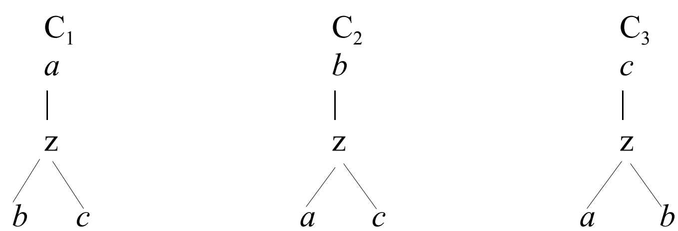
• Case III. C([q~*a*]) = L. Anything goes.

In short, if [q~*a*]⊢[q~*b*], trivial ERCs aside, then for every C, we must have a≤b in the C/q order, where equality of order means membership in the same equivalence class.

This result shows why bounding yields entailment: if *a* bounds *b*, then #*a*≤#*b* on every constraint – the violations of *a* are never greater than those *b* – which imposes a stronger ordering condition. Diagram (91) also illustrates exactly where the converse fails: when *a,b* are both in class W or both in class L, the relation between #*a* and #*b* is hidden in C/q, and indeed any relation is possible.

We conclude with an observation about the relationship between optimality, harmonic bounding, and entailment. An ERC [q~*z*] over a constraint set ℂ completely determines the ranking conditions on ℂ under which q is *better than or equal to z* evaluatively. But this one ERC provides no more than a necessary condition for the *optimality* of q, since optimality requires that q never be bested by any competitor at all. How many competitors must be examined to ensure optimality? It is certainly sufficient to examine all the possible optima for any ranking of ℂ; if q survives comparison with all of these under a certain ranking R, it will necessarily be optimal on R [SLP:5,43]. Along the same lines, if q is compared successfully with a set B, where {q}∪B covers all other members of q's candidate set, then q is sure to be optimal on R as well. But neither of these tactics is necessary, and there are in fact cases where optimality can established by comparison with elements that are neither optimal themselves nor bounders for the set of competitors. To see this, consider the following example (constructed in collaboration with Vieri Samek-Lodovici).

(92)



Candidate a is rendered optimal when C_1>> {C_2, C_3} and only then. This result is secured directly from a single comparison with z, which is never optimal under any ranking.

(93)

	C1	C2	C3
a~z	W	L	L
a~*b*	W	L	
a~*c*	W		L

As shown, the single comparison may be (inefficiently) replaced with two, each running against one of the other possible optima. But it suffices to compare with the never-optimal, non-bounder z.

The key is that [a~z] *entails* both [a~b] and [a~c]. The limit on the set of competitors is this: if the ERCs constructed from a competitor set K′ *entail all possible ERCs* projectable from elements of the entire candidate set K, then comparison with K′ guarantees optimality; and conversely.

7. The Logic of Optimality Theory

Ex falso quodlibet.
– *J. D. Scotus*, overoptimistically
Rank hath its privileges.
– *C.G. Prince*, on the US Navy

Summary. The fusion and negative operations on ERCs explored above correspond to a three-valued propositional logic, the implication/negation fragment of the relevance logic RM. Taken whole, RM itself provides the logic of OT. Theorems of RM may therefore be freely imported. Returning the favor, it is seen that the 'system' – a set of RM valuations related by permutation of ranking – has properties that make it a structure of independent logical interest.

7.1 OT as Logic, and *v.v*

The pattern of relations examined here fits into a larger formal structure that can be identified as propositional calculus-like logic with a semantics based on not two but three truth-values: T,F,*e*, corresponding to the W,L,*e* of ranking theory. An ERC evaluates the assertion '*a is better than b*' for two candidates *a*,*b*, and each constraint C_k records one of three views on the matter:

- W/T – the assertion is true for C_k, and strictly so, since $a \succ b$ on C_k
- L/F – the assertion is false for C_k, and strictly so, because $b \succ a$ on C_k
- *e* – C_k offers no illumination because neither of the above holds.

As seen in §2, there is a natural definition of 'negation' in this system. This leads immediately to another operation, dual to fusion, which resembles logical disjunction in the same way that fusion resembles logical conjunction. Following standard usage in the literature, we notate this connective as '+' and (biting the terminological bullet) refer to it as 'fission'.[12]

(94) **Major connectives**

Negation

φ	¬φ
T	**F**
e	*e*
F	**T**

Fusion

ψ…

φ ⋮

φ ∘ ψ	T	*e*	F
T	**T**	T	**F**
e	T	*e*	F
F	**F**	F	**F**

Fission

ψ…

φ ⋮

φ + ψ	T	*e*	F
T	**T**	T	**T**
e	T	*e*	F
F	**T**	F	**F**

12 Passing with regret over András Kornai's droll proffer: "co-fusion".

Bolded values emphasize the complete identity of these connectives with their PC correlates over the truth values T and F. Fission is not an independent notion, but can be defined in terms of fusion and negation, in very much the usual way:

(95) $\varphi+\psi =_{df} \neg(\neg\varphi\circ\neg\psi)$.

We also have a notion of 'arrow', the conditional connective, again defined as expected:

(96) $\varphi\rightarrow\psi =_{df} \neg(\varphi\circ\neg\psi) = \neg\varphi+\psi$

Choice of '$\circ$' as basic is arbitrary: the connectives $\{+, \circ, \rightarrow\}$ are interdefinable, given '$\neg$'. The truth table for '$\rightarrow$' then runs as follows:

(97) **The conditional connective**

	ψ…		
$\varphi\rightarrow\psi$	T	*e*	F
φ… T	**T**	F	**F**
e	T	*e*	F
F	**T**	T	**T**

The notion of 'arrow' shown here underlies the properties of entailment established in §1. The inferential legitimacy of 'W-extension' is echoed in the fact that '$e\rightarrow$T' is awarded T; and that of 'L-retraction' in the T awarded to 'F$\rightarrow e$'.

We define the *designated values* in the system to be $\{T,e\}=\Delta$. If a wff assumes a *designated value* under an assignment of truth values, it is said to be *true* under that assignment. If a wff always assumes a designated value, regardless of the evaluation of its constituents, it is *valid*. This corresponds to the fact that any $\varphi\in\mathcal{W}^*$ interprets to a ranking condition that is true under every ranking, in the ordinary 2-valued sense.

A remarkable fact about this logic is that there are no *invalid* wffs – no formula takes on the nondesignated value F for every assignment. If all the prop letters are assigned *e,* then the whole thing evaluates to *e*. For example, the canonical contradiction pattern is $\varphi\wedge\neg\varphi$, but the analogous formula $\varphi\circ\neg\varphi$ is *e* (and designated) whenever φ is *e*.

With these value assigments in hand, it is easy to check that the resulting logic shows many of the familiar properties of propositional systems. For '$(\alpha\rightarrow\beta) \wedge (\beta\rightarrow\alpha)$', a strong form of equivalence demanding that α and β assume the same value under any valuation, we write '$\alpha=\beta$'.

(98) **Basic Properties of the Connectives**.

(i) $\neg\neg\varphi = \varphi$	Double negation
(ii) $\alpha\rightarrow\varphi = (\neg\varphi)\rightarrow(\neg\alpha)$	Contraposition
(iii) $\neg(\alpha\circ\varphi) = \neg\alpha + \neg\varphi$	De Morgan
(iv) $\neg(\alpha+\varphi) = (\neg\alpha)\circ(\neg\varphi)$	De Morgan
(v) $(\alpha\circ\varphi \rightarrow \psi) = (\alpha\rightarrow(\varphi\rightarrow\psi))$	Trans(ex-/im)portation

Certain properties are missing, though, as we have seen repeatedly above: for example, $\alpha\circ\varphi\rightarrow\varphi$ and the closely related $\varphi\rightarrow(\alpha\rightarrow\varphi)$ are nontheorems. Furthermore, we do not have the fusional analog of '*ex falso quodlibet*', which would read $(\alpha\circ\neg\alpha)\rightarrow\varphi$, because $\alpha\circ\neg\alpha$ is not in fact 'false' on all occasions. Nor do we have its contrapositive congener '*verum ex quodlibet*', $\varphi\rightarrow(\neg\alpha+\alpha)$: even though the consequent is universally designated, it is not 'arrowed' by a random wff. (To see this, calculate its value with the assignment φ=T, α=*e*). Finally, we note that distribution of '∘' over '+' and of '+' over '∘' also fails (see p. 84 below for some discussion).

The system laid out in (94) might be called 'propositional calculus with identity': the value *e* serves as an identity with respect to the connectives '∘' and '+', but relations between T and F are arranged in the normal way. The sense of a third, or intermediate, truth value is often subject to philosophical disputation; especially when its role is to illuminate certain refractory conundra. In the present case, the *tertium quid* offers no difficulties whatever: it serves simply to indicate that no assessment is made; that the assessing item (constraint, predicate, prop letter, formula) is *irrelevant* to the outcome. The system, then, might be called the pure logic of irrelevance, or *LI* for short.

The pattern of truth value assignments just laid out has been interpreted in terms of 'ambivaluation', by which propositions assume *sets* of truth values, and can therefore be *both* truth and false. Under this conception, what we call '*e*' is represented as {T,F}, and is therefore often named 'b' in the literature. This difficult notion receives some clarification, perhaps, from the concrete, order-theoretic model we are dealing with. A stratified partial order, such as is imposed by a constraint on a set of candidates, recognizes three relations: $a \succ b$, $a \prec b$, and $a \approx b$, with the last meaning that *a* and *b* occupy the same stratum. If the basic unit of analysis is an assertion of the form $a \succeq b$, then the meaning of the values is straightforward:

- $T \in val(a \succeq b)$ iff $a \succeq b$.
- $F \in val(a \succeq b)$ iff $b \succeq a$.

If $a \approx b$ holds, then *both* $a \succeq b$ and $b \succeq a$ are true. Hence, by these rules, $val(a \approx b)=\{T,F\}=e$.

With the truth table for '→' in hand (97), it becomes easy to recognize that '∘' and '+' are the connectives originally discussed in the context of relevance logic as 'intensional conjunction' and 'intensional disjunction', more recently 'fusion' and 'fission'.[13] Sobociński 1952 presents an

---

[13] Anderson & Belnap 1975 [AB], p. 344-6, prefer to call '∘' *co-tenability*; other terms include 'consistency' (from Dunn 1966, who introduced the notation) and 'fusion' (from Meyer, as noted in http://www.cis.upenn.edu/~bcpierce/types/archives/1992/msg00002.html), with the last eventually winning out in the literature. Meyer 1973:228ff cites Church (no specific ref.) as the originator of these connectives, and notes several dissertations in which they are studied: Belnap 1959, Dunn 1966, Meyer 1966; AB also

axiomatization of a system using only '¬' and '→', with the truth tables given above. This is exactly the basis for our logic, which is better but more opaquely called S, after its first investigator; as noted, the connectives '∘' and '+' can be treated as abbreviations for '¬(p→¬q)' and '¬p→q' respectively (designated 'K' and 'A' in Sobociński 1952:52, §4.6). Parks 1972 shows that Sobociński's system constitutes the implication-negation fragment of RM (Cf. AB:148), the significance of which will emerge shortly. Logics of the RM family also include the standard connectives '∧' and '∨'.[14] The logic RM3 uses Sobociński's values for fusion, fission, negation, and arrow. Here we give the RM3 truth tables for '∧' and '∨'.(As in AB:470; note also that these are the usual Łukasiewicz/Kleene assignments: Rescher 1969:23.) Taken together, the assignments in (97) and (99) are known as the Sugihara 'matrix' called M_3 in AB and S_3 in Meyer 1973.

---

mentions Woodruff 1969. [Dissertationēs nondum vidī]. Belnap 1960 introduces fusion into relevance logic [AB:345]. Meyer 1973 and Meyer 1975 (i.e. AB,§29.3:393-420, and §29.12, esp. p. 470), a chapter of AB, provides further discussion in the context of the logics R, RM and RM3. AB: 344-346 reviews both 'intensional' connectives; Anderson, Belnap & Dunn 1992 [ABD] contains more discussion *passim*. Restall 2000 defines fusion in various contexts. Fusion corresponds as well to the 'multiplicative conjunction' or 'tensor (product)' of Linear Logic. AB also cites Fisk 1964, who provides (pp. 51-53, 58,60) some brief general discussion of intensional uses of disjunction and of implication, from which it appears that the connectives we are dealing do not entirely resolve these notions. For Fisk, intensional uses of 'either/or' are 'true' if relevant alternatives are exhaustively mentioned, and not merely when at least one disjunct is true; intensional implication requires some kind of appropriate connection between antecedent and consequent, and not merely that the latter be true when the former is. It seems little of this is guaranteed for + and →.

[14] Relevance logics include ∧, ∨, ⊃, ¬ and typically include the theorems of Propositional Calculus. But '→' is not defined in terms of these : it has its own axioms. Modus ponens works over '→' not over PC '⊃', so certain deductions are disallowed. E.g. A⊃(B⊃A) is a theorem, but A→(B→A) is not, therefore from the premise A, one may *not* deduce B→A, by virtue of which something true would be implied by anything at all. The originating idea is to formalize the notion of 'entailment', which when put to use involves some relation of 'relevance' between premises and conclusions. (More extremely, some relevantists have objected to admitting 'A→A', on the grounds that one doesn't go around saying "2+2=4 entails 2+2=4.") As we will find here, the logics have concrete applications that transcend their embattled philosophical origins; this is also made clear in the accessible treatment of Restall 2000.

The relevance logic of greatest importance in the present context is RM , an extension of the basic relevance logic R, obtained by adding the axiom A→(A→A). A theorem-determining semantics for RM requires twice the number of truth values as there are letters in the formula under scrutiny; so that it requires an infinite set of truth values to handle the set of all wffs. RM3 is obtained by adding on a couple more axioms, and has the virtue that only 3 values are required for a complete semantics. Aspects of Meyer's illuminating discussion in AB:§29.3.2 will be presented below. Perhaps the most surprising feature of relevance logics, for the newcomer, is the drastic increase in complexity that is caused by dropping a few features of familiar Propositional Logic (forget quantification): even finding a semantics requires major skill and insight, and some of the most basic systems (E,R) are undecidable. One is struck by the staggering level of success achieved within the simplicities of classical logic.

(99) RM3 $\wedge$ and $\vee$

Conjunction

$\psi\ldots$

	$\varphi \wedge \psi$	T	*e*	F
φ	T	T	***e***	F
$\vdots$	*e*	***e***	*e*	F
	F	F	F	F

Disjunction

$\psi\ldots$

	$\varphi \vee \psi$	T	*e*	F
φ	T	T	T	T
$\vdots$	*e*	T	*e*	***e***
	F	T	***e***	F

The places where *e* behaves as a non-identity in (99) are shaded. With '$\wedge$' and '$\vee$' in hand, we can also develop a notion of horsehoe '$\supset$', defined as $\neg A \vee B$. To facilitate comparison, we present the RM3 tables for '$\supset$' and '$\rightarrow$' side by side.

(100) **The conditional connectives in RM3**

Horseshoe

$\psi\ldots$

	$\varphi \supset \psi$	T	*e*	F
φ	T	**T**	*e*	**F**
$\vdots$	*e*	T	*e*	e
	F	**T**	T	**T**

Arrow

$\psi\ldots$

	$\varphi \rightarrow \psi$	T	*e*	F
φ	T	**T**	F	**F**
$\vdots$	*e*	T	*e*	F
	F	**T**	T	**T**

The cells where the two differ are shaded. As with the other connectives shown above, the difference between the extensional ($\wedge$, $\vee$, $\supset$) and the intensional ($\circ$,+,$\rightarrow$) connectives lies in the treatment of combinations involving *e*: the intensional connectives impose further distinctions among combinations that neutralize to *e* in the extensional set.

(101) **Intensional/extensional differences in RM3**

$e \circ e \neq \mathrm{T} \circ e = \mathrm{T}$ $\quad$ $e + e \neq \mathrm{F} + e = \mathrm{F}$ $\quad$ $e \rightarrow e \neq \mathrm{T} \rightarrow e = \mathrm{F}$ $\quad$ $e \rightarrow e \neq e \rightarrow \mathrm{F} = \mathrm{F}$

$e \wedge e = \mathrm{T} \wedge e = e$ $\quad$ $e \vee e = \mathrm{F} \vee e = e$ $\quad$ $e \supset e = \mathrm{T} \supset e = e$ $\quad$ $e \supset e = e \supset \mathrm{F} = e$

All these connectives share certain fundamental properties: associativity, commutativity, idempotence (a *op* a = a), and what we might call 'T/F-dominance'. Both '$\wedge$' and '$\circ$' are F-dominant, while '$\vee$' and '+' are T-dominant: the composition takes on a dominant value held by either of its parts. The Łukasiewicz assignments extend the notion of dominance: assuming the ternary scale T>*e*>F, '$\wedge$' takes on the least and '$\vee$' the greatest of its component values. The valuation of '+' can be described as taking on the greatest value on the auxiliary scale[15] T>F>*e* (AB:470); but this amounts to treating *e* as an identity. The innovation in these relevance logics, and in Sobociński's precursor, is defining the implication connective '$\rightarrow$' not from '$\wedge,\vee$' but from '$\circ,+$' (or vice versa), a move that we heartily endorse.

To make use of the logic S, we need to bridge the gap between assigning values to single formulas and assigning values to the vectorial aggregates formed by individual ranking arguments. Above, we have treated the ERC vector as a basic object and we have simply extended operations to it coordinatewise. Viewing the entries as truth values suggests a change in perspective, which will embed OT in an appropriate logic.

In ordinary circumstances, a propositional logic recognizes a valuation function which assigns truth values to all wffs: a valuation maps the basic propositional variables ("prop letters") to truth values, and is extended to complex formulae via the assumption of civilized relations with the connectives: for example, $v(A\wedge B)=v(A)\cdot v(B)$ in standard propositional logic, with T=1 and F=0. For the logic S or RM3, a valuation v is a function from the atomic prop letters to {T,*e*,F} or {−1,0,1} or some such 3 element set with the same structure, which is extended to general wffs in accord with the tables given above.

With an ERC represented as single letter, an ERC vector consists of *multiple* valuations of the same entity. Each constraint is then a single basic valuation function. We may think of an ERC vector, then, as representing the *product* of a number of such basic valuations, which we will call a *polyvaluation*.

(102) **Vectorial S** (VS). **Polyvaluation**.

Let v_i, $1\leq i\leq n$ be valuations, not necessarily distinct, from the set Ω of S-wffs to a three element set $\mathfrak{Z}$, so that we have v_i: $\Omega\rightarrow\mathfrak{Z}$. The *product* $V=\prod v_i$ is the polyvaluation function $V:\Omega\rightarrow\mathfrak{Z}^n$, with $[V(\alpha)]_k=v_k(\alpha)$ for any $\alpha\in\Omega$.

This definition immediately gives us the 'coordinatewise' effect we are seeking for the treatment of the connectives . From it, we have:

$$V(\alpha\circ\beta) = (v_1(\alpha\circ\beta),\ldots,v_n(\alpha\circ\beta))$$
$$V(\neg\alpha) = (v_1(\neg\alpha),\ldots,v_n(\neg\alpha))$$

The coordinate valuations $v_k(\alpha\circ\beta)$ and $v_k(\neg\alpha)$ are just those defined in S from $v_k(\alpha)$ and $v_k(\beta)$.

As noted above, a wff ψ of S is 'true' under a valuation v if $v(\psi)\in\Delta=\{e,T\}$. For this we can write $\Vdash_v\psi$ or $v\Vdash\psi$, using the special symbol to emphasize that we are working outside standard propositional logic.

It is natural to extend the notion of truth to a polyvaluation V by requiring that *every* v_i in V yields a designated value.

(103) **Truth and Validity in** VS.

Let $V=\prod v_i$ be a polyvaluation. A wff ψ of VS is *true* under V iff $\forall v_i$ inV, $v_i(\psi)\in\Delta$. For this we write $\Vdash_V\psi$ or $V\Vdash\psi$. A wff ψ is *valid* iff it is true in every polyvaluation, written $\Vdash\psi$.

---

[15] The notion of an auxiliary scale is motivated by Meyer 1975 (AB:400, ex. iii) in the context of RM, where more truth values are at play. See below p. 56 ff. for development and reformulation.

With two ways of evaluating the same set of formulae, it's worth asking what relationship holds between them. From one point of view, it is particularly simple: any theorem of S is a theorem of VS, and vice versa.[16] If ψ is true on *all* S valuations, then it is true over any collection of them, that is, for any polyvaluation; and if ψ is true for all polyvaluations, it is certainly true for all simple valuations (because every simple valuation is also a polyvaluation, over a one-element set).

(104) **Remark. Theorem-Equivalence of S and VS.** Any theorem (equivalently, valid wff) of S is a theorem of VS and vice versa.

Pf. As in text.

This doesn't mean that S and VS are indistinguishable. Generalizations framed with reference to *single* (poly)valuations, rather than with respect to *all* valuations, can come out differently in the two logics. Let's call a wff 'antidesignated' under a given valuation if its negative is designated; and 'nondesignated' if it is merely not designated. In S, the notions 'antidesignated' and 'nondesignated' are coextensive, but they part company in VS.

In S, it is true that for any valuation v, we must have (nonexclusively) either $v \Vdash \psi$ or $v \Vdash (\neg\psi)$: a wff is either designated, or antidesignated, with a designated negative, or both, if $v(\psi)=e$. But this meta-statement is not true for single polyvaluations: there are wffs ψ such that neither $V \Vdash \psi$ nor $V \Vdash (\neg\psi)$. For example, suppose $V(\psi)=(T,F)$; then $V(\neg\psi)= (F,T)$ and neither of the mutual negatives ψ and $\neg\psi$ is true under V: neither consists only of designated values.

Along the same lines, consider that under simple valuation and polyvaluation alike, if φ is designated and ψ is not, then $\varphi\circ\psi$ is nondesignated; but in the case of simple valuation though not of polyvaluation, we also have it that $\varphi\circ\psi$ is antidesignated, *i.e.* that $\neg(\varphi\circ\psi)$ is designated. Under simple valuation, the fusion of designated and nondesignated wffs is false, because here 'antidesignated' and 'false' are synonymous. The failure of this property to extend to polyvaluations shows up in the kind of 'nontruthfunctionality' of fusion noted above (§2, p. 9). For example, designated (T,e) and antidesignated (e,F) fuse to (T,F), but (T,F) is not 'false' or antidesignated. Finally, observe that under any simple valuation, we must have $v \Vdash \alpha \rightarrow \beta$ or $v \Vdash \beta \rightarrow \alpha$; yet this is manifestly not guaranteed under polyvaluation, due to the requirement that the same formula must hold in every coordinate.

The upshot is that polyvaluation distinguishes more finely among wffs than does simple valuation. S already introduces a distinction within the set of 'true' or designated wffs that is not made in standard two-valued logic. Polyvaluation imposes yet further distinctions throughout the system. For a product of n valuations, there are fully 2^n distinct kinds of designated wffs, ranging from $(e,e,...)$ to (T,T,...), with every combination of e's and T's in between. Of these, only (T,T,....) acts exactly like the simple (T) of the single valuation. Similarly, among antidesignated wffs, there is a parallel set of distinctions running from (F,F,...) to $(e,e,...)$. And between these two classes, there is brought into existence a class of wffs which are neither designated nor antidesignated. In essence, polyvaluation blows up the middle region of value space, replacing the simple 3-point lattice ($\mathsf{3}$) based on the relations $F \leq e \leq T$ with a far richer one ($\mathsf{3}^n$).

---

[16] S (and RM3) is complete with respect to the semantics given, so there is no harm in using 'theorem' to mean also 'valid wff'.

To reach terra firma, we must relate calculations in VS to the familiar kind of PC logical relations among ERCs and sets of ERCs. Above, we have done this by associating each polyvaluation with a expression in ordinary predicate calculus. We may then pass back and forth between statements in VS and PC: for example, writing h for the function assigning V(α) to an ERC, if $V \Vdash \alpha \rightarrow \beta$, then $\vdash h(V(\alpha)) \supset h(V(\beta))$.

As we have seen throughout, handling the converse requires some care with valid and invalid wffs, because '$\vdash$' has a coarser sense of these notions than '$\Vdash$'. For example, $V \Vdash \alpha \rightarrow \beta$ means that both the W-condition and the L-condition are satisfied without qualification, since by the the laws of S valuation for '$\rightarrow$', given in (97), T at $[\alpha]_k$ must correspond to T at $[\beta]_k$, and F at $[\beta]_j$ must correspond to F at $[\alpha]_j$. Thus, although $\vdash (T,e) \supset (e,T)$ under the rubric 'verum ex quodlibet', it is manifestly not the case that $\Vdash (T,e) \rightarrow (e,T)$.

7.2 Beyond VS to RM

To minimize shuttling between VS and PC, the entire scheme of relations can be brought under the one roof of the logic RM, which contains both the intensional and the extensional connectives. RM has the following property, to be demonstrated shortly.

(105) **RM/PC**. Suppose a formula A contains none but the connectives $\wedge$, $\vee$, $\supset$, $\neg$. Then A is a theorem of RM iff A is a two-valued tautology.[17]

In essence, RM hospitably takes in PC in its entirety, along with S, and regulates the behavior of every mixture of the two. This allows us to operate safely with any wff using the entire range of intensional and extensional logical apparatus. And when we cash out intensional expressions of whatever complexity for single ERCs, we are left with a sentence that is fully and correctly interpretable in standard terms.

ERC fusion puts VS expressions like $A \circ B$ in 1:1 correspondence with PC expressions containing only extensional connectives. The relation between '$\neg_{RM(k)}$' and PC '$\neg$' is only slightly more complicated. This means that the transition from RM to PC will be smooth.

For example, $(A \wedge B) \rightarrow (A \circ B)$ is a theorem of RM (R55, AB:397), which leads immediately to $(A \wedge B) \supset (A \circ B)$, *i.e.* $A \wedge B \vdash A \circ B$, since $X \rightarrow Y$ entails $X \supset Y$. But this is just Lemma (14), which will be given to us outright by RM/PC (105) if we are successful in identifying RM as the logic of OT. To this end, we develop the basic semantics of RM. (For RM proof theory, see Avron 1987, 1991.)

The relation between RM and ERC logic is most easily determined through the RM semantics developed by R. K. Meyer (reported as Meyer 1975:400ff.) and further developed in Dunn 1970. Meyer shows that RM is complete with respect to a multivalued semantics. The fundamental insight is that the relevant sets of truth values have the property, when represented as integers, that the integer $-k$ is included in the set whenever k is; staying close to Meyer's usage, let us call such a collection of integers a 'Sugihara set' (Sugihara 1952).

[17] AB:396 state this for $\wedge$, $\vee$, $\supset$, $\equiv$ only. But see Anderson & Belnap 1959, Dunn 1985:149.

The RM3 values{F,*e*,T}, for example, can be represented by {−1,0,1}, by {−71,0,71}, or indeed by any set {−k,0,+k}, though for concreteness and ease of exposition, we will typically instantiate Sugihara sets with the smallest possible integers that fit the description. Let us label each Sugihara set as S_N, where N is its cardinality (here, we take the liberty of modifying Meyer's notation). Two kinds of Sugihara sets will play a role in the following discussion: those which contain 0, and those which do not. These can be distinguished as S_{2N} and S_{2N+1}. For example, S_5={−2,−1,0,1,2}and S_4={−2,−1,1,2}, under the convention of smallest integer representation. Extending to infinite sets, we have, writing Z for the integers and borrowing from Dunn the notation Z* for the integers except 0,

S_Z = {...,−3,−2,−1,0,1,2,3,...}

S_{Z*} = {...,−3,−2,−1,1,2,3,...}

The non-negative values will always be designated.

Valuation of the extensional connectives and of negation contains no surprises.

(106) **Valuation on Sugihara Sets (Negation, Conjunction, Disjunction)**

Let I be Sugihara, *i.e.* a set of integers such that k∈I => −k∈I. Let v map prop letters to I. We extend v to complex sentences as follows:

(i) Negation v(¬P) = −v(P)
(ii) And v(P∧Q) = min[v(P),v(Q)]
(iii) Or v(P∨Q) = max[v(P),v(Q)]
(iv) H'shoe v(P⊃Q) = max[v(¬P),v(Q)]

Note that (ii) - (iv) are not independent, because '∧', '∨', and '⊃' are interdefinable, given negation.

These definitions extrapolate straightforwardly from the familiar patterns:

- P∧Q is designated iff neither P nor Q is negative – 'false'.
- P∨Q is designated iff either is *non*negative – 'true'.
- P⊃Q is designated iff P is nonpositive or Q nonegative (beware zero in the antecedent!).

The negative and nonegative stretches of value space form vast equivalence classes with respect to these connectives, retaining amid apparent multiplicity the simple familiar sense of F and T. To give teeth to this intuitive formulation, we state it as the the following:

(107) **Lemma. Extensionality.** Let φ be any sentence using only the extensional connectives and negation. Let S_K be any Sugihara set. Let v be any valuation on S_K conducted according to the rules in (106). Let v* be the valuation defined as follows, for every prop letter P:

v*(P) = +1 if v(P)>0
v*(P) = 0 if v(P)=0
v*(P) = −1 if v(P)<0.

Then v(φ) <0 iff v*(φ)<0 and v(φ)=0 iff v*(φ)=0.

Pf. By straightforward induction on the number of connectives in φ. If none, the thesis follows trivially from the definition. Now assume that the thesis holds up to n connectives, and examine formulae containing n+1. Since all extensional connectives can be defined in terms of $\neg$ and $\wedge$, we have only two cases to consider.

(i) φ is $\neg\varphi_1$. Then $v(\varphi)<0$ iff $v(\varphi_1)>0$ (definition of $\neg$) iff $v^*(\varphi_1)>0$ (induction hypothesis) iff $v^*(\varphi)<0$ (def. of $\neg$). Along the same lines, $v(\varphi)=0$ iff $v(\varphi_1)=0$ iff $v^*(\varphi_1)=0$ (induction hypothesis) iff $v^*(\varphi)=0$.

(ii) φ is $\varphi_1\wedge\varphi_2$. $v(\varphi)<0$ iff at least one of the conjuncts is negative, by the definition of $\wedge$; call one such conjunct φ_{neg}. Then $v^*(\varphi_{neg})<0$ iff $v^*(\varphi_{neg})<0$ (induction hypothesis) iff $v^*(\varphi)<0$ (definition of $\wedge$). Along the same lines, $v(\varphi)=0$ iff one of the conjuncts evaluates to 0 and the other to a nonnegative integer; the argument proceeds similarly, though both conjuncts must be kept track of. ☐

Note that for S_{2N}, which does not contain 0, the lemma guarantees that valuation of these connectives tracks PC valuation perfectly: negative to negative, positive to positive.

Dealing with the intensional connectives requires further development. Meyer postulates an auxiliary order that runs 0, −1, 1, −2, 2, −3, 3, … , *i.e.* one in which $a<b$ if $|a|<|b|$, but which reverts to the usual order between a and b when $|a|=|b|$, so that $-|x|<|x|$. In terms of this auxiliary order, v(P+Q) is defined as the maximum of {v(P),v(Q)}, paralleling disjunction in the ordinary order. But the auxiliary order is limited in its generality of application. For example, as Meyer notes, in this order v(P∘Q) is *not* the minimum of v(P) and v(Q). (E.g. $1\circ-2$ is -2, not 1.) Let us therefore approach the matter in a slightly different way, one that highlights other symmetries in the system and turns out to mesh more closely with ERC logic.

Let the *outermost* members of a set of integers to be those with the greatest absolute value; we can then state the definitions of the intensional connectives as follows:

(108) **Valuation of Intensional Connectives on Sugihara Sets**

(i) Fusion $v(P\circ Q)$ = Outermost of {v(P),v(Q)} if unique, else min[v(P),v(Q)].
(ii) Fission $v(P+Q)$ = Outermost of {v(P),v(Q)} if unique, else max[v(P),v(Q)].
(iii) Arrow $v(P\rightarrow Q)$ = Outermost of {v(¬P),v(Q)} if unique, else max[v(¬P),v(Q)].

The relation between the two clauses in each definition is essentially that of constraint domination in the optimality theoretic sense, and the resulting order is lexicographic, with the primary sort by absolute value and the subsidiary sort by value *tout simple*.

We can give a more compact definition if we recognize a function OUT[X], defined on sets of integers, that returns a set containing the outermost member(s) of X, *i.e.* $x\in$OUT[X] iff $|x|\geq|y|$ for all $y\in X$. For example, with X={−n,+n}, we have OUT[X]=X, but otherwise for X={m,n}, $|m|\neq|n|$, we have a singleton set. Then we can achieve the desired effects through composition with the functions min and max, which return the extremal element:

(109) **Valuation on Sugihara Sets (Fusion, Fission, Arrow)**

(i) Fusion $v(P\circ Q)$ = min OUT[v(P),v(Q)]
(ii) Fission $v(P+Q)$ = max OUT[v(P),v(Q)]
(iii) Arrow $v(P\rightarrow Q)$ = max OUT[v(¬P),v(Q)]

The intensional component, which distinguishes these connectives from their extensional congeners, lies entirely in the appeal to absolute value via OUT.

All these connectives are of course interdefinable, given negation. An important consequence is that $v(P\rightarrow Q)$ is designated iff $v(P) \leq v(Q)$.

(110) **Remark.** $v(P\rightarrow Q)\geq 0$ iff $v(P)\leq v(Q)$.

Pf. Left to reader.

A further consequence worth noting is that 0 is always an identity.

(111) **Remark.** 0 is an identity for fusion and fission in S_{2N+1} and S_Z .

Pf. 0 is never dominant. □

As a final observation along these lines, we note that the intensional connectives, including negation, evaluate to 0 iff all of their arguments are 0. (This is not true for the extensional connectives!)

(112) **Remark.** $v(P\circ Q) = v(P+Q) = v(P\rightarrow Q) = v(\neg P) = v(\neg Q) = 0$ iff $v(P) = v(Q) = 0$.

Pf. Direct from (111). □

We can now define a "Sugihara matrix" M_k to be the Sugihara set S_k equipped with the apparatus of valuation just outlined. This yields the definition of validity in M_k.

(113) **Def. Valid in M_k.** A sentence A is valid in M_k, written $M_k \Vdash A$, iff $v(A)\geq 0$ for all v on M_k.

Note that A is valid in M_Z iff A is valid in *all* M_k. More generally, we have from Dunn (1970:7, Theorem 4) the result that validity holds downward in the sequence of M_k's.

(114) **Validity Descent.** For M_n and M_k, with $n\geq k$, $M_n \Vdash A \Rightarrow M_k \Vdash A$.

Pf. The challenge is going from a 0-lacking M_n to 0-containing M_k: see Dunn.

Equally useful is the contrapositive, 'invalidity ascent': for S_k and S_n, with $k\leq n$, if A is *not* valid in M_k then it is *not* valid in M_n.

Meyer (1971;1975), reporting work from 1966, establishes the following key result.

(115) **Completeness , Soundness, and Decidability of RM.** A sentence of RM with N prop letters is provable from the axioms of RM iff it is valid in M_{2N} .

Pf. See AB:413, RM81.

This can be generalized mildly in the light of Validity Descent: an N-letter sentence of RM is provable iff it is valid in *some* M_k, for $k\geq 2N$.

From (115), the central properties of RM semantics follow immediately. In particular, M_Z and M_{Z*} provide a general semantic setting for RM logics of unbounded size.

(116) **Corollary.** Let RM be formulated with denumerably many prop letters. Then all and only the theorems of RM are valid in M_Z and M_{Z*}.

Pf. See AB:414, Corollary 3.5.

In addition, Dunn 1970 shows that each M_k is associated for finite k with a logic of its own: RM[k], which has the property that all and only its theorems are valid in M_k. When this property holds, one says that the matrix is 'characteristic' for the logic. Thus, M_Z and M_{Z*} are characteristic for RM.

(117) **Characterization.** Each M_k for finite k is characteristic for an extension of RM, *i.e.* RM with additional axiom(s). (Dunn 1970: 9, corollary 2).

These results establish that RM and its extensions have a remarkably elegant and simple structure. We have a hierarchy of logics and symmetry-breaking among the connectives, characterized by the Sugihara matrices:

M_1 with $S_1=\{0\}$: the peaceable kingdom in which everything is true, and in which all connectives are the same;[18]

M_2 with $S_2=\{-1,1\}$: classic Propositional Calculus, in which the extensional and their corresponding intensional connectives are identified;

M_3 with S3=$\{-1,0,1\}$: RM3, in which $\{\circ,\wedge\}$, $\{+,\vee\}$, and $\{\rightarrow,\supset\}$ all part company

⋮

$$\left.\begin{array}{ll} M_Z & \text{with } S_Z=\{\ldots-3,-2,-1,0,1,2,3,\ldots\} \\ M_{Z*} & \text{with } S_{Z*}=\{\ldots-3,-2,-1,1,2,3,\ldots\} \end{array}\right\rangle \text{RM}$$

To get a feel for the system, let us examine how these logics treat the sentence φ: $A\circ B\supset A$.

RM1 Valid, like everything else.

RM2 Valid, and coincident with $A\wedge B\supset A$, $A\wedge B\rightarrow A$, $A\circ B\rightarrow A$.

RM3 Valid, although $A\circ B\rightarrow A$ is not.

RM4 Invalid: because $(-1)\circ 2\supset(-1)$ yields $2\supset(-1)$ yields (-1).

⋮ Invalid: ditto, with more and more problems along the same lines.

RM Invalid: ditto.

Observe that the invalidity of φ is established in RM4, exactly as determined by Meyer's central result (115): 4 = 2 × number of prop letters in φ. And since all successor matrices contain this valuation pattern, invalidity persists from RM4 on up, as guaranteed by (114). Further, since the valuation possibilities of RM2 are a proper subset of those in RM3, the validity established at RM3 persists downward.

[18] In heaven, I suppose, lie down together / Agonized Pilate and the boa-constrictor / That swallows anything … – C. Day Lewis.

Meyer's (115) also makes it easy to nail down the relation between theoremhood in PC and theoremhood in RM, validating our claim above that RM includes PC.

(118) **Remark**. **RM/PC**. Let φ be a sentence using only connectives drawn from the set $\{\neg, \wedge, \vee, \supset\}$. Then φ is a theorem of RM iff φ is a theorem of PC.

> *Pf.* The LR direction is straightforward, since RM valuation under S_Z includes every assigment of $\{-1,+1\}$ to φ. Now suppose φ is a theorem of PC, and therefore assumes a positive value under any valuation over $S_2=\{-1,+1\}$. By the 2N property (115), φ is a theorem of RM iff $M_{2N} \Vdash \varphi$, where N is the number of prop letters in φ. By the extensionality lemma (107), the existence of a valuation v on M_{2N} with $v(\varphi)<0$ implies the existence of v* on S_2 with $v^*(\varphi)<0$. Since there can be no such v*, there is no such v, and $RM \Vdash \varphi$. □

7.3 Ordered Polyvaluation as RM Semantics

Establishing the relation between ERC logic and RM is now within reach. First we show how the notion of propositional variables representing ERCs over constraint hierarchies generates a Meyer/Sugihara semantics, via what we will call the *ordered polyvaluation* (OP).[19] Just as for VS, we assume that a prop letter represents an ERC by virtue of a polyvaluation $V=\prod v_i$ over a collection of ternary valuations $v_i:\Psi\rightarrow\{-1,0,1\}$, $1\leq i\leq N$, where Ψ is a set of prop letters.

To incorporate the notion of a constraint hierarchy, let the valuations v_i be subject to a *total order* R, essentially a permutation on the indices. Each v_i represents a constraint C_i, and R represents the domination order among the constraints when they are placed in a (totally ordered) hierarchy. For $C_i >> C_j$ let us have $v_i >_R v_j$, so that the greater element in the R order is the dominating constraint.

Given an ordered polyvaluation $\langle V,R\rangle$, consisting of a polyvaluation V with N individual valuations v_i and an order relation R on them, each $\psi\in\Psi$ can be assigned a numerical value drawn from the set of integers $\{-N,\ldots,N\}$. The choice of sign is determined by the distinction between T and F; the absolute value is determined by the position in the ranking order of the highest (first, greatest) constraint assigning a T or F value. This will allow us to re-construct the Sugihara-set semantics in its entirety.

To implement this progam, let us recall the notion of 'rank' for an ERC, from §3, (53), p. 27. The rank |C| of a constraint C in a hierarchy H is the stratum that contains it. This notion carries over straightforwardly to the setting of ordered polyvaluations, where each stratum contains but one element, due to the totality of R.

To measure position in the ordering R, we assign to each v_i a positive numerical weight, its *dominance*. The greatest v_i under R is assigned N, and we count downward until we reach the least v_i, which earns a value of 1.

[19] See Appendix 4 for a 'world' or Kripke-style semantics (Dunn 1976) that covers the same territory.

(119) **Def. Dominance.** Given an ordered polyvaluation $\langle V,R\rangle$, the *dominance* of any v_i in V, written dom(v_i), is the number of valuations v in V such that $v_k \geq v$ in the R order.

$$\text{dom}(v_i) = \text{card}\ \{v|\ v_k \geq v\}$$

Using dominance as a weight, we derive a numerical valuation from each v_i, which is the *value* that it assigns.

(120) **Def. Value.** Given $\langle V,R\rangle$, the *value* assigned by v_i to an argument ψ, written $\text{val}_i(\psi)$, is equal to $v_i(\psi)$ weighted by dom(v_i).

$$\text{val}_i(\psi) =_{df} \ v_i(\psi) \times \text{dom}(v_i)$$

For example, if V(ψ)= (e, T, e, F), with R following the coordinate order, then the corresponding sequence of values would be (0×**4**, 1×**3**, 0×**2**, −1×**1**) = (0,3,0,−1).

RM semantics is based entirely on the value associated with the highest-ranking polar valuation. Given the weighting system, this will be the one with the greatest *absolute value*, i.e. the *outermost* of the set of values, and it will be unique.

(121) **Def. Principal Value.** The *principal value* of ψ in an ordered polyvaluation $\langle V,R\rangle$, written $\mathsf{val}(\psi)$, is the value associated with the highest-ranked valuation in V assigning a polar value to ψ, and is equal to 0 if no valuation assigns a polar value.

$$\begin{aligned}\mathsf{val}(\psi) &=_{df} \min \text{O}\scriptstyle{\text{UT}}\displaystyle[\{\text{val}_i(\psi),\ 1\leq i\leq N\}] \\ &= \max \text{O}\scriptstyle{\text{UT}}\displaystyle[\{\text{val}_i(\psi),\ 1\leq i\leq N\}]\end{aligned}$$

Since there is only *one* outermost value in this case, the min and max functions are equivalent and merely serve to extract from the singleton set its one member. It turns out to be useful to express things this slightly artificial way, because it eases interaction with the intensional connectives.

Dominance assigns to each v_i a numerical value in the range $\langle 1...N\rangle$ and the v_i themselves assign values from $\{-1,0,+1\}$. The principal value of any ψ will then lie in the Sugihara set $S_{2N+1}=\{-N,...,+N\}$. A small-scale example should make this clear. Here are the results of all permutations of order on a 2-valuation system of polyvaluations.

(122)

	R_1 : 1>2						R_2: 2>1			
	v_1	val_1	v_2	val_2	**val**	**val**	v_2	val_2	v_1	val_1
i	**T**	2	**F**	−1	2	−2	**F**	−2	**T**	1
ii	***e***	0	**T**	1	1	2	**T**	2	***e***	0
iii	**T**	2	***e***	0	2	1	***e***	0	**T**	1
iv	***e***	0	***e***	0	0	0	***e***	0	***e***	0
v	**T**	2	**T**	1	2	2	**T**	2	**T**	1
vi	***e***	0	**F**	−1	−1	−2	**F**	−2	***e***	0
vii	**F**	−2	***e***	0	−2	−1	***e***	0	**F**	−1
viii	**F**	−2	**F**	−1	−2	−2	**F**	−2	**F**	−1

This simplest nontrivial case exemplifies various key properties:

1. A system of N constraints can assign 2N+1 values under a linear order R, and 2N if we exclude the degenerate vector $\delta = (0, \ldots, 0)$.

2. A given polyvaluation of α need not run through all possible assignments to α as we run through all permutations of the constituent valuations.

3. An ERC true in every model corresponds to a prop letter that assumes a designated (nonnegative) value under every permutation of order on the polyvaluation: here (ii)-(v), shaded. An ERC that has no models assumes only negative values, like those of (vi)-(viii). A nontrivial ERC assumes both positive and negative values, like (i) does.

Now that we have established a polyvaluation-based semantics for RM, we can inquire about the relation between validity in VS and valuation in RM. Because the theorems of VS are exactly those of S, as noted in remark (104), and because S is the implication-negation fragment of RM (Parks 1972), anything valid in VS is also valid in RM, and vice versa for sentences in the intensional vocabulary. This is unsurprising, and rather weak, in that validity requires truth in all valuations. A stronger result relates valuation in VS to valuation in RM: anything that is designated in VS under a given polyvaluation V is also true (i.e. designated) under RM rules in all orderings of that polyvaluation. This is a straightforward result – validity in VS means no F's in the polyvaluation, which ensures a nonnegative RM valuation regardless of order – nonetheless but has useful consequences. For example, if $V \Vdash \alpha \rightarrow \varphi$ in the VS sense, then $\langle V,R \rangle \Vdash \alpha \rightarrow \varphi$ for any R. This means that $\mathsf{val}(\alpha) \leq \mathsf{val}(\varphi)$, and then, taking the OP point of view, we are assured that for *any* ranking at all of which α holds, i.e. with $\mathsf{val}(\alpha) \geq 0$, the ERC φ has a rank that is above or co-stratal with that of α.

7.4 RM as the logic of OT

Two essential tasks remain before we can declare that we have successfully identified the logic of OT, as explored above in §§1-6, with RM.

[1] First, we determine the relation between the ERC format as in (4) ("some W must dominate all L's") and the semantics of ordered polyvaluation.

[2] Second, an important house-keeping task, showing that the various modes of valuation we have introduced all lead to the same conclusion. Specifically, we show the the result of polyvaluation in VS, which is defined over complex sentences built with intensional connectives, coheres with that obtained by the ordered polyvaluation semantics for RM.

[1] **The ERC and The Ordered Polyvaluation**

> … the next time he makes the aquinatance of the Ondt after this they have met themselves, these mouschical umsummables, it shall be motylucky if he will beheld not a world of differents. FW, 417.

The relation between ordered polyvaluations and constraint hierarchies is clear and straightforward. An expression ψ is designated or 'true' in a given $\langle V,R\rangle$ iff $\mathsf{val}(\psi)\geq 0$, *i.e.* iff its principal value is nonnegative. Designation does not occur when the principal valuation v_ψ, highest ranked under R, serves up a negative value on ψ. This is perfectly parallel to evaluation in a constraint hierarchy, where an ERC fails iff its highest-ranked polar assessment is L.

An ERC expresses a limitation on contraint hierachies. What limitation does a polyvaluation $V(\psi)$ impose on the orders R under which ψ is designated? Given an ordererd polyvaluation $\langle V,R\rangle$, designation of ψ,with $\mathsf{val}(\psi)\geq 0$, can happen only under these conditions:

- If there is any v_j in V with $v_j(\psi)$=F/−1, then there must be a v_i in V with $v_i(\psi)$=T/+1, such that dom(v_i)>dom(v_j) under R.

This is of course nothing more than a restatement of (4), the definition of an elementary ranking condition. Thus, the ordered-polyvaluation semantics of RM contains ERC theory within it.

To make the relationship more precise, let us specify the function h mapping from polyvaluated expressions to PC ERC expressions of type (4). Quantification is over the valuation functions in V.

(123) $\mathsf{h}: (\psi,V) \mapsto \exists g\forall f\,[\,(f(\psi)=F) \supset (g(\psi)=T \;\wedge\; dom(g)>dom(f))\,]$

The function h associates a polyvaluated expression with a statement of the conditions required for it to be true under an ordering of its polyvaluation. Since this has exactly the form of an ERC, we identify {V} with the set of constraints that track it, as well as T with W, L with F. The following relation then holds:

(124) **Remark**. If a certain ψ is designated by RM rules under an ordered polyvaluation $\langle V,R\rangle$, then $\mathsf{h}(\psi,V)$ holds of R. Conversely, if $\mathsf{h}(\psi,V)$ holds of R, then ψ is designated under $\langle V,R\rangle$.

In this way, the truth (or falsity) of a logical expression under an OP exactly tracks the status of an ERC over a constraint hierarchy.

This a mere first step to ERC logic. The real strength of the approach is that the *sentences* of RM, of arbitrary logical complexity, translate under polyvaluation to *sentences* of ordinary logic about constraint rankings.

PC logic does not, of course, contain connectives like '$\circ$' and '$\rightarrow$' or '$\neg_{RM3}$' (i.e. negation with the possibility of $\neg\alpha=\alpha$). But the interest of RM lies precisely in its intensional connectives and the way that it handles their relations to the familiar extensional system. In the present context, it is crucial that the set of ERCs is closed under the intensional operations: ERCs in, ERCs out. This is a direct correlate, via h, of the fact that VS polyvaluation applies to entire intensional expressions as well as to the prop letters they are made up of, by virtue of definition (102). Because any ERC vector or polyvaluation corresponds to an ordinary PC logical expression, all taint of intensionality can be *eliminated*, by replacing intensional subformulae with single polyvaluated expressions, when the time comes to speak only of truth and falsity. In short, $\alpha\circ\beta$ looks just a single prop letter to the interpretive mechanism.

The eliminative claim holds not just for VS but for the full range of RM sentences, employing any mixture of connectives. Because RM licenses distribution of $\{\wedge,\vee\}$ over $\{\circ,+\}$ and *vice versa*, with De Morgan laws relating the connectives within each set, we have a normal form in which the negative takes scope only over single prop letters, and in which no extensional connective is in the scope of any intensional connective (AB:396-399). The purely intensional subformulae are polyvaluable expressions that map to single ERCs, and the extensional remainder may be interpreted exactly as in standard logic, by virtue of RM/PC (118).

VS supplies the theory of intensional connectives under polyvaluation, but the desired application is to RM, with a different overall semantics based on order. To complete the argument, we need to certify that fusion and negation as defined in VS function properly in the ordered polyvaluation semantics of RM.

[2] **VS and RM**

For each *intensional* connective, including negation, there are now two distinct routes to calculating its value: one depends only on the Meyer-type rules for RM (109); the other makes us of the S recipe for evaluating intensional expressions (102) , which underlies VS. We must make sure that these two routes give the same result.

To completely specify the OP semantics for RM, we extend the function val to arbitrarily complex formulae, in the manner of (106) and (109)

(125) **General Valuation in OP semantics**

(i) Negation $\mathrm{val}(\neg\alpha) = -\mathrm{val}(\alpha)$
(ii) And $\mathrm{val}(\alpha\wedge\beta) = \min\{\mathrm{val}(\alpha),\mathrm{val}(\beta)\}$
(iii) Fusion $\mathrm{val}(\alpha\circ\beta) = \min \textsc{Out}\{\mathrm{val}(\alpha),\mathrm{val}(\beta)\}$

The function val itself is defined on the constituent valuations in V and their order R, as follows, repeated from (120) and (121). The definition assumes N valuations v_i in V.

(126) **Definition of val**

(i) $\text{val}_i(\alpha) =_{df} v_i(\alpha) \times \text{dom}(v_i)$

(ii) $\mathsf{val}(\alpha) =_{df} \min \text{OUT} \{\text{val}_i(\alpha), 1 \le i \le N\}$

The ambiguity in valuation paths comes about because the constituent valuations v_i are themselves empowered to handle not just atomic prop letters but general intensional expressions. This is what gives us VS and the entire theory of ERC logic explored in §§1-6.

(127) **S definition of v**

(i) $v(\neg\alpha) = -v(\alpha)$

(ii) $v(\alpha\circ\beta) = \min \text{OUT } \{v(\alpha), v(\beta)\}$

Under a given $\langle V,R \rangle$ within RM, we may calculate, for example, $\text{val}(\alpha\circ\beta)$ from $\mathsf{val}(\alpha)$ and $\mathsf{val}(\beta)$, according to the Sugiharan RM recipe for fusion as a function of its arguments.

$$\mathsf{val}(\alpha\circ\beta) = \min \text{OUT}\{\mathbf{val}(\alpha), \mathbf{val}(\beta)\} \qquad [(125)\ (iii)]$$

But we may also circumvent the recursive val-to-val recipe, using S/VS to do fusion coordinatewise, and only then evaluate the result under R, ultimately recursing on v_i.

$$\mathsf{val}(\alpha\circ\beta) = \min \text{OUT } \{\mathbf{val_i}(\alpha\circ\beta), 1 \le i \le N\} \qquad [(126)\ (ii)]$$

$$= \min \text{OUT } \{\mathbf{v_i}(\alpha\circ\beta) \times \text{dom}(v_i), 1 \le i \le N\} \qquad [(126)\ (i)]$$

Well-definition requires that these two routes yield the same answer.

$$\min \text{OUT}\{\mathsf{val}(\alpha), \mathsf{val}(\beta)\} \stackrel{?}{=} \min \text{OUT } \{\text{val}_i(\alpha\circ\beta), 1 \le i \le N\}?$$

The following illustrative examples give a sense of the different valuation paths.

(128) **Fusion both ways**

	$\mathbf{v_1}$	val_1	$\mathbf{v_2}$	val_2	$\mathbf{v_3}$	val_3	**val**
α	**T**	3	**F**	−2	**T**	1	3
β		0	**T**	2	**F**	−1	2
α∘β	**T**	3	**F**	−2	**F**	−1	3

We may obtain the value of $\alpha\circ\beta$ from the last column by computing with $\mathsf{val}(\alpha)$ and $\mathsf{val}(\beta)$, whereby the outermost of {3,2} is 3. Or we may first establish the bottom-row polyvaluation of $\alpha\circ\beta$ as $V(\alpha\circ\beta)$, and apply val to the resulting expression, selecting 3 as the outermost of {3,−2,−1}, arriving at the same answer.

Negation shows a similar behavior.

(129) **Negation both ways**

	$\mathbf{v_1}$	val_1	$\mathbf{v_2}$	val_2	$\mathbf{v_3}$	val_3	**val**
α	**T**	3	**T**	2	**F**	−1	3
$\neg\alpha$		−3	**F**	−2	**T**	1	−3

We may compute the value of $\neg\alpha$ from val(α) in the last column by negating it directly; or we may negate each component valuation of α to arrive at the bottom-row VS representation of $\neg\alpha$, whose outermost value will be the same.

Why does this work? The negation case is clear enough – the principal valuation (here v_1) is determined by absolute value, which remains the same under multiplication by −1. And since the coordinatewise definition of negation, *via* v_i, is the same (arithmetic negation) as the val definition, we are guaranteed sameness of result. In essence, the 'outermost' function for choosing val(α) commutes with the function that interprets the '$\neg$' connective:

$$-\ [\text{outermost of}\{val_i(\alpha)\}] = \text{outermost of } \{-val_i(\alpha)\}$$

The success of the fusional calculation turns, somewhat more subtly, on the same considerations. Perhaps this can be seen most clearly if we construct the order implicit in the (composite) function 'minimal outermost'. Fusion is minimal in the following linear order on Z:

$$\text{MinO order:}\quad \ldots < -3 < 3 < -2 < 2 < -1 < 1 < 0.$$

Simple properties of the general min function on linear orders then give us the desired result. The min function is idempotent: min(min X) = min X. It also has a kind of associative property: given a set of integers X and Y, we may directly compute min[$X \cup Y$] or we may first compute min[X] and min[Y] and only then compute the global minimum min[{min[X], min[Y]} – the outcome must be the same either way, by transitivity of the order. Generally, we can divide up a set of integers in any ways we like, take min over the pieces, then min over those piece-wise minima, always arriving at the same result, which is the minimum for the whole.

In example (128), it matters not whether we take min OUT row-wise first, computing val(α) from the $val_i(\alpha)$ and val(β) from the $val_i(\beta)$, and then apply min OUT to the result; or go columnwise first, computing each $val_i(\alpha\circ\beta)$ from $\{val_i(\alpha), val_i(\beta)\}$ by min OUT, and then obtain the min OUT of the resulting set of pairwise minima. In either case, we are ultimately computing min OUT over the entire collection of $val_i(\alpha)$'s and $val_i(\beta)$'s.

Our desired conclusion, then, is that the val function for any $\langle V,R\rangle$ yields the same results under two conditions, given a complex expression φ with proper subparts φ_i.

1) Recurse on val. val(φ) is determined from the val(φ_i) by OP rules (125).
2) Recurse on v_i. val(φ) is determined from the $val_i(\varphi)$, $1 \le i \le N$, where the val_i call directly on v_i and therefore on $v_i(\varphi)$ and S rules (127).

To deconstruct this distinction, let us separate out the part of val that recurses only at the v_i level, which we will name 'VAL'. We then show that any derivation using val on an intensional expression achieves the same value as one using only VAL. This shows that all possible derivations involving val, with any combination of recursions, will necessarily agree.

In the interest of conciseness, we use the following abbreviations: we write MO for 'min OUT', and whenever a subscripted item appears within braces, like $\{x_i\}$ or $\{f(x_i)\}$, we mean the entire set of things obtained by assigning the subscript all of its values.

The theory of valuation then comes out like this, using the concise notation where appropriate. The definitions are tagged with names in the left-hand column.

$\circ$/val $\quad \mathsf{val}(\alpha_1\circ\alpha_2) =_{df} \text{MO}\{\mathsf{val}(\alpha_k)\}$
$\neg$/val $\quad \mathsf{val}(\neg\alpha) =_{df} -\mathsf{val}(\alpha)$

These definitions recurse on val. The function val may also be cashed in for the v_i indirectly by this, when α is free of extensional connectives:

val/VAL $\quad \mathsf{val}(\alpha) = \text{VAL}(\alpha)$

This simply provides a name, 'VAL', for the part of the derivation that depends entirely on the v_i to compute the value of intensional collocations. The function VAL itself knows nothing of connectives, and is defined as follows:

Def. VAL $\quad \text{VAL}(\alpha) =_{df} \text{MO}\{\text{val}_i(\alpha)\}$

This establishes VAL as precisely that subpart of valuation that plucks out the dominant value from a single polyvaluated expression. The val_i provide the call to the valuations in V as well as to the order R. Their definition is recalled here:

Def. val_i $\quad \text{val}_i(\alpha) =_{df} v_i(\alpha)\times\text{dom}(v_i)$

And the v_i make use of S definitions of the connectives, paralleling those for val:

$\circ$/v $\quad v(\alpha_1\circ\alpha_2) =_{df} \text{MO}\{v(\alpha_k)\}$
$\neg$/v $\quad v(\neg\alpha) =_{df} -v(\alpha)$

We will first show that a derivation involving just *one* recursive call via the rules $\circ$/val or $\neg$/val will give the same result as a direct computation from VAL. This will then allow us to bootstrap our way to the desired general conclusion.

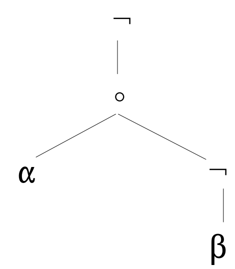
(130) **Lemma. One-Step Well-Definition.** Let α_1, α_2 be expressions of S. For any OP $\langle V,R\rangle$, if $\mathsf{val}(\alpha_1\circ\alpha_2)$ and $\mathsf{val}(\neg\alpha)$ are each computed with a single recursive call to val, then the resulting valuation is identical to that obtained from VAL, without such calls.

Pf.

$\mathsf{val}(\alpha_1\circ\alpha_2)$	$= \mathbf{MO}\ \{\mathbf{val}(\alpha_k)\}$	$\circ$/val
	$= \text{MO}\{\text{VAL}(\alpha_k)\}$	val/VAL
	$= \text{MO}\{\text{MO}\{\text{val}_i(\alpha_1)\}, \text{MO}\{\text{val}_i(\alpha_2)\}\}$	Def. VAL
	$= \text{MO}\{\text{MO}\{\text{val}_i(\alpha_k)\}\}$	Assoc. MO
	$= \text{MO}\{\text{val}_i(\alpha_k)\}$	Idemp. MO
	$= \text{MO}\{v_i(\alpha_k)\times\text{dom}(v_i)\}$	Def. val_i
$\text{VAL}(\alpha_1\circ\alpha_2)$	$= \mathbf{MO}\ \{\mathbf{val_i}(\alpha_1\circ\alpha_2)\}$	Def. VAL
	$= \text{MO}\{v_i(\alpha_1\circ\alpha_2)\times\text{dom}(v_i)\}$	Def. val_i
	$= \text{MO}\{\text{MO}\{v_i(\alpha_k)\times\text{dom}(v_i)\}\}$	$\circ/v_i$
	$= \text{MO}\{v_i(\alpha_k)\times\text{dom}(v_i)\}$	Idemp. MO

$$
\begin{array}{llll}
\mathsf{val}(\neg\alpha) & = -\mathbf{val}(\boldsymbol{\alpha}) & & \neg/\mathsf{val} \\
 & = -\mathrm{VAL}(\alpha) & & \mathsf{val}/\mathrm{VAL} \\
 & = -\mathrm{MO}\ \{\ \mathrm{val}_i(\alpha)\ \} & & \text{Def. VAL} \\
 & = -\mathrm{MO}\ \{\ v_i(\alpha)\times \mathrm{dom}(v_i)\ \} & & \text{Def. } \mathrm{val}_i \\
 & = \mathrm{MO}\ \{\ -v_i(\alpha)\times \mathrm{dom}(v_i)\ \} & & \text{Comm. MO/}- \\
 & & & \\
\mathrm{VAL}(\neg\alpha) & = \mathbf{MO\ \{\ v_i(\neg\alpha)\times dom(v_i)\ \}} & & \text{Def.VAL} \\
 & = \mathrm{MO}\ \{\ -v_i(\alpha)\times \mathrm{dom}(v_i)\ \} & & \neg/\ v_i \qquad \square
\end{array}
$$

Now consider the general character of valuations involving val. These track the syntactic structure of a complex expression, which we may represent as an at-most binarily branching tree, with each nonterminal node labeled as '∘' or ' ¬'. For example,



$= \quad \neg\,(\alpha\circ\neg\beta)$

Given any course of valuation for the whole expression, we may represent it by labeling each node in the tree with the functions that are responsible for its valuation. The key nodes will be those where valuation is handed over from 'val' to 'VAL' by the val/VAL rule; they will then be labeled with both function names. Any node labeled with 'val' may have its daughters labeled with either 'val' or 'val/VAL', depending on the evaluative route taken. The key fact is that every path from the root to a terminal node must contain a 'VAL' at some point, because it is only through 'VAL' that contact is made with the polyvaluation. Furthermore, no more than the one 'VAL' will appear in each path, since 'VAL' itself is not recursive. Nor may a node labeled 'VAL' dominate any labeled 'val'. This means that every such valuation tree, given a val-recursive derivation, has a 'frontier' of nodes labeled 'val/VAL' that separates the (upper) val-recursive parts from the (lower) parts where v_i is at work.

The lemma allows us to push the frontier back (or 'up'0. The nodes on the frontier come in two types: those exhaustively dominated by a 'val' node; those that are paired with another frontier node to form a constituent dominated by a 'val' node. In either case, the lemma tells us that the dominating mother node can be replaced by a 'VAL' node, retaining the same valuation. This procedure creates a new frontier, which can then be pushed back in exactly the same way. Because trees are finite, this procedure will eventually replace every 'val' with 'VAL', never changing the valuation. It follows that any recursive derivation will compute same overall valuation as would be computed by one that started right off with VAL. This establishes that all derivations give the same answer, as desired.

(131) **Proposition 7.1**. **Well-definition.** Let φ be a sentence of S, *i.e.* one containing connectives from the set {¬,→,∘,+}. All courses of valuation using recursion of val (125) and/or recursion of the v_i (127) yield the same result.

Pf. By formalizing the textual narrative. □

It may also be observed that the VS definition of fusion and negation is the only coordinatewise one that will fit with RM semantics.[20]

To complete all the house-keeping with respect to the VS/OP relation, we want to show that arbitrary sentences of RM come out the same over any course of valuation, *i.e.* no matter if or when or to what extent VS-collapse is done.

(132) **Corollary to Proposition 7.1.** Let φ be an arbitrary sentence of RM, using any mix of intensional and extensional connectives. For any OP ⟨V,R⟩, let $[\mathrm{val}]_1(\varphi)$ and $[\mathrm{val}]_2(\varphi)$ represent two distinct courses of valuation. Then $[\mathrm{val}]_1(\varphi)$ and $[\mathrm{val}]_2(\varphi)$ deliver the same valuation for φ.

Pf. Because VAL only works over expressions lacking extensional connectives, the only differences between any such $[\mathrm{val}]_k(\varphi)$ lie in the valuation of intensional subformulae. But these are evaluated equivalently, by Prop. 7.1. And since the value of the whole is determined by the values assigned to the parts, which do not change under changes in course of valuation, the result is guaranteed. □

The import of these results is that we can freely translate *any* RM formula, under polyvaluation, for a purely PC expression involving ERCs, obtained by using the correspondence function h. From the normal form result (AB:399), we know that any RM sentence is equivalent to one in which intensional expressions are segregated off from the extensional structure, in the sense that no intensional connective contains an extensional connective in its scope. Any such intensional subformula corresponds to a single ERC, by VS rules. Thus, given a polyvaluation, we may de-intensionalize any RM sentence by reducing it to normal form and replacing every intensional subformula with a prop letter that has the same polyvaluation. We are guaranteed by Proposition 7.1 that the resulting sanitized wff has the same valuation as the original, under any R. And the resulting wff has, by RM/PC (118), the same semantic status it has in straight extensional logic.

[20] It might be worth noticing that the matter can be treated syntactically as well, possibly with profit. Let us reconstruct a polyvaluated expression as a fusion of elementary components. For any α, let α_i be a prop letter whose polyvaluation agrees with α on v_i but evaluates to 0 everywhere else; then α is logically equivalent, under V, to a fusion over the α_i's. We have $V\Vdash \alpha = f_k\alpha_k$.(In essence, we identify a 'basis' for the polyvaluation.) From this we obtain $V\Vdash \alpha\circ\beta = f_k\alpha_k \circ f_k\beta_k = f_k(\alpha_k\circ\beta_k)$, the last step by associativity and commutativity of fusion. But $f_k\alpha_k\circ\beta_k$ is nothing more than a rewrite of the definition of V(α∘β), since it performs fusion in each coordinate. To deal with negation, note along the same lines that $V\Vdash \neg f_k\alpha_k = f_k(\neg\alpha_k)$, where f is self-dual here because of the mutual 'orthogonality' of the α_k's. (Since no fusion ever involves two polar values, we have α∘β=α+β.)

We conclude with the observation that we *cannot* extend VS-style polyvaluation to the extensional connectives ($\wedge$, $\vee$, $\supset$). Their properties are determined only within OP, where order is crucial. To see why such extension is impossible, consider the fate of RM3 conjunction under the coordinatewise logic of VS. The special symbol $\vec{\sqcap}$ emphasizes the distinctness of coordinatewise pseudo-conjunction from authentic $\wedge$.

(133) **Pseudo-Conjunction $\vec{\sqcap}$ $\neq$ Conjunction $\wedge$**

Coordinatewise $\vec{\sqcap}$	**Value under R = $v_1 > v_2 > v_3$**	
$V(\alpha) = (T,F,e)$	$val(\alpha)=3$	$v_\alpha = v_1$
$V(\beta) = (e, T,F)$	$val(\beta)=2$	$v_\beta = v_2$
• $V(\alpha \vec{\sqcap} \beta) = (T \wedge e, F \wedge T, e \wedge F) = (e,F,F)$	• $val(\alpha \vec{\sqcap} \beta) = -2 \notin \Delta$	
	• $val(\alpha \wedge \beta) = 2 \in \Delta$	

Were $\vec{\sqcap}$ and $\wedge$ to be conflated, the result would be classic ill-definition. In the example, two OP-designated expressions pseudo-conjoin to a falsity, an impossibility for conjunction under OP rules.

This negative result emphasizes the necessity of establishing, as we have just done, that the same thing never happens with fusion or negation (and thus with any of the intensional connectives). In this regard, recall the observation in §2, p.9, that fusion was not 'truth functional', in the sense that even a univerally true ERC and a universally false ERC could fuse to an ERC true under some rankings. Thus, taking (W,*e*) as an example of the true and (*e*,L) as an example the false, we have (W,*e*)$\circ$(*e*,L)=(W,L). The apparent oddity is that simple fusion gives T$\circ$F=F, which looks unimpeachably truth functional. Current understanding resolves the apparent conundrum: the mistake lies in identifying (W,*e*) with 'T', (*e*,L) with 'F'; a more nuanced appreciation is required.

(134) **Nontruthfunctionality of Fusion Ranked Away**

$V(\alpha) = (T,e)$	$val(\alpha) = 2$	(nb. R: $v_1 > v_2$)
$V(\beta) = (e,F)$	$val(\beta) = -1$	
• $V(\alpha \circ \beta) = (T,F)$	• $val(\alpha \circ \beta) = \min \mathrm{OUT}\{2, -1\} = 2$	

From the fusional point of view, the fact that $val(\beta)$ is negative ('false') matters not at all, in the face of the greater dominance (rank, absolute value, position of principal valuation in the order) carried by α. For conjunction, by contrast, where the intensional notion of *dominance* is not at play, negativity is all. OP finds no difficulties for fusion in the global fact that truth and falsity can combine to yield truth, because at the key local juncture (the decisive constituent valuation, α's principal valuation, $v_\alpha = v_1$) the expression β is in fact (locally) designated and not false at all, because $v_\alpha(\beta)=v_1(\beta)=0$. This is, of course, just the story of constraint ranking, retold.

7.5 From RM to PC

The discussion above leads to the conclusion that RM is the logic of OT, in the sense that any statement about ERCs made in PC is true iff the corresponding statement in RM is designated under the relevant orderered polyvaluation.

There is a minor wrinkle to be smoothed over: not every sentence of RM has a correspondent sentence about ERCs, because the function h only deals in the language of S. Thus an expression like $(A\wedge B)\circ(C\vee\neg D)$ doesn't translate directly. We know, though, that any such sentence is, by distribution, strictly equivalent to one in which all conjunction and disjunction are outside the scope of fusion,which can be read via h as a statement about conjunctions and disjunctions of four ERCs: $(A\circ C)$, $(A\circ\neg D)$, $(B\circ C)$, $(B\circ\neg D)$. Let us call 'normal' any sentence in which extensional connectives $\wedge$, $\vee$, $\supset$ are not in the scope of fusion, fission, or arrow. We lose nothing in focusing on such normal sentences, and may then state the RM/OT relation as follows:

(135) **Proposition 7.2**. Let φ be a normal sentence of RM subject to evaluation under a given $\langle V,R\rangle$. Let H be a constraint hierarchy in which each v_i in V corresponds to a constraint C_i which assigns W wherever v_i assigns T, assigns L wherever v_i assigns F, and asssigns *e* wherever v_i assigns *e*. Let the constraints in H be ranked according to the sequence imposed on their corresponding valuations by R. Let φ' be the sentence obtained by using the correspondence function h to replace all prop letters and intensional subformulae of φ with ERCs.

Then φ is designated under $\langle V,R\rangle$ iff φ' holds of H.

Pf. Suppose first that φ contains only intensional connectives and negation. Then φ' is just a single ERC. We know from remark (124), p. 62, that the thesis holds, because the ERC φ' is exactly an expression of the conditions under which φ is true under $\langle V,R\rangle$.

Now suppose that φ contains extensional connectives. Let φ^* be an expression in which each maximal intensional subexpression ψ_i of φ is replaced by a prop letter P_i not in φ; let V^* be a polyvaluation exactly like V over the prop letters in φ, and which assigns to each P_i exactly the valuation $V(\psi_i)$. Clearly, $\langle V^*,R\rangle \Vdash \varphi^*$ iff $\langle V,R\rangle \Vdash \varphi$. Now consider the PC valuation of φ' wrt H. Each maximal intensional subexpression ψ_i of φ corresponds to an ERC $h(\psi_i)$, which is either true or false of H, accordingly as $\langle V,R\rangle \Vdash \psi_i$ or not. Now construct a PC valuation v of φ^* which assigns T to P_i if $\langle V,R\rangle \Vdash \psi_i$ and F to P_i otherwise; it follows that $v(\varphi^*)$ is T iff φ' holds of H. Suppose that $\langle V^*,R\rangle \Vdash \varphi^*$. We have it from the extensionality lemma (107), p. 55, that $v(\varphi^*)=T$, and conversely, if $v(\varphi^*)=T$, $\langle V^*,R\rangle \Vdash \varphi^*$. Therefore, if $\langle V,R\rangle \Vdash \varphi$, then φ' holds of H, and conversely. □

It follows from Proposition 7.2 that any *theorem* of RM, stated in terms of normal sentences, will cash out as a true statement about logical relations between ERCs, holding generally over all hierarchies. A number of the basic results established above may now be recognized as reflexes of RM theorems, rather than peculiarities of OT. Consider the relations between connectives:

i.	$A\wedge B \rightarrow A\circ B$	(R55, AB:397)	Lemma (14) p.10
ii.	$A\circ B \rightarrow A+B$	(RM71, AB:397)	
iii.	$A+B \rightarrow A\vee B$	(R54, AB:396)	
iv.	$A\circ B \rightarrow A\vee B$	(ii) & (iii)	Proposition 2.2, p. 11

Notice too that an RM argument for the truth of assertions like these, based on the Meyer/ Sugihara semantics, is likely to be rather direct. In the case of the all-important conjunction-fusion relation (i), for example, we need merely note that $val(A\wedge B)=min[val(A),(B)] \leq val(A\circ B)$.

Transition from RM formulae to those of propositional/predicate calculus is straightforward Consider expression (i), which contains the non-PC connectives $\rightarrow, \circ$. Fusion is eliminated by taking $A \circ B$ to be equivalent to a single propositional variable denoting $\mathsf{h}(A \circ B)$, which is just another ERC. Of the arrow, we know from RM that $\Vdash (P \rightarrow Q) \rightarrow (P \supset Q)$, therefore we are licensed to conclude by modus ponens, which works on '$\rightarrow$' in RM, the following:

i.′ $\quad A \wedge B \supset [A \circ B]$.

This expression is intension-free, so we have the PC result

i.′′ $\quad \vdash A \wedge B \supset [A \circ B]$

which can also be written equivalently as $A \wedge B \vdash A \circ B$ or $\{A,B\} \vdash A \circ B$, as we have throughout.

Finally, we note that the relations involved here give an instructive view of fusion. "How then to interpret $\circ$? We confess puzzlement," write Anderson & Belnap *ca.* 1975 [AB]. By Restall 2000, various clear interpretations are known: "In the world of function typing, an object is of type $A \circ B$ just when it can be obtained by applying something of type A to something of type B. A string is of type $A \circ B$ just when it is made up of a string of type A concatenated with a string of type B (p.28-29)." Nevertheless, it remains somewhat obscure what fusion is when put to work in the home territory of logic, combination of propositions. For example, although one grasps the meaning of "The river is wide $\wedge$ the river is deep", it is murkier what "the river is wide $\circ$ the river is deep" commits one to, and perhaps unclear that it has any sense at all. However, in the case of the kind of order relations such as we have been dealing with here, we get substantial patches of clarity. For example,

$$(a>b)\circ(c>d) = \quad (a>b \wedge a>d) \vee (c>b \wedge c>d)$$

faithfully translates the fusion relation for nontrivial ERCs. The system does not allow us to interpret '$-$' as '$\neg$' though: $-(a>b) =_{df} (b>a)$ but $\neg(a>b)=(b \geq a)$, and this disparity leads to ruination with expression like b>b. So $(b>b) = -(b>b)$, but we need b>b to be false. And we cannot simply ban expressions like b>b from our calculus, since they will arise from perfectly ordinary expressions via fusion conjunction. The full-scale ERC of (4) of course supports a definition of fusion in terms of familiar logical operators, and negation works there for everything but degenerate ERCs.

7.6 Systems of Ordered Polyvaluations

Any theorem of RM is guaranteed to hold in all ordered polyvaluations. Optimality Theory, however, focuses on a collective structure that is intermediate between the single ordered polyvaluation and the entire set of them: the system of all rankings on a set of constraints, which is equivalent to a single polyvaluation V considered under every ordering of its constituent valuations.

OP semantics for RM depends on both V and R, but makes rather light use of V: although there may be many consituent v_i in V, and many assigning polar values, only *the greatest of these* plays a role in any given case; the rest are silent. Optimality Theory, by contrast, is deeply interested in each v_i, because its force – which may be occulted under a given R – must becomes apparent at some point in the run through all the possible orders. Within OP semantics for RM, let us define a *system* as the collection of all orderings on a fixed polyvaluation; we will argue that the *system* has a natural place in RM semantics as well as in linguistic theory.

(136) **Def. System.** An OT System (OTS), notated $\Sigma(V)$, is a set of ordered polyvaluations $\langle V,R_k\rangle$ sharing the same V, and including all permutations R_k on the indices i of the v_i in V.

It is only within the *system* that the major characteristics of ERC logic emerge. An ERC, after all, corresponds to a polyvaluated expression $V(\alpha)$ whose conditions for truth are of interest within the system $\Sigma(V)$.

The notion of 'entailment' used throughout is entirely system-based. When we say for ERCs α,β that $\alpha\vdash\beta$, we don't mean that entailment holds theoremwise and generally, for any V and any R; we mean that it holds precisely inside the one system $\Sigma(V)$, where V refers to the constraint system of interest. Our logical focus is then on expressions of the form $\Sigma(V)\Vdash\varphi$, 'some RM sentence φ holds throughout any given system $\Sigma(V)$'.

With this understanding, we can further explicate the notation used in the §§1-6 above. When we write $\alpha\vdash\beta$, we mean that the PC expression corresponding to α entails, in the PC sense, the expression corresponding to β; equivalently, by the deduction theorem, we have $\vdash\alpha\supset\beta$. Construing α,β as sentential variables within RM, we mean that for a specific polyvaluation V, we have $\Sigma(V)\Vdash\alpha\supset\beta$. The expression '$\alpha\supset\beta$' is strictly weaker in RM than '$\alpha\rightarrow\beta$', but it corresponds perfectly to the PC notion, and can be immediately cashed in for it, so long as we are sure that α and β are free of intensional connectives.

If φ is a *theorem* of S or of RM, then $\Sigma(V)\Vdash\varphi$ holds *a fortiori* for any V, because $\langle V,R\rangle\Vdash\varphi$ holds for any V and any R, and $\Sigma(V)$ is just a collection of $\langle V,R\rangle$'s. The converse is clearly not true: if φ holds in some $\langle V,R\rangle$, there is no guarantee whatever that it holds in any other. The situation is similar for collections of ordered polyvaluations: merely holding in various $\langle V_i,R_k\rangle$ for various V_i and R_k guarantees nothing of general interest. But certain collections of ordered polyvaluations – *systems* – have structural properties that make them rather like down-sized versions of the entire model space. Let us consider three examples.

[A] **VS and RM**

The relation between VS and RM is significantly strengthened at the level of the system. We know from Parks 1972 that a sentence φ, with intensional connectives only, is a theorem of S iff it is a theorem of RM.

The situation is less salutary for single polyvaluations: we only have one direction of implication, from VS to RM: if $V\Vdash\varphi$ under VS rules, so that every coordinate is designated, then clearly $\langle V,R\rangle\Vdash\varphi$ for any R. But the converse can scarcely be expected: the insensitivity of OP semantics to valuations beyond the principal one means that $\langle V,R\rangle\Vdash\varphi$ need barely penetrate the polyvaluation. For example, we have

$$\langle V,R\rangle\Vdash(e,T)\rightarrow(T,F) \qquad \text{(Sugihara: } 1\leq 2)$$

taking R to be the same as the coordinate order, but we don't have coordinatewise VS success, because the second coordinate fails:

$$(e\rightarrow T,\ T\rightarrow F) = (T,F) \notin \Delta\times\Delta$$

Similarly for random collections of polyvaluations. Within an entire system $\Sigma(V)$, however, good order re-emerges, and the converse sails through. The fact that φ must hold under every R means that all coordinatewise relations will be searched out.

(137) **Remark**. For any given polyvaluation V and expression φ of S, $V \Vdash \varphi$ iff $\Sigma(V) \Vdash \varphi$.

Pf. LR. Since $v_i(\varphi)>=0$, $val(\varphi)\geq 0$ under every R. For the RL direction, we need only note that every v_i in V is greatest (initial) under some R, and therefore that $v_k(\varphi)=F$ cannot be tolerated for any k, so that every coordinate must be designated. □

This means that any system behaves like the entire space of RM valuations with respect to this property.

[B] **Fusion and Conjunction.**

It is within the *system* that fusion emerges as a worthy rival to conjunction. The main result of §2 finds that entailments from a conjunction of ERCs lead to entailments from a fusion of ERCs. We repeat its statement here for convenience.

(138) (Proposition 2.5) Let $\mathcal{A}$ be a set of ERCs. $\mathcal{A}\vdash\varphi$ iff there is a $\Psi\subseteq\mathcal{A}$ such that $f\Psi\vdash\varphi$.

Interpreted in the present context, this identifies circumstances under which the truth of $\wedge\mathcal{A}\supset\varphi$ guarantees that there is a $\Psi\subseteq\mathcal{A}$ with $f\Psi\supset\varphi$. This is true rather trivially for a single RM valuation, when $\mathcal{A}$ is any set of RM sentences at all. It is certainly not guaranteed for a random collection of RM valuations, but – by Proposition 2.5 – it is true for all OT systems, with $\mathcal{A}$ restricted to ERCs.

(139) **Remark.** Let $\mathcal{A}$ be a set of sentences in the language of RM. Let val be the RM valuation under a single given OP $\langle V,R\rangle$. Then $val \Vdash \wedge\mathcal{A}\supset\varphi$ iff there is a $\Psi\subseteq\mathcal{A}$ such that $val \Vdash f\Psi\supset\varphi$.

Pf. RL is trivial because $RM \Vdash \wedge\Psi\rightarrow f\Psi$, for any Ψ. For the LR direction, note that $val(\wedge\mathcal{A}\supset\varphi)\geq 0$ iff $val(\varphi)\geq 0$ or $val(\wedge\mathcal{A})\leq 0$. If the first, then the value of the antecedent is irrelevant, and $f\mathcal{A}$ will work. If the second, then there is some $\psi\in\mathcal{A}$ such that $val(\psi)\leq 0$. In that case, let $\Psi=\{\psi\}$, and $val(f\{\psi\})=val(\psi)\leq 0$. □

(140) **Remark.** There are collections $\{\$_k\}$ of RM valuations in which the thesis of Proposition 2.5 fails, i.e. under which $\{\$_k\} \Vdash \wedge\mathcal{A}\supset\varphi$ but it is not the case that there is a $\Psi\subseteq\mathcal{A}$ such that $\{\$_k\} \Vdash f\Psi\supset\varphi$.

Pf. We need only a single example. Consider the expression $A\wedge B\supset C$, under the following two valuations.

	$\$_1$	$\$_2$
A	2	−1
B	−1	1
C	−1	−1

The antecedent of the Proposition 2.5 thesis is satisfied because $\{\$_1,\$_2\} \Vdash A\wedge B\supset C$.

But the consequent of Prop. 2.5 fails for every Ψ, due to the following counterexamples:

NOT $\$_1 \Vdash A\supset C$ $(2 \not\supset -1)$

NOT $\$_2 \Vdash B\supset C$ $(1 \not\supset -1)$

NOT $\$_1 \Vdash A\circ B\supset C$ because $2\circ -1=2$ (NB: $2\wedge -1=-1$). □

Given the parallelism we are seeking between the behavior of systems and that of the entire model space, it is natural to ask whether Proposition 2.5 corresponds to a theorem about RM. In the broadest construal of its terms, this is manifestly not the case. From RM$\Vdash \wedge\mathcal{A}\supset\varphi$, where $\mathcal{A}$ is a set of RM sentences, we are not licensed to conclude that there is a $\Psi\subseteq\mathcal{A}$ such that RM$\Vdash f\Psi\supset\varphi$.

(141) **Remark. A non-metatheorem.** Proposition 2.5 fails for RM. There are set of sentences $\mathcal{A}$ such that RM$\Vdash \wedge\mathcal{A}\supset\varphi$, without $\mathcal{A}$ containing a subset Ψ with the property that RM$\Vdash f\Psi\supset\varphi$.

Pf. Let $\mathcal{A} = \{A\vee B, A\vee\neg B, \neg A\vee B)$, let $\varphi= A\wedge B$. Then

(*) RM$\Vdash [(A\vee B) \wedge (A\vee\neg B) \wedge (\neg A\vee B)] \supset A\wedge B$

but

(**) NOT RM$\Vdash [(A\vee B) \circ(A\vee\neg B) \circ(\neg A\vee B)] \supset A\wedge B$

Since all PC tautologies are valid in RM, it is clear that (*) is true. Now consider the valuation val(A)= 1, val(B)= −2.

[i] $\text{val}(A\wedge B) = \min(1,-2) = -2$

[ii] $\text{val}((A\vee B) \wedge (A\vee\neg B) \wedge (\neg A\vee B)) = \min(\max(1,-2), \max(1,2),\max(-1,-2))$
$= \min(1,2,-1) = -1$

[iii] $\text{val}((A\vee B) \circ(A\vee\neg B) \circ(\neg A\vee B)) = \min \textsc{Out}(1,2,-1) = \textsc{Out}(1,2,-1) = 2.$

Ergo val$\Vdash$ [ii]$\supset$[i], as expected, but NOT val$\Vdash$ [iii] $\supset$[i]. Note that this suffices to establish the claim, since no *proper* subset of $\mathcal{A}$ can possibly work. If there were such a $\Psi\subseteq\mathcal{A}$, with RM$\Vdash f\Psi\supset\varphi$, then we'd have RM$\Vdash \wedge\Psi\supset\varphi$ (because conjunction entails fusion), which is manifestly false, as the reader may easily verify via a PC argument. □

We are undone here by the behavior of the extensional connectives within the scope of conjunction. Recalling that an ERC corresponds not to a general sentence of RM but to a polyvaluated expression, and that only intensional expressions are reducible, by VS rules, to single polyvaluated expressions, we advance a more modest version of the thesis:

(142) **Claim**. Let $\mathcal{A}$ be a set of sentences of S. RM$\Vdash \wedge\mathcal{A}\supset\varphi$ iff there is a $\Psi\subseteq\mathcal{A}$ such that $\Vdash f\Psi\supset\varphi$.

To prove this, it is useful to establish certain basic properties of valuation within systems.

Observe first that the range of valuations generated by a system is typically going to be defective in certain respects. It will not be the case, in general, that cycling through all permutations of order on the constituent valuations will succeed in assigning every possible relevant valuation pattern to a set of prop letters. A system may fall short in two, complementary ways – horizontally and vertically, as it were: when no single letter receives all the values the system can assign; and when there are correlations across prop letters, so that not all patterns of valuations are assigned.

To see this, recall that the Sugihara value assigned by $\langle V,R\rangle$ to a polyvaluated expression ψ is computed from the following formulas:

$$val(\psi)=\min \text{OUT}\{val_i(\psi)\}$$
$$val_i(\psi)= v_i(\psi)\times dom(v_i)$$

A nonzero value is determined from the *principal valuation* under R, the v_k in V assigning a polar value which is greatest in the R order, which yields the outermost value of the val_i's. For N valuations, the weighting factor $dom(v_i)$, determined from the position of v_i in the order R, ranges from N to 1, while $v_i(\psi)$ ranges over $\{-1,0,+1\}$. The possible Sugihara values therefore range from N to −N inclusive. But individual assignments, under permutation, may be restricted so as not to exhaust the combinatorics.

• 'Horizontally'. If the polyvaluation of prop letter V(A) contains *k* polar specifications of N constituent valuations, then the dominance of its assignments, determined from possible principal valuations, runs only from N to k.

• 'Vertically'. For any two prop letters A and B, $v_k(A)$ and $v_k(B)$ are coupled permanently, which will prevent A, B from assuming all values independently of each other.

Patterns of valuation thus differ in structure between the system and the whole model space. But the apparent divergence proves to be inconsequential, when examined in the light of Meyer's foundational results. As recorded in (115), a sentence of RM with N prop letters is valid iff it holds in a Sugihara set containing 2N values. Furthermore, the key feature of the Sugihara set of integers is that it contains −k if it contains k. Above, for simplicity of presentation, we narrowed our focus to sets containing ranges of integers −N...N (with or without 0), but any set with the basic property will do just as well. Thus, {-117, −73, −2, 2, 73, 117} is every bit as useful as the more staid {−2,−1,1,2}. By 'Sugihara set', let us reinstate general reference to any set of integers with the basic property, regardless of gaps in the range covered.

Observe that any set of integers that *contains* a Sugihara set of cardinality 2N will also be completely diagnostic for the theoremhood in RM of sentences with N prop letters. To show that systems in general provide an adequate setting for validity determination, we need only show that there are systems for every N that *contain* combinatorically exhaustive assignments of 2N Sugihara values to N prop letters.

Such systems may be constructed as follows. For k prop letters $A_1...A_K$, let the polyvaluation V contain *all* mappings $v:\{A_1,...,A_K\} \rightarrow \{-1,0,1\}$; this will de-correlate the assignments in the way we desire. In addition to these, let V contain K−2 copies of the constant mapping $z:A_i\mapsto 0$; this will allow access to all the polar values. Let us call such a polyvaluation 'K-complete'. We will find that Σ (V) generates a set of values that contain a Sugihara set of cardinality 2K.

(143) **Def. K-complete.** Let V be a polyvaluation of a set of K prop letters. V is said to be 'K-complete' if V includes (i) every mapping $v:\{A_1,...,A_K\}\rightarrow\{-1,0,1\}$, and, additionally, (ii) K copies of the mapping $z:\{A_i\}\rightarrow\{0\}$.

A quick example will give a sense of how this works. Consider the following polyvaluation, which meets the criterion of 2-completeness.

	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8	v_9
A	T	T	T	*e*	*e*	*e*	F	F	F
B	T	*e*	F	T	*e*	F	T	*e*	F

Here's some representative Sugihara values obtained from various orderings:

$v_1> \ldots$	$v_7>\ldots$	$v_2>v_6>\ldots$	$v_5>v_2>v_8>v_6>\ldots$
val(A) = 9	val(A) = −9	val(A) = 9	val(A) = 8
val(B) = 9	val (B) = 9	val(B) = −8	val(B) = −5

In this case, the system $\Sigma(V)$ contains within it the Sugihara set {−9,−8,8,9), as well as various other assignments.

(144) **Lemma.** Let V be K-complete for some K. Then $\Sigma(V)$ assigns every combination of values in a Sugihara set S_{2K} of cardinality 2K to a set of K prop letters.

Pf. It suffices to show that any Sugihara valuation function $\$:\{A_1,...,A_K\} \to S_{2K}$ can be represented by OP semantics under some ordering, where $\mathcal{A}=\{A_1,...,A_K\}$ is a set of K prop letters. If N is the total number of valuations in V, then let S_{2K} be the set $\{-N, -N+1, \ldots, -N+K-1, N-K+1, N-1, \ldots, N\}$. For any such \$, let us distinguish between the dominance and the polarity of each assignment:

$$\mathrm{dom}(\$A_i) = |\$A_i|$$
$$\pi(\$A_i) = \mathrm{sign}(\$A_i) = \$A_i/|\$A_i|$$

For conciseness, we write \$X for \$(X). The definition of π is sound: $\$A_i \neq 0$ because $0 \notin S_{2K}$.

Let us now construct an ordering on V that produces \$. We sort the A_i by dominance of the $\$A_i$. Let $\mathcal{A}_d = \{A_i |\ \mathrm{dom}(\$A_i)=d\}$.We now construct a valuation $v_d:\mathcal{A}\to\{-1,0,+1\}$, according to the following recipe:

$$v_d(A_i) = \pi(\$A_i) \text{ for } A_i\in\mathcal{A}_d$$
$$= 0 \text{ otherwise.}$$

In short, v_d assigns −1 or +1 to the prop letters whose \$-value is −d or d, and it assigns 0 to all others. Any such v_d must belong to V, since by the definition of K-completeness, V contains all maps from $\{Ai, 1\leq i\leq K\}$ to $\{-1,0,+1\}$. Further, any two such valuations v_{d1}, v_{d2} must be distinct for $d_1\neq d_2$. (In fact, we can say more: if $v_{d1}\in\{-1,+1\}$ then v_{d2}=0. Thus, if we consider the subpolyvaluation V' consisting of the v_d's, we will find that $V'(A_i)$ contains precisely one polar coordinate, with the rest 0.) The upshot is that the the v_d's form a subset of V, with no repetition among them.

Let us now arrange the v_d's in the appropriate order. Number the positions in an ordering of V from N to 1. Let each v_d occupy position d. Let any other positions among the first K-1 be filled by valuations that assign 0 to all A_i. (We have enough of these by definition of K-completeness.) This arrangement reproduces val(A_i). To see this, consider any instance of valuation, say val(A_i)=m. By construction, $v_{|m|}$ = sign(m). But $v_j(A_i)=0$ for all other v_j with $N\geq j\geq N-K+1$. Ergo $v_{|m|}$ is the principal valuation of A_i, and it assigns the value $|m|\times\mathrm{sign}(m)=m$. □

We may now use Proposition 2.5, which is concerned with the structure of systems, to prove the claim (142), which is about the structure of the entire model space for RM.

(145) **Proposition 7.3**. Let $\mathcal{A}$ contain sentences of S. RM$\Vdash \bigwedge\mathcal{A}\supset\varphi$ iff there is a $\Psi\subseteq\mathcal{A}$ such that $\Vdash f\Psi\supset\varphi$.

Pf. Only the LR direction is of interest. Assume RM$\Vdash \bigwedge\mathcal{A}\supset\varphi$. Let K be number of prop letters in $\mathcal{A}\cup\{\varphi\}$. Let V be a K-complete polyvaluation. By the lemma, $\Sigma(V)\Vdash \bigwedge\mathcal{A}\supset\varphi$. By Proposition 2.5, there is a $\Psi\subseteq\mathcal{A}$ such that $\Sigma(V)\Vdash f\Psi\supset\varphi$. But $\{\Psi\cup\varphi\}$ has no more prop letters than $\mathcal{A}$. Therefore, by Meyer's result (115) and Dunn's (114), RM$\Vdash f\Psi\supset\varphi$. □

[C] **Arrow and Horseshoe**

We conclude this brief foray into system theory by observing that the relationship between '$\supset$' and '$\rightarrow$' also takes a strong form within the system. RM gives us a general one-way implicational relation:

$$\text{RM}\Vdash (\alpha\rightarrow\varphi)\rightarrow(\alpha\supset\varphi)$$

This is just a variant of the RM theorem $\Vdash$ (A+B) $\rightarrow$ (A$\vee$B), *v.* AB:396(R54), obtained by substituting ¬A for A. The statement 'if $\Vdash\alpha\rightarrow\varphi$ then $\Vdash\alpha\supset\varphi$' follows immediately by modus ponens.

The converse does not hold: it is not guaranteed that $(\alpha\supset\varphi)\rightarrow(\alpha\rightarrow\varphi)$; consider the valuation v(α)=1, v(φ)=0. Nor do we have 'if $\Vdash\alpha\supset\varphi$ then $\Vdash\alpha\rightarrow\varphi$'; for example, the familiar PC theorem $\Vdash(A\wedge\neg A)\supset B$ holds in RM, as do all PC theorems, but $(A\wedge\neg A)\rightarrow B$ is not valid. We can obtain a near converse, however, if we limit antecedent and consequent to be 'nontrivial' in relevant respects. When neither φ nor ¬α itself valid, the validity of '$\alpha\supset\varphi$' does guarantee the validity of '$\alpha\rightarrow\varphi$'.

The result does not carry over to random collections of RM valuations, however. Consider this pair:

	$\$_1$	$\$_2$
α	2	−1
φ	1	−1
$\alpha\supset\varphi$	1	1
$\alpha\rightarrow\varphi$	−2	1

Here neither ¬α nor φ is designated across $\{\$_1, \$_2\}$, yet despite the fact that $\alpha\supset\varphi$ *is* designated under both valuations, we find that $\alpha\rightarrow\varphi$ fails

Within the *system*, however, where we restrict ourselves to polyvaluated expressions, the two species of implication connectives fall into line. The required nontriviality conditions should have a ring of familiarity from Prop. 1.1 (6), p. 6, and Remark (33), p. 18.

(146) **Proposition 7.4**. Let V be any polyvaluation . Let α, φ be any sentences in the language of S, such that it is *not* the case that V$\Vdash\varphi$ and *not* the case that V$\Vdash\neg\alpha$. The following holds:

$$\text{if } \Sigma(V)\Vdash\alpha\supset\varphi \text{ , then } \Sigma(V)\Vdash\alpha\rightarrow\varphi.$$

Pf. We show the contrapositive of the main assertion. Assume that it is not the case that $\Sigma(V)\Vdash\alpha\rightarrow\varphi$. From Remark (137), it follows that it is not the case that V$\Vdash\alpha\rightarrow\varphi$, coordinatewise. Therefore some constituent valuation of V, call it v_k, has $v_k(\alpha\to\varphi)$=F. This happens in only three conditions.

[i] $v_k(\alpha)$=T, $v_k(\varphi)$=F. In this case, consider the OP in which v_k is greatest; clearly, $\alpha\supset\varphi$ fails on this OP, and so it fails to be valid throughout $\Sigma(V)$.

[ii] $v_k(\alpha)$=T and $v_k(\varphi)$=e. Because by assumption φ is not valid throughout V, there must be some v_j with $v_j(\varphi)$=F. Consider the OP in which v_k is greatest and v_j is second-greatest. In the corresponding RM valuation val, we have val(α)>0, val(φ)<0, so that $\alpha\supset\beta$ fails again to valid throughout $\Sigma(V)$.

[iii] $v_k(\alpha)=e$ and $v_k(\varphi)$=F. By assumption it is not the case that $V \Vdash \neg\alpha$, it must be $V(\neg\alpha)$ contains F at some coordinate, call it v_i, so that $v_i(\alpha)$=T.. Consider the OP in which v_k is greatest and v_i second-greatest. In the corresponding RM valuation val, we have val(α)>0, val(φ)<0, so that $\alpha\supset\beta$ fails yet again to be valid throughout $\Sigma(V)$. □

An immediate corollary is that the thesis of Proposition 7.4 generalizes to RM, via Lemma (144), since it holds in all systems.

7.7 Syntactical Manipulations of Paramount Interest

We conclude with a look at how syntactic considerations can be used to arrive at results semantically achieved in earlier sections of this paper. Let us examine the conditions under which fusion can be coerced into behaving like conjunction. Basic principles involving conjunction include these:

Conjunction Out	$\alpha\wedge\beta \to \beta$
Weakening	If $\alpha\to\beta$, then $\alpha\wedge\varphi\to\beta$

Neither holds for fusion, as we have seen repeatedly:

"Fusion Out" (FO)	$\alpha\circ\varphi\to\varphi$	falsified by v(α)=T, v(φ)=*e*.
"Fusion Weakening" (FW)	If $\alpha\to\beta$, then $\alpha\circ\varphi\to\beta$	falsified by v(α)=v(β)=*e*, v(φ)=T.

Falsifying valuations from S/RM3 are shown; by 'invalidity ascent' (114), these falsify for all RM(k) for k≥3, including RM itself. The situation is not entirely dismal, however, as noted in §3 above, since we can find useful conditions under which the differences between fusion and conjunction vanish.

Upon observing that the only falsifying evaluation for FO is the one given, we can ensure FO as in $V \Vdash \alpha\circ\varphi\to\varphi$ by demanding of the polyvaluation that the places where α is assigned T are subset of those where φ is assigned a *polar* value, i.e. that $v_k(\alpha)=T \Rightarrow v_k(\varphi)\in\{T,F\}$ for each v_k in V: compare Corollary 2 to Proposition 3.3, (37).

We may also approach the problem syntactically. Fusion is subject to importation and exportation of premises, just like conjunction.

(147) **Transportation**. $[\alpha\circ\beta\to\varphi] = [\alpha\to(\beta\to\varphi)]$

From this we have, substituting φ for β,

$$[\alpha\circ\varphi\rightarrow\varphi] = [\alpha\rightarrow(\varphi\rightarrow\varphi)].$$

This gives us a useful syntactic condition on FO: $\alpha\circ\varphi\rightarrow\varphi$ holds iff $\alpha\rightarrow(\varphi\rightarrow\varphi)$ holds.

The expression $\varphi\rightarrow\varphi$, equivalently $\neg\varphi+\varphi$, repays scrutiny. To cut down on the profusion of arrows let us abbreviate it to $T(\varphi)$. Valid in S/RM, under polyvaluation, $T(\varphi)$ contains a T at every coordinate where φ has either of the polar values T,F. We might say that $T(\varphi)$ detects all the polar values in $V(\varphi)$ and marks them as T. We know that any implicational expression $\alpha\rightarrow\beta$ will be valid under a polyvaluation V iff two conditions are met:

[i] every T in $V(\alpha)$ is matched to a T in $V(\beta)$ – the "W-condition"

i.e. $v_k(\alpha)=T \Rightarrow v_k(\beta)=T$

[ii] every F in $V(\beta)$ is matched to an F in $V(\alpha)$ – the "L-condition"

i.e. $v_k(\beta)=F \Rightarrow v_k(\alpha)=F$

Since $T(\varphi)$ contains no F's, only the first condition applies. The expression $\alpha\rightarrow T(\varphi)$ will therefore be valid iff $v_k(\alpha)=T \Rightarrow v_k(\varphi\rightarrow\varphi)=T$ i.e. iff $v_k(\alpha)=T \Rightarrow v_k(\varphi) \in\{T,F\}$, i.e. iff the T coordinates of α are a subset of the polar coordinates of φ.

W-compliance emerges when we ask for equivalence between conjunction and fusion:

$$\alpha\circ\varphi\rightarrow\varphi \quad \text{iff} \quad \alpha\rightarrow T(\varphi)$$
$$\alpha\circ\varphi\rightarrow\alpha \quad \text{iff} \quad \varphi\rightarrow T(\alpha)$$

Equivalence, as in $\alpha\circ\varphi = \alpha\wedge\varphi$, occurs when both conditions hold, so that any T in either $V(\alpha)$ or $V(\varphi)$ is matched to a polar value in the other.

The conditions that allow for the stronger property "Fusion Weakening" are slightly more elaborate.

(148) FW. Suppose $\alpha\rightarrow\beta$. Then $\alpha\circ\varphi\rightarrow\beta$ iff $\varphi\rightarrow T(\alpha)\circ T(\beta)$, equivalently $\varphi\rightarrow T(\alpha)+T(\beta)$.

The two expressions are equivalent because $A+B \rightleftarrows A\circ B$ when the only values assumed by A, B are the designated ones, *i.e.* when $v\Vdash A$ and $v\Vdash B$. The sense of the condition $\varphi\rightarrow T(\alpha)\circ T(\beta)$ is that, under polyvaluation, every T assigned to φ is matched to a polar value in either α or β.

To establish FW syntactically, it is useful to note a couple of facts about S. First, fusion is well-behaved with respect to $\rightarrow$:

(149) **Double Weakening**. If $\alpha\rightarrow\beta$ then $\alpha\circ\varphi\rightarrow\beta\circ\varphi$ for any φ.

Pf. When φ evaluates to F, the consequent implication holds regardless of $\alpha\rightarrow\beta$. When φ evaluates to e, $\alpha\circ\varphi\rightarrow\beta\circ\varphi$ evaluates the same as $\alpha\rightarrow\beta$. When φ evaluates to T, it only induces changes in valuation of the fusions when either of $\alpha,\beta = e$, as in $e\rightarrow T$, $F\rightarrow e$, and $e\rightarrow e$, but then the result of double weakening by φ is valid when $\alpha\rightarrow\beta$ is. □

Second, $T(\alpha)$ functions as a kind of fusional identity with respect to α:

(150) **Local Identity**. $\alpha\circ T(\alpha) = \alpha$

Pf. $T(\alpha)$ assumes the same value as α for $v(\alpha)=T$ and $v(\alpha)=e$. Where $v(\alpha)=F$, we have $v(T(\alpha))=T$, but $T\circ F=F$ so that the result equals $v(\alpha)$. □

Let us now put these to work.

(151) **FW I**. Suppose $\alpha\rightarrow\beta$ and $\varphi\rightarrow T(\alpha)\circ T(\beta)$. Then $\alpha\circ\varphi\rightarrow\beta$.

Pf.	i	$\varphi\rightarrow T(\alpha)\circ T(\beta)$	assumed
	ii	$\alpha\circ\varphi \rightarrow \alpha\circ T(\alpha)\circ T(\beta)$	DW from (i)
	iii	$\alpha\circ T(\alpha)\circ T(\beta) \rightarrow \alpha\circ T(\beta)$	LI
	iv	$a\circ\varphi\rightarrow\alpha\circ T(\beta)$	transitivity of $\rightarrow$, from (ii), (iii)
	v	$\alpha\rightarrow\beta$	assumed
	vi	$\alpha\circ T(\beta) \rightarrow \beta\circ T(\beta)$	DW from (v)
	vii	$\beta\circ T(\beta) \rightarrow \beta$	LI from (vi)
	vii	$\alpha\circ\varphi \rightarrow \beta$	transitivity of $\rightarrow$, from (iv),(vi), (vii).

(152) **FW II**. Suppose $\alpha\rightarrow\beta$ and $\alpha\circ\varphi\rightarrow\beta$. Then $\varphi\rightarrow T(\alpha)+T(\beta)$.

Pf.	i	$\alpha\circ\varphi\rightarrow\beta$	assumed
	ii	$\varphi\rightarrow(\alpha\rightarrow\beta)$	exportation from (i)
	iii	$(\alpha\rightarrow\beta) \rightarrow (\alpha\rightarrow\beta) + (\beta\rightarrow\alpha)$	thm. of S
	iv	$\varphi \rightarrow (\alpha\rightarrow\beta) + (\beta\rightarrow\alpha)$	transitivity of $\rightarrow$, from (ii),(iii)
	v	$\varphi \rightarrow (-\alpha)+\beta + (-\beta) +\alpha$	definition of $\rightarrow$
	vi	$\varphi \rightarrow (\alpha\rightarrow\alpha)+(\beta\rightarrow\beta)$	definition of $\rightarrow$

Paralleling the line of attack in §3, we observe that FW implies FO: we need merely identify α and β in the statement of FW, and observe that $\alpha\rightarrow\alpha$ is a theorem.

8. Constraint Logic

Summary. When a constraint is understood as a vector of comparative values, the logic S can be applied to constraints just as it is applied to ERCs. Key constraint relations are definable in logical terms, and the notion of a constraint hierarchy is itself expressible in terms of fusion and fission. Logical relations between constraints determine properties of their ranking; for example, a fusion is a greatest lower bound for its fusands in the MSH.

The full utility of the logical approach, and of VS in particular, only becomes apparent when we allow the logical connectives to relate *constraints* as well as ranking arguments. We may then work with fusion of constraints $C_1 \circ C_2$, fission $C_1 + C_2$, implication $C_1 \rightarrow C_2$, and negation $\neg C_1$, as well as with more complicated collocations.

As noted above, p. 22, a constraint can be construed extensionally over an ERC set as a column vector in the array where the ERCs are row vectors. Any vectorial operation may then be applied to columns as well as to rows. Given a set of desired optima, the 3-valued comparative structure is defined on candidate sets, and the logical connectives become fully meaningful.

In this transformation of perspective, we treat constraints as propositional atoms and ERCs as functions evaluating *them*. Give a valuation v_i: $\{\alpha_k\} \rightarrow \mathbf{3}$ of the familiar type, we construct a corresponding propositional object v_i^* and a valuation α_k^* defined by the relation $\alpha_k^*(v_i^*) = v_i(\alpha_k)$. In the present context, expressions such 'constraint C_i' refer to such v_i^*.

We may now write $\alpha^*(C)$ for the value of C at α. A set of ERCs $\mathcal{A}$ dualizes to a polyvaluation $\mathcal{A}^*$, and $\mathcal{A}^*(C)$ is a (column) vector in the valuation table of ERCs $\mathcal{A}$. Over collections of constraints, we have expressions such as

$$\mathcal{A}^* \Vdash \otimes_i C_i \qquad \mathcal{A}^* \Vdash \oplus_i C_i \qquad \mathcal{A}^* \Vdash \neg C_i \qquad \mathcal{A}^* \Vdash Ci \rightarrow C_j$$

and indeed, we may evaluate any intensional expression over constraints in the language of S. For purposes of clarity, we use a special notation for collective fusion $\otimes$ and collective fission $\oplus$, as a reminder that we are operating over constraints rather than ERCs.

One application that we have already seen lies in the development of RCD: in the definition of satisfaction of nontrivial argument vector by a hierarchy H, (41), p. 22, repeated here in the current notation:

(153) $H \vDash \alpha$ iff , for some $m \leq n$, $\alpha^*(\otimes_{i \leq m} C_i) = T$.

An ERC α is trivial iff the following condition is met

(154) **Triviality of ERC α.** $\alpha^* \Vdash \otimes_i C_i = \oplus_i C_i$

Important relations between constraints may be directly formulated in these terms. Two constraints P and Q cannot possibly conflict iff their fusion and fission are equal, *i.e.* iff every ERC when restricted to {P,Q} is trivial.

(155) **Remark**. **Necessary Nonconflict.** Let P, Q be constraints over an ERC set $\mathcal{A}$. Then P and Q do not conflict iff $\mathcal{A}^* \Vdash P\circ Q = P+Q$

Pf. The expression 'P∘Q' differs from 'P+Q' on a valuation v only when v(P)=T and v(Q)=F, or vice versa, which is necessary for conflict. □

In the MSH for $\mathcal{A}$, having P sit in a higher stratum than Q does not mean that P conflicts with Q, because P could have earned its place in |P| merely by virtue of being ranked as high as possible. Thus there need be no $\alpha\in\mathcal{A}$ with $\alpha^*(P)$=T and $\alpha^*(Q)$=F. But with a consistent ERC set, there must be a conflict relation that puts Q against $\otimes|P|$, the fusion of all the constraints in |P|. If not, then there is no reason why Q should not sit in |P|, or even higher. The notion of fusion of constraints thus allows us to state a tight relation between a stratum and the constraints subordinate to it.

(156) **Remark. Stratal Conflict**. Let P,Q be constraints over an ERC set $\mathcal{A}$, with |P|>|Q| in the MSH. Then there is an $\alpha\in\mathcal{A}$ such that $\alpha^*(\otimes|P|)\in\{T,e\}$ and $\alpha^*(Q)$=F, with $\alpha^*(C)=e$ for any constraints C with |C|>|P|. If $\mathcal{A}$ is consistent, then for at least one such α, it must be the case that $\alpha^*(\otimes|P|)$=T with $\alpha^*(C)=e$ for any constraints C with |C|>|P|.

Pf. Suppose |P|>|Q| in $\mathcal{H}(\mathcal{A})$, where $\mathcal{A}$ need not be consistent. The set $\mathcal{A}$ is partitioned into two disjoint subsets by rank of arguments, defined in (53). Let A_0 denote the set of arguments given a polar valuation (necessarily T) by a constraints ranked above P or which are degenerate. Let A_P denote the complement set $\mathcal{A}-A_0$, and let Σ_P denote the set of constraints at rank |P| and below. Then, as is clear from the recursive definition of RCD (49), we must have $A_P^*\Vdash\otimes|P|$, and in fact |P| is precisely the maximal set of constraints from Σ_P for which this holds. But if there is no $\alpha\in A_P$ such that $\alpha^*(Q)$=F, then $Q\in|P|$, contrary to assumption. So there must be an $\alpha\in A_P\subseteq\mathcal{A}$ with $\alpha^*(Q)$=F and $\alpha^*(\otimes|P|)\in\{T,e\}$. It is also clear from the RCD construction that $\alpha^*(C)=e$ for any C with |C|>|P|.

Now by the RCD construction of strata, we can have $\alpha^*(\otimes|P|)=e$ for *all* $\alpha\in A_P$ only if |P| is the top stratum. The constraint set divides into those, like P, for which $C(f\mathcal{A})=\delta$ and the remainder for which $C(f\mathcal{A})$=L. Whence $f\mathcal{A}\in\mathcal{L}^+$ and $\mathcal{A}$ is inconsistent. Contrapositively, if $\mathcal{A}$ is consistent, some $\alpha\in\mathcal{A}$ must have $\alpha^*(\otimes|P|$=T. □

Let us say that a constraint P *occludes* a constraint Q if when P is ranked above Q it is guaranteed that Q can never decide an ERC. This means that Q never assumes a value from {T,F} when P assumes the value *e*. We take this semantic relation as the definition of occlusion.

(157) **Def. Occlusion.** P *occludes* Q over an ERC set $\mathcal{A}$ iff for every $\alpha\in\mathcal{A}$ it is the case that $\alpha^*(Q)\in\{T,F\}$implies $\alpha^*(P)\in\{T,F\}$; equivalently, by contraposition, iff $\alpha^*(P)=e \Rightarrow \alpha^*(Q)=e$.

The following syntactic condition determines this relationship; recall that T(P) abbreviates P→P, equivalently ¬P+P.

(158) **Remark. Occlusion.** P occludes Q on $\mathcal{A}$ iff $\mathcal{A}^* \Vdash T(Q) \rightarrow T(P)$, equivalently iff $\mathcal{A}^* \Vdash P \circ \neg P \rightarrow Q \circ \neg Q$.

Pf. Right to left. The formula $X \circ \neg X$ takes on only 2 possible values: F,*e*. If $\nu(P)=e$ then $\nu(P \circ \neg P)=e$, and then from the assumption, it follows that $\nu(Q \circ \neg Q)=e$. Whence Q=*e*, so Q doesn't decide either.

Left to right. The only valuations on which $P \circ \neg P \rightarrow Q \circ \neg Q$ fails are $\nu(P)=e$ and $\nu(Q) \in \{T,F\}$. But the definition of occlusion requires $\nu(Q)=e$ when $\nu(P)=e$. □

It is apparent that P occludes P, and that P∘Q occludes both P and Q. Indeed, *any* intensional combination of P_i occludes any other intensional combination of the P_i. Occlusion of Q by P also means that there can be no ranking argument between Q and any constraint ranked below P. (Grimshaw 1999/2001). We will return to the logic of occlusion below.

More surprising, perhaps, the very notion of a constraint hierarchy can be rendered exactly in terms of fusion and fission. Consider the following expression:

$$P + (P \circ Q)$$

The value of the expression P + (P∘Q) is determined in the following way:

1. **P decisive.** P + (P∘Q) is true if v(P)=T. It is false if v(P)=F.
2. **Else Q.** If v(P)=*e*, then it assumes the value of Q.

This corresponds precisely to the way an ERC is evaluated over the hierarchy P>>Q. The generalization to more than two constraints is immediate.

(159) **[P>>Q>>R>>S …]** $P + (P \circ Q) + (P \circ Q \circ R) + (P \circ Q \circ R \circ S) \ldots$

With a little notation, an expression can be concocted that represents the general notion of a constraint hierarchy over a set of constraint $\{P_i\}$, where the subscript indicates position in the total ordering, starting with 1 for the highest position. With any totally-ordered strict-domination hierarchy

(160) $$\mathop{\gg}_{i=1}^{n} P_i$$

we associate the following logical expression

(161) $$\bigoplus_{k=1}^{n} \bigotimes_{i=1}^{k} P_i$$

Call the fission/fusion formula the 'VS representation' of the domination hierarchy. It is clear that a hierarchy satisfies a set of ERCs $\mathcal{A}$ iff its VS representation is valid (in the VS sense) under $\mathcal{A}^*$, *i.e.* evaluates to a vector with every entry designated. The requirement that every ERC must hold – that $\bigwedge\mathcal{A}$ holds – is matched by the coordinatewise requirement of VS validity for $\mathcal{A}^*(\oplus\otimes P_i)$.

Note that the expression 'P+(P∘Q)' is *not* the generic equivalent of 'P>>Q', with the meaning 'P dominates Q in some hierarchy'. In particular, from

$P + (P \circ Q) + (P \circ Q \circ R)$ 'P>>Q>>R' one cannot deduce

$P + (P \circ R)$ 'P>>R'.

(Consider the valuation v(P)=*e*.) Thus, the expressions in (159) and (161) characterize entire hierarchies, not pairwise order relations.

Note also that the quasi-dual expression, with ∘ and + interchanged,

$P \circ (P+Q) \circ (P+Q+R)$

is logically equivalent to the expression given above for 'P>>Q>>R'. In both cases, the value of the expression is determined by the value of the first prop letter that does not evaluate to *e*. This has a kind of distributive look to it, though the general distributive law doesn't hold in S. Distribution fails in one direction: what's missing is the implication from P+(Q∘R) to (P+Q)∘(P+R). (To see this, consider the assignment of F to P, and (*e*,T) or (T,*e*) to Q,R.) A restricted form remains valid, however, obtained by identifying one of the inner elements of P+(Q∘R) with the outer, thereby eliminating the fatal valuation:

$P+(P\circ Q) \rightarrow (P+P)\circ(P+Q) = P\circ(P+Q).$

Bringing in the half of the distributive law that does hold, $P\circ(Q+R)\rightarrow P\circ Q+P\circ R$, we have

$P\circ(P+Q) \rightarrow P\circ P+P\circ Q = P+P\circ Q.$

Taken together, these give us the desired equivalence:

$P+(P\circ Q) = P\circ(P+Q).$

This may be generalized as follows:

(162) **Generalized restricted distributivity**

$$\bigoplus_{k=1}^{n} \bigotimes_{i=1}^{k} P_i = \bigotimes_{k=1}^{n} \bigoplus_{i=1}^{k} P_i$$

A *stratified* hierarchy is obtained by using fusional agglomerations in place of the single prop letters in a linearly orderered hierarchy. From A>>B, by setting A:=P∘Q and B:=R∘S, we obtain, for example, the expression

(163) $(P\circ Q) + (P\circ Q\circ R\circ S)$

which represents the hierarchy {P,Q}>>{R,S}. Expression (163) holds under a valuation α* only if both of P and Q are designated; if both are *e*, then both of R,S must be designated. This mirrors exactly the conditions under which the associated stratified hierarchy validates the ERC α.

To construct the general expression for a stratified hierarchy, we define a *net* N(*n*) on positive integer *n* as a set of positive integers meeting the following conditions:

[1] $n\in$N,

[2] $k\in$N$\Rightarrow k\leq n$.

A net on *n* is just a set of positive integers less than *n*, which also includes *n*. The members of a given net defined the way the elements of $\{P_i\}$ are clumped together into strata. For example, a net {1,3,5} will be used to define a 3-stratum hierarchy on five constraints:

$P_1 >> P_2, P_3 >> P_4, P_5$	→	$P_1 + P_1\circ P_2\circ P_3 + P_1\circ P_2\circ P_3\circ P_4\circ P_5$
{1, 3, 5}		{1, 3, 5}

Any net N thus gives rise to a stratified hierarchy on $\{P_1, P_2, ...P_n\}$. If the net is maximal, in that it includes all positive $k\leq n$, then each stratum has just one constraint in it, and we are back at the totally-ordered version of constraint ranking. If the net is minimal, with only n in it, then the hierarchy has one stratum. A stratifed hierarchy over a net N meets this description:

(164) **Stratified Hierarchy** $$\bigoplus_{k\in N(n)} \bigotimes_{i=1}^{k} P_i$$

Each fusional subconstituent $\otimes_1^k P_i$ corresponds to a stratum in the constraint hierarchy, and we will carry over the term to refer to it. This represensentation makes it clear that a stratum functions like a single constraint that it is the fusion of the constraints that it contains.

The VS representation of a stratified hierarchy is in general *not* equivalent to the one obtained from it by swapping $\circ$ and + throughout. Observe that where

$$(P\circ Q) + (P\circ Q\circ R\circ S)$$

is false if v(P)=F, regardless of the valuation of other letters, its $\circ$/+-quasi-dual

$$(P+Q)\circ(P+Q+R+S)$$

is true whenever v(Q)=T even if v(P)=F. The correct application of $\circ$/+ swapping would be this:

$$(P\circ Q)\circ(P\circ Q + R\circ S)$$

where the relevant fusional agglomerates $P\circ Q$ and $R\circ S$ are treated as units, like prop letters in (162)

The expression of the form $\otimes_N \oplus P_i$ obtained by complete $\circ$/+-swapping holds when *at least one* constraint in the highest nonneutral stratum (fissile and shaped $\oplus P_i$) is T, or when all strata are neutral. This formalizes a notion of crucial partial order among constraints, by which an argument is validated by a stratum if there is *some* ranking of the constraints in the stratum that validates it.

It is true, however, that whenever $\oplus_N \otimes P_i$ holds under some valuation, then $\otimes_N \oplus P_i$ holds as well; equivalently, that whenever $\otimes_N \oplus P_i$ fails, then $\oplus_N \otimes P_i$ fails as well. This establishes the one-way implication to the effect that $\Vdash \oplus_N \otimes P_i \rightarrow \otimes_N \oplus P_i$.

The minimal stratified hierarchy occupies a special position in the set of formulae, just as it does among hierarchies. Of all stratified hierarchies true under some given valuation, represented as in (164), the minimal stratified hierarchy has the minimal number of (fusional) strata and given that, the *maximal* number of elements per fusional stratum.

To measure minimality, let us define the (Boolean) *deviation* of fusional stratum in (164) as the number of prop letters missing from it: for n letters, a k letter fusion has a deviation of $(n - k)$. Let the deviation of the whole hierarchy be the sum of the deviations of its components. The minimal stratified hierarchy has the smallest deviation of all stratified hierarchies that are consistent with a given set of ERCs. Consider, for example, the following hierarchies and their statistics:

		Strata	Deviation
P>>Q>>R	$P + P\circ Q + P\circ Q\circ R$	3	3
P>>{Q,R}	$P + P\circ Q\circ R$	2	2
{P,Q}>>R	$P\circ Q + P\circ Q\circ R$	2	1
{P,Q,R}	$P\circ Q\circ R$	1	0

Deviation measures the distance from the Boolean ideal, which is one stratum with all relevant constraints unanimous in support of the desired optima. The MSH deviates minimally from this paragon.[21]

With this representation in hand, we may discern certain ranking relations that necessarily hold in the MSH between logically-related constraint forms. These will resemble (but not mirror) the relations found to hold among logically-related argument forms in §4.

Consider the effect of a constraint equivalent to P∘Q under some polyvaluation $\mathcal{A}$*, call it R, so that we have $\mathcal{A}^* \Vdash$ R=P∘Q. Two situations may be distinguished. First, the constraints P or Q may not appear in a stratum higher than |R|. In this case the stratum in which R=P∘Q first appears will look like this:

$$\ldots\ X\ + (\mathbf{R}\circ X\ldots\ ) + \ldots =$$
$$\ldots\ X\ + (\mathbf{P}\circ\mathbf{Q}\circ X\ldots\ ) + \ldots \qquad \text{, where } P,Q,R \notin X.$$

But this is indistinguishable from a hierarchy in which the first appearance of P and Q *individually* is in that very stratum.

In the second configuration, one of P or Q may dominate R; for concreteness, let's say P. Here we have

$$\ldots\ (P\circ X) + (P\circ\mathbf{R}\circ X)\ \ldots \quad =$$
$$\ldots\ (P\circ X) + (P\circ\mathbf{P}\circ\mathbf{Q}\circ X)\ \ldots \quad =$$
$$\ldots\ (P\circ X) + (\mathbf{P}\circ\mathbf{Q}\circ X)\ \ldots$$

This is indistinguishable from a hierarchy in which Q appears at R's stratum, in the place of R. Thus, when |P|≥|P∘Q|, we have P∘Q functioning as the equivalent of Q.

In the MSH, there can be no third way, in which |P∘Q| >|P|. Once P∘Q appears, P appears with it.

In other words, the fusion of a set of constraints is a greatest lower bound for that set within the MSH. It is also completely redundant: with P,Q in the hierarchy, P∘Q adds absolutely nothing. The general result can be stated this way:

•The fusion of a set of constraints has the same rank in the MSH as the lowest-ranked subset of those constraints. The hierarchy without the fusion is exactly equivalent to the hierarchy with it.

In proving this, it is useful to adapt some terminology:

According to the VS representation of a stratified hierarchy, as in (164), it is a fission of fusional expressions, with an overall $\oplus_k \mathsf{F}_k$ shape, where the fusional strata $\mathsf{F}_k = \otimes P_i$ are enumerated according to length, from shorter to longer. We write $\mathsf{F}_m \geq \mathsf{F}_n$ if F_m appears before F_n in this enumeration, *i.e.* if *m*<*n*; equivalently, if F_m is a subconstituent of F_n. The *rank* |P| of a constraint P is the *first* (equivalently, minimal) such F_k that it appears in.

[21] The deviation of the MSH can only be reduced by removing a fusional cluster, equivalent to removing a stratum and putting its content in the next lowest stratum, which we know is not possible, or by adding a prop letter to a fusional cluster that it is not a member of, which is equivalent to raising a constraint to a higher stratum. But constraints in the MSH are as high as they can go (Tesar & Smolensky 1998/2000).

(165) **Remark**. If a constraint P appears in some F_k, then we are guaranteed that $|P| \geq \mathsf{F}_k$. This holds in any hierarchy.

For a hierarchy H that validates a given ERC set, this means that the occurrence of a constraint $R = \otimes P_k$ in a stratum guarantees that H may be modified, salve veritate, so that any of the P_k ranked below R are ranked with it. The additional assumption of hierarchy minimality leads to a stronger result.

(166) **Proposition 8.1**. Let $\mathcal{H}(\mathcal{A})$ be a minimal stratified hierarchy containing all of $\mathsf{P} = \{P_i, 1 \leq i \leq n\}$, as well as Q, among possibly other constraints, where $Q = \otimes\mathsf{P}$. Then the following assertions hold:

(a) $|P_i| \geq |Q|$ for all i.
(b) $|Q| = |P_i|$ for at least one i.
(c) $\mathcal{H} \equiv \mathcal{H} - Q$.

Pf. Consider the VS representation of $\mathcal{H}$. Constraint Q first appears in some stratum F_k. Replace Q by $\otimes\mathsf{P}$ and we have an equivalent formula in which all of the P_i appear in F_k. By the remark (165), there is a hierarchy equivalent to $\mathcal{H}$. for which $|P_i| \geq \mathsf{F}_k$, for all i. The assumption of minimality for $\mathcal{H}$ entails that all constraints in it occupy their highest possible stratum, so $|P_i| \geq \mathsf{F}_k$ for $\mathcal{H}$ as well. This establishes (a).

As for (b) and (c), let F_k be the lowest rank of any $P_i \in \mathsf{P}$. Then F_k is also the first rank in which all of the P_i appear, *i.e.* it is the first rank in which Q may appear, therefore the rank of Q in $\mathcal{H}$, by minimality. This establishes (b). Now, if we replace Q by its equivalent $\otimes\mathsf{P}$, we merely double the occurrences of each P_i in that stratum. By idempotence, commutativity, and associativity of '$\circ$', this is equivalent to having a single occurrence of each of the P_i – that is, to having the same hierarchy sans Q. □

Proposition 8.1 can also be established by semantical considerations. A constraint is rankable (above the bottom rank) by RCD as soon as all of its L (truthvaluewise, F) coordinates have been removed. But each L coordinate of $\otimes\mathsf{P}$ arises where and only where some P_i has L at the same coordinate. Therefore by the time $\otimes\mathsf{P}$ is rankable, so is every P_i. But individual P_i's may become L-free without $\otimes\mathsf{P}$ becoming so. This establishes Prop. 8.1a. Now consider the set of lowest-ranked P_i's. At the point where they become rankable, there are no L's left among the coordinates of any P_i. Therefore $\otimes\mathsf{P}$ is rankable exactly then, as well (8.1b). Finally, note that if these lowest-ranked P_i's ever participate in the ratification of an ERC, all higher-ranked P_i's must take on the value *e*. Therefore, the value of $\otimes\mathsf{P}$ is exactly determined by these lowest-ranked P_i's and $\otimes\mathsf{P}$ makes no independent contribution to the calculation. Removed from the hierarchy, it changes no outcomes (8.1c).

The result makes good intuitive sense. A constraint $P \circ Q$ evaluated with respect to an ERC [*a*~*b*] is a repository of bad news from its components: it is as pessimistic about *a*'s chances of losing to *b* as the most pessimistic of P and Q. Since the MSH is constructed by collecting at each cycle of RCD the constraints most favorable to the desired outcome (Samek-Lodivici & Prince 1999), it is to be expected that the least favorable constraints will be ranked beneath their more favorable companions.

Proposition 8.1 shows how the logical composition of a constraint can directly affect both its ranking and its occlusion relations with other constraints that share logical components. The phenomenon is quite general, and a key to it lies in the transmission of polar values from atomic constraints to complex logical constructions that contain them. Let us say that a sentence G is *polar* in a certain variable P_j if G assumes a polar value whenever P_j does.

(167) **Def. Polar in a variable**. A formula G containing the prop letter P_j, among possibly others, is *polar in* P_j *over a set of valuations* V, if for every $v \in V$, $v(P_j) \in \{T,F\} \Rightarrow v(G) \in \{T,F\}$. Equivalently, by contraposition, if for every $v \in V$, $v(G)=e \Rightarrow v(P_j)=e$. If G is polar in all of its variables, we will simply say that G is *polar*.

The converse property, which we can call 'transparency', by which the assignment of *e* to all variables results in the assignment of *e* to the whole, is also worthy of note.

(168) **Def. Transparent.** A logical expression G is *transparent* in a three-valued system $\{T,F,e\}$ if $v(G)=e$ when $v(P_i)=e$ for all variables P_i in G. Equivalently, by contraposition, G is transparent if $v(G) \in \{T,F\} \Rightarrow v(P_i) \in \{T,F\}$ for all P_i in G.

Transparency is induced by all connectives we have been concerned with (though of course there are logics that lack it[22]). Polarity is more selective. All intensional formulas (those of S) are polar in all of their variables for any valuation, but other logical collocations may be polar only under certain valuations. For example, $P \wedge Q$ is not polar in P under RM3 rules when $v(P)=T$ and $v(Q)=e$, but it is polar in P in all other circumstances.

A polar formula occludes other constraints that are *transparently* made up of its variables.[23]

(169) **Proposition 8.2. Polar Occlusion.** Let G be a logical expression and let P be the set of prop letters used in G. Suppose that G is polar in all $P_i \in \mathsf{P}$. Then G occludes all constraints formulable as transparent expressions using only prop letters from P. If G is transparent, then G is occluded by any set of of polar expressions that collectively uses all the prop letters of P.

Pf. Suppose $|G| \geq |K|$ in some stratified hierarchy, where G is polar and K is some transparent logical combination of (some of) the variables of G. By transparency, if K assumes a polar value, then one of its prop letters assumes a polar value. By polarity of G, G will also assume a polar value. Now assume $\mathsf{K}=\{K_i\}$ is a set of formulas in which all the variables of G appear. Assume further that the K_i are polar in these variables. Say $|Ki| \geq |G|$, for all K_i. Suppose that G assumes a polar value. By transparency of G, at least one of its prop letters will assume a polar value. Since the K_i are assumed to be polar in the variables of G, at least one of them will assume of polar value. Since it is higher ranked or co-stratal with G, by assumption, G will be occluded. □

---

[22] E.g. those using Łukasiewicz's '→' or Bochvar's 'external connectives' (Rescher 1969:23,31).

[23] Occlusion results can also be stated order-theoretically. Proposition 8.2 guarantees that a polar formula G will occlude all formulas transparently made up from its prop letters. Suppose that K is such a set, and suppose that in a hierarchy H we have $|G| \geq |K_i|$ for all $K_i \in \mathsf{K}$. Then there is an equivalent hierarchy $H' = H - \mathsf{K}$ in which G is a lower bound for transparent formulas made up of prop letters from G.

A more fine-grained conclusion can be reached when we limit ourselves to intensional expressions. A logically-composite constraint G which is polar in P_j over an ERC set $\mathcal{A}$ has a potent effect on any intensional K that shares P_j, whenever $|G|\geq|K|$. (We do not presuppose that the relevant hierarchy is in MSH form.) In such a situation, K will behave as if it lacked P_j entirely in its logical composition. There is a kind of 'knockout' effect – the involvement of P_j in the constraint G eliminates the need for its involvement in subordinate or co-stratal constraints. We see this general effect in special form in Prop. 8.1, from which it follows, for example, that P>>P∘Q is equivalent to P>>Q: since P is polar (trivially) in P, it is effectively erasable from lower-ranked P∘Q.

More precisely, suppose G is polar in some prop letter P_j: then whenever $|G|\geq|K|$ in a hierarchy, K won't be needed to evaluate an ERC α unless $\alpha^*(G)=e$, i.e. unless $\alpha^*(P_j)=e$. Therefore, K may be replaced by any formula K' which is equivalent to K over all valuations v* for which $v^*(P)=e$. In the case of intensional formulae from the language of S, these are easy to construct by simple omission.

(170) **Def. G\\P_j.** If G is a formula of S containing a propositional variable P_j, then the formula G\P is arrived at by replacing binary constituents in G containing P_j, possibly repeatedly, according to the following recipe:

$$P_j \circ X \rightarrow X$$
$$\neg P_j \circ X \rightarrow X$$

until such replacement may no longer apply.

Since the formulas of S can be written using only '¬' and '∘', this covers all cases. Observe that in the case of the implicational connective, the result can involve more than erasure: for example, (X→P)\P is ¬X.

We may now state the general 'knockout' property, whereby a constraint disables part of a lower-ranked or co-stratal constraint.

(171) **Proposition 8.3. Knockout.** Let G be a constraint over some ERC set $\mathcal{A}$, where G is a logical expression that is polar in P. Suppose K, a constraint over $\mathcal{A}$, is an expression of S with P among its variables. Then any hierarchy H in which $|G|\geq|K|$, with $H\vDash\mathcal{A}$, is equivalent over $\mathcal{A}$ to one in which K is replaced by K\P, or, if K = P or K= ¬P, omitted entirely.

> *Pf.* K is clearly dispensable in the evaluation of ERC α when $\alpha^*(P)\in\{T,F\}$; in that case, it doesn't matter what K is replaced by. K can only have an effect when $\alpha^*(G)=e$. Since G is intensional and hence polar in P, we have $\alpha^*(P)=e$. In this case, if K contains more prop letters than just P, $\alpha^*(K)=\alpha^*(K\backslash P)$, by the definition of K\P. If K only contains P, then $\alpha^*(K)=e$ and K has no effect on the valuation of α. □

The Knockout property allows us to derive Proposition 8.1c from 8.1a. If Q=⊗P, then by 8.1a, $|P_i|\geq|Q|$ in the MSH. But each P_i is polar in itself, so all may be removed from Q, establishing that $\mathcal{H}=\mathcal{H}-Q$ for any MSH $\mathcal{H}$ containing Q and its variables. Similarly, if |⊗P|≥|P| in some hierarchy (necessarily non-minimal if |⊗P|>|P|), then P may be dispensed with, because ⊗P is polar in each of the P_i.

Occlusion *presupposes* a dominance relation; but logical structure may also impose limitations on ranking, as is seen in the special case of fusion in Proposition 8.1. A bridge between occlusion and ranking is provided by Grimshaw's observation that an occluded constraint is freely rankable with any other other constraint dominated by the occluder.

(172) **Proposition 8.4.Occlusion and Ranking.** (Grimshaw). Suppose C_1 occludes C_2 over $\mathcal{A}$. Let H, H′ be stratified hierarchies, in which $|C_1|\geq|C_2|$ and which otherwise agree on all ranking relations except those involving C_2. Then $H\vDash\mathcal{A}$ iff $H'\vDash\mathcal{A}$.

> *Pf.* Since C_1 occludes C_2, in any hierarchy H_i with $|C_1|\geq|C_2|$, we have $H_i=H_i-C_2$. Therefore $H=H-C_2$ and $H'=H'-C_2$. But $H-C_2$ is identically $H'-C_2$. □

In other words, when C_1 occludes and dominates C_2, we may rank C_2 anywhere below C_1. This has consequences for the structure of MSH: there $|C_2|$ can lie no further down the hierarchy than the immediately adjacent lower stratum.

(173) **Corollary 1 to Prop. 8.4.** Let C_1 occlude C_2 over $\mathcal{A}$, with $|C_1|\geq|C_2|$ in the MSH for $\mathcal{A}$. Then there is no constraint D with $|C_1|>|D|>|C_2|$ in $\mathcal{H}(\mathcal{A})$.

> *Pf.* By Prop. 8.4, C_2 can be ranked anywhere below C_1. Since a constraint assumes its highest possible rank in the MSH, C_2 will be co-stratal with C_1 if that is possible, else in the next rank down. □

We may also draw a general conclusion about the ranking behavior of intensional formulas with respect to each other: when variable-sharing results in occlusion, Corollary 1 will apply..

(174) **Corollary 2 to Proposition 8.4**. Let K be a constraint formulable in the language of S and let $\mathsf{G}=\{G_i\}$ be a set of such constraints, where the prop letters of K are included among those of G. Then if $|G_i|\geq|K|$ for every $G_i\in\mathsf{G}$, K will be ranked in any K,G-containing MSH in a stratum that is no lower than the stratum immediately below the lowest ranked G_i.

> *Pf.* Since the G_i are polar in all the variables of K, they collectively occlude K. and Corollary 1 to Proposition 8.4 guarantees the result. □

Corollary 2 (174) contrasts with Proposition 8.1, which gives a tighter result: in the MSH relation between a fusional constraint and its prop letters, we know that the fusion lies no lower than the lowest ranked prop letter, indeed lies with it. The more complex condition on (174) puts the occluded constraint at most one rank further down than its occluders. Situations where the next lower rank is needed are easily found. Consider the following relation between P+Q and P∘Q, when $\alpha^*(P)=T$, $\alpha^*(Q)=F$.

	P+Q	P∘Q
α	W	L

Here P+Q includes all the prop letters of P∘Q, and occludes P∘Q, yet must be ranked above it. This effect does not go away when we require the presence of the prop letters in the constraint set:

	P+Q	P∘Q	P	Q
α	W	L	W	L
β	W	L	L	W

If we wish to tighten things up so as to put the occluded formula exactly in the lowest stratum occupied by its occluders, we can require (first) that the prop letters of K both be present in the hierarchy and *not* be dominated by K. And (second) that K be a 'positive' formula – lacking negation, and composed only through fission and fusion (or possibly conjunction and disjunction, if we extend to RM3). Any such positive sentence has the valuable property that it takes the value L only when one of its letters is L.

(175) **Remark**. Let G be a positive sentence, *i.e.* one containing only $\circ$,+, $\wedge$, $\vee$. Then v(G)=F only if there is a prop letter P in G such that v(P)=F, where v is an RM3 valuation.

Pf. The assertion holds by basic RM3 definitions if there is only one such connective in the formula. Suppose it holds up to n connectives; any formula with n+1 connectives will have the form A *op* B, where A, B have n or fewer connectives. If v(A *op* B)=F, then by RM3 definitions we have v(A)=F or v(B)=F. Apply the induction hypothesis to the one that evaluates to F and we're done. □

A quick lemma expanding the purview of Proposition 8.1 sets us on our way.

(176) **Lemma. Fusion as Lower Bound.** Let let G=$\{G_i\}$ be a set of positive expressions constructed in the language of S from prop letters P=$\{P_i\}$. Let G and $\otimes$P belong to an MSH H. Then $\otimes$P is a lower bound for G in H, i.e. for every $G_i \in$G, $|G_i| \geq |\otimes$P$|$.

Pf. Suppose $|\otimes$P$| > |G_i|$, for some G_i. Then H must be consistent with an ERC α in which $\alpha^*(\otimes|\otimes$P$|)$=T and $\alpha^*(G_i)$=F and where $\alpha^*(|C|)=e$ for all strata $|C| > |\otimes$P$|$, i.e.an ERC α that embodies the argument for ranking G_i below the stratum of $\otimes$P, which says that the fusion of P's stratum, namely $\otimes|\otimes$P$|$, must conflict with G_i. If H is not consistent with α, then nothing restricts G_i to a stratum below $|\otimes$P$|$. However, if α*(Gi)=F, then by the remark some P_j in G_i must have $\alpha^*(P_j)$=F, since G_i is positive intensional. Then $\alpha^*(\otimes$P$)$=F, contrary to assumption. So no such α can exist. This establishes that $\forall G_i \in$G, $|G_i| \geq |\otimes$P$|$. □

Under the assumption that all prop letters involved in the constraint under discussion are independently present in the hierarchy, we can now get our desired result:

(177) **Remark.** Let K be a positive formula, and let P=$\{P_i\}$ be the set of prop letters in K. Let $\mathcal{H}$ be an MSH containing K and P, and suppose further that $|P_i| \geq$K for every $P_i \in$P. Let $P_\perp$ denote the lowest ranked of the P_i. Then $|K| = |P_\perp|$.

Pf. By the lemma, $\otimes$P is a lower bound for K. In the VS representation of $\mathcal{H}$, $|P_\perp|$ contains $\otimes$P. Therefore $|K| \geq |P_\perp|$. Since $|P_\perp| \geq |K|$ by assumption, we have $|K| = |P_\perp|$. □

To illustrate the range of effects just noted, let us examine the ranking properties of the fissile constraint $\oplus$P. These are more various than those of its fusional cognate $\otimes$P, even though the occlusional properties are the same.

Unlike $\otimes$P, a constraint $\oplus$P can sit in a variety of ranking positions with respect to its components in the MSH. Furthermore, the presence of $\oplus$P in a hierarchy in addition to its component prop letters can radically alter the range of outcomes. A couple of examples will make this clear. Imagine a set of constraints C_1, C_2, which impose orders on candidates {a,b,z}, as follows.

(178)

$$\begin{array}{cc} C_1^{>} & C_2^{>} \\ a & b \\ | & | \\ z & z \\ | & | \\ b & a \end{array}$$

Example (178) is the kind of system studied in Samek-Lodovici & Prince 1999 and discussed in §5 above. Cndidate *z* is collectively bounded by {*a,b*} and no ranking can make it optimal. In ERC representation, the failure of *z* is quite apparent:

(179) Collective Bounding

	C_1	C_2
[z~a]	L	W
[z~b]	W	L
[z~a] ∘[z~b]	L	L

Both columns fuse to L, so no ranking satisfies both ERCs. The only optima from {*a*,*b*,*z*} are *a* and *b*. With $C_3 = C_1{+}C_2$ added to the mix, however, candidate *z*'s situation improves dramatically.

(180) $C_1{+}C_2 >> C_1, C_2$

	$C_1{+}C_2$	C_1	C_2
[z~a]	W	L	W
[z~b]	W	W	L
[z~a] ∘[z~b]	W	L	L

Now *z* becomes optimal under the ranking $C_3 >> \{C_1, C_2\}$.

The fission of a set of constraints may also be ranked *below* some of its components, though in this case the intervention of another constraint is required. Here's an example.

(181) C_1 >>Q>> $\{C_2, C_1{+}C_2\}$

	C_1	**Q**	C_2	$C_1{+}C_2$
α		W	L	L
β	W	L	…	W

The general situation can be portrayed graphically like this. For $\mathsf{P}=\{P_i, 1 \leq i \leq n\}$ arrayed in possibly many strata in an MSH:

(182)

$$\ldots\ |P_1 \ldots (\ \ldots\ |\ \ldots\ )\ \ldots\ P_n|\ \ldots$$

$$\oplus\mathsf{P}$$

The constraint $\oplus\mathsf{P}$ can, grossly speaking, fit in anywhere from just above the highest ranked of the P_i 's to a position co-stratal with the lowest ranked. Under the assumption that all the prop letters of $\oplus\mathsf{P}$ are present with it in the hierarchy, the ranking properties of $\oplus\mathsf{P}$ can be laid out as follows:

(183) **Ranking Properties of $\oplus\mathsf{P}$**

I. $|\oplus\mathsf{P}| \geq |P_i|$.

(a) $\oplus\mathsf{P}$ may crucially dominate a subset of P with two or more members.

• Example (180) shows how this can happen. Why two or more? Suppose $\oplus\mathsf{P}$ dominates only one of its prop letters, call it $P_\perp$. Then all the others either dominate or are co-stratal with $\oplus\mathsf{P}$. Call these P'. By **Knockout** (171), Prop. 8.3, $\oplus\mathsf{P}$ will function like $\oplus\mathsf{P}\backslash\mathsf{P}'$. But this is just $P_\perp$ and self-domination cannot be crucial.

(b) In the MSH, any $P_i \in \mathsf{P}$ that is crucially dominated by $\oplus\mathsf{P}$ falls into the stratum *immediately below* that of $\oplus\mathsf{P}$. Any such $\oplus\mathsf{P}$-dominated P_i's contribute nothing to operation of the hierarchy and can be removed without changing its optima.

• The second clause follows from Prop. 8.2, **Polar Occlusion**, since $\oplus\mathsf{P}$ occludes its variables, $\oplus\mathsf{P}$ being polar. The first clause then follows from Corollary 1 to Prop. 8.4, **Occlusion and Ranking**.

II. $|P_i| \geq |\oplus\mathsf{P}|$

(a) $\oplus\mathsf{P}$ may itself be crucially dominated by a $P_i \in \mathsf{P}$, but only through transitivity of domination involving the agency of non-member(s) of P .

• Since $\alpha^*(\oplus\mathsf{P})$=T whenever $\alpha^*(P_i)$=T, the two constraints $\oplus\mathsf{P}$ and P_i cannot directly conflict: no α yields $\alpha^*(P_i)$=T and $\alpha^*(\oplus\mathsf{P})$=F. Crucial domination is possible by a transitivity argument, as illustrated in example (181).

(b) In the MSH, if $|P_i| \geq |\oplus\mathsf{P}|$ for *all* $P_i \in \mathsf{P}$, then $|\oplus\mathsf{P}|$ occupies the *same* stratum as the lowest-ranked P_i.

• Since $\oplus\mathsf{P}$ is positive, this follows directly from **Remark** (177).

(c) When $|P_i| \geq \oplus\mathsf{P}$ for some certain $P_i \in \mathsf{P}$ in any hierarchy, then $\oplus\mathsf{P}$ functions the same as $\oplus(\mathsf{P} - \{P_i\})$. Hence, in condition II(b), where this relation holds for all P_i, $\oplus\mathsf{P}$ contributes nothing to the operation of the the hierarchy.

• This is **Knockout**, obtained because the P_i are trivially polar and $\oplus\mathsf{P}$ transparent.

To this point we have considered logical relations based on syntactic overlap between constraints. Of equal interest is the logical relation we began with: *entailment*. Suppose P→Q holds under some polyvaluation $\mathcal{A}$*. We find excellent algebraic behavior in the MSH: we are guaranteed that |P|≤|Q| and, further, that P is occluded in hierarchies that validate $\mathcal{A}$.

Proposition 8.5. Let P,Q be constraints in the MSH $\mathcal{H}(\mathcal{A})$. If $\mathcal{A}^* \Vdash$ P→Q, then |Q|≥|P| Furthermore, if $\mathcal{H} \vDash \mathcal{A}$., then $\mathcal{H} = \mathcal{H} - \text{P}$.

Pf. Say that |P|>|Q| in $\mathcal{H}(\mathcal{A})$. Then, by Remark (156) on stratal conflict, there must be an $\alpha \in \mathcal{A}$ with α*(Q)=L. Since $\mathcal{A}^* \Vdash$ P→Q, it must be that α*(P)=L and consequently α*(⊗|P|)=L, an impossibility in $\mathcal{H}(\mathcal{A})$. So |P|≤|Q|. Now suppose $\mathcal{H}$ validates $\mathcal{A}$. We claim P is occluded. Because $\mathcal{A}^* \Vdash$P→Q, the only valuation β* for which β*(Q)=*e* and β*(P)∈{W,L} requires β*(P)=L. Thus, Q occludes P except on this valuation. But P cannot be the highest-ranked issuing a polar valuation of β, else β is unsatisfied, contrary to the assumption that $\mathcal{H} \vDash \mathcal{A}$. Therefore, some other constraint C occludes P on β* in $\mathcal{H}$. Since P is occluded on all valuations, $\mathcal{H} = \mathcal{H} - \text{P}$. □

Remark. This shows that occlusion in *successful* hierarchies requires only Occluded=W ⇒ Occluder∈{W/L}.

The converse of Proposition 8.5 is of course not true: |P|≤|Q| does not force P→Q. (Obvious, perhaps, because subordination is not occlusion, which is a concomitant of P→Q in a hierarchy that validates its ERC set .) For example, with no other constraints in sight, the column vector $\mathcal{A}$*(Q)=(W,*e*,) must dominate $\mathcal{A}$*(P)=(L,W), but '→' fails in the second coordinate (going from subordinate P to dominant Q).

	Q	P
α	W	L
β	*e*	W

Proposition 8.5 reveals that '→' over constraints construed as VS vectors works in a way that is opposite to naive Boolean expectations. When one constraint entails another in a Boolean system, it is the entailed constraint that is superfluous. Here it is the entailer that plays no role.[24] Within the current terms, however, the result makes good sense. For an ERC [*a*~*b*], the expression 'P→Q' means that *a* does no worse against *b* on Q that it does on P (where earning *e* counts as doing worse than W). Since the sense of the MSH is, broadly, that constraints more favorable to the desired winners are ranked higher than less favorable constraints, it is unsurprising that the at-least-as-favorable Q should be at least as high-ranked as P.

[24] This disparity would be rectified if we switched the meaning of T and F; after all, assignment of the polar values to poles of the 'better than' relation is arbitrary. Numerous architectural contortions would ensue, however, and we leave the matter for future contemplation.

By contraposition, whenever we have P→Q, we also we have (¬Q)→(¬P), and therefore, by Proposition 8.5, we have |¬Q|≤|¬P|, a fact that can be put to use in various ways; for example, by considering the hierarchy obtained by replacing P and Q with their negations. It is reasonable, then, to put aside implication, and ask how general the relation is between negation and ranking inversion. It turns out that it *cannot* be generally guaranteed in the MSH. Here's an example showing that the very opposite can hold:

(184) **Preservation of relative rank under negation in MSH**

|P| > |Q|	I	II		III
	C_1	**P**	C_2	**Q**
α	W	**L**	L	**W**
β			W	L

|¬P|> |¬Q|	I		II	
	C_1	**¬P**	C_2	**¬Q**
α′	W	**W**	L	**L**
β′			W	W

Notice that we have neither implication nor crucial domination between P and Q. P's rank is as high as possible, this being the MSH , but a (nonminimal) stratified hierarchy in which P joined Q in rank III would ratify the same ERCs.

When there is *crucial* domination between two constraints, however, ranking contraposition does hold. The notion of 'crucial domination' is subject to some ambiguity of use, depending on whether the focus is on a specific hierarchy or on the class of hierarchies consistent with the ERC set; let us disambiguate.

Consider the case of constraints defined by an ERC vector (W,L,W). A hierarchy with domination order C_1>>C_2>>C_3 will satisfy the ERC, and the relationship C_1>>C_2 might be called 'locally crucial' with respect to that hierarchy inasmuch as C_1 cannot be shifted to a lower position, while retaining the rest of the ranking relations. But it is also *non*crucial inasmuch as the work it does can be replaced by C_3, yielding an equally satisfactory hierarchy C_3>>C_2>>C_1, which denies the 'crucial' C_1>>C_2. By contrast the relation C_1>>C_2 over the sole ERC α=(W,L,*e*) is crucial in the stronger sense that *any* hierarchy consistent with α will require it. It is this latter, global sense of 'crucial' that we are interested in. In this sense, a ranking relation is crucial when it is entailed by ERC set.

(185) **Def. Crucial domination.** Let P,Q be constraints over a consistent ERC set $\mathcal{A}$. Let us say that P>>Q is *crucial* iff P>>Q is entailed by $\mathcal{A}$, equivalently iff Q>>P is inconsistent with $\mathcal{A}$.

Under this definition, P>>Q will be present in every hierarchy H with $H \vDash \mathcal{A}$. We can now proceed to demonstrate that P>>Q, with domination crucial, implies ¬P>> ¬Q. The only complexities in the proof are superficial and due to the necessity of keeping track of the parallels and anti-parallels in the two hierarchies being compared.

(186) **Proposition 8.6**. Assume constraints P,Q in a constraint set S over a consistent $\mathcal{A}$, with P>>Q crucial. Construct a modified constraint set S′ by replacing P,Q in S with ¬P, ¬Q. Let $\mathcal{A}'$ be the ERC set resulting from this modification and let H′ be any hierarchy on S′ validating $\mathcal{A}'$. If such H′ exists, then (¬Q)>>(¬P) crucially in H′.

Pf. Consider the ERC φ, not necessarily a member of $\mathcal{A}$, in which φ*(P)=W and φ*(Q)=L and φ*(C)=*e* for all other C∈S. Since φ says precisely P>>Q, we have $\mathcal{A} \vdash \varphi$. By Proposition 2.5, (25) p. 14, we must have some subset Ψ of $\mathcal{A}$ with $f\Psi \vdash \varphi$.

Let us examine the structure of Ψ. Note first that we must have $\mathcal{A} \Vdash f\Psi \rightarrow \varphi$, because both φ and $f\Psi$ are nontrivial. This gives us a tight hold on the coordinatewise behavior of the antecedent in the implication. Briefly, $f\Psi$ must have W at P, L at Q, and L or *e* elsewhere, with φ derived by L-retraction, at most.

(i) For all constraints C∈S, with C≠P,Q, we have φ*(C)=*e*, by construction. Therefore, since $C \Vdash f\Psi \rightarrow \varphi$ for each C, we must have $f\Psi$*(C)∈{L, *e*}.

(ii) Since φ*(Q)=L, we must have $f\Psi$*(Q)=L.

(iii) We must therefore have $f\Psi$*(P)=W, else $f\Psi \in \mathcal{L}_+$, contrary to the assumption that $\mathcal{A}$ is consistent.

Let $\mathcal{A}'$ be the ERC set corresponding to S′, in which P is replaced by P′ = ¬P and Q is replaced by Q′ = ¬Q. Observe that this transforms φ into φ′=(−φ). Let Ψ′ be like Ψ and except that in each ψ∈Ψ the values ψ*(¬P) and ψ*(¬Q) are substituted for those of ψ*(P), ψ*(Q), yielding ψ′∈Ψ′. Claim: $\mathcal{A}' \Vdash f\Psi' \rightarrow (-\varphi)$. Briefly, $f\Psi'$ will have L at P, W at Q, and will remaining the same as $f\Psi$ at all other coordinates, so these will be L or *e*. Once again, L-retraction will do the job.

(i′) For C∈S, with C≠P,Q, we have C∈S′ and everything remains the same. From (i), it follows that $C \Vdash f\Psi' \rightarrow (-\varphi)$, since (−φ)*(C)=φ*(C)=*e* and ($f\Psi'$)*(C)=($f\Psi$)*(C)∈{L,*e*}. Therefore, the interesting action is in the coordinates P′ and Q′, corresponding to P and Q.

(ii′) By construction, we have (φ′)*(¬P)=L. From (iii), we have ($f\Psi$)*(P)=W. This means that for some ψ∈Ψ, ψ*(P)=W and therefore (ψ′)*(¬P)=L. Therefore($f\Psi'$)*(P′) = L.

(iii′) By construction, (φ′)*(¬Q)=W. (We also note that ($f\Psi'$)*(¬Q)=W else $f\Psi'$ is inconsistent, which would lead to inconsistency in $\mathcal{A}'$, contrary to assumption.)

From these remarks, it follows that $\mathcal{A}' \Vdash f\Psi' \rightarrow (-\varphi)$, and consequently $\mathcal{A}' \vdash (-\varphi)$. But (−φ) means exactly Q>>P. □

Implication and crucial ranking do not exhaust the conditions under which ranking contraposition can be found. Here's a case where neither is at issue, but in which ranking does contrapose under negation.

(187) **Ranking contraposition without implication or crucial domination**

$\|P\| > \|Q\|$	I	II		III
	C_1	**P**	C_2	**Q**
φ	W	L	L	W
ψ			W	L
ξ		W	W	

$\|\neg Q\| > \|\neg P\|$	I	II		III
	C_1	**¬Q**	C_2	**¬P**
φ′	W	L	L	W
ψ′		W	W	
ξ′			W	L

We have seen that logical relations between constraints, represented in VS, correlate with ranking relations, in ways that are parallel – or anti-parallel – to what we found for ERCs. This recalls to mind the hopes of some that Paninian special-general relations, which typically involve entailment in some form, would also limit domination patterns. This is manifestly untrue for constraints construed as assessors of violation (Prince 1997 et seq.), where special-case/general-case constraints are crucially rankable, via transitivity, in any order. But by tranferring operations from violation space to comparison space, we have located the setting in which logic and ranking interact systematically.

To conclude this survey of constraint logic, let us switch perspective and think in terms of sets of candidates, in order to ascertain what fusion means for candidate-set structure. Each constraint C_i imposes a stratified partial order $C_i^>$ on a candidate set. For example, we could have the following structure imposed over a candidate set {a,b,c,d,e,f,g,z.}, represented in a Hasse-type diagram, with better being higher:

$C_1^>$
a
|
b
|
c,d,z
|
e
|
f

From the point of view of any individual candidate z, the candidate set falls orderwise into three equivalence classes on each constraint C_i:

L_i the set of candidates to which z loses. For $\lambda \in Li$, $(\lambda > z; C_i)$.
E_i the set of candidates which C_i does not distinguish from z. $\forall \xi \in E_i\ (\xi \approx z; C_i)$.
W_i the set of candidates which z beats on C_i. $\forall \omega \in W_i(\ z > \omega; C_i)$

Recast in familiar terms, this description puts into the set L_i all those candidates λ such that $C_i([z{\sim}\lambda])=L$; into E_i all those candidates ξ such that $C_i([z{\sim}\lambda])=e$; intoW_i all those candidates ω such that $C_i([z{\sim}\omega])=W$.

For the triple $\langle L_i, E_i, W_i \rangle$, let's write $C_i^{>}/z$. Our example transforms as follows:

	$C_1^{>}/z$
L:	a,b
	|
E:	c,d,z
	|
W:	e, f

Given two such constraint-orders $C_i^{>}/z$ and $C_j^{>}/z$, their fusion can be constructed in a way that accords completely with previous usage:

(188) $(C_i^{>}/z) \circ (C_j^{>}/z) = \langle\ L_i \cup L_j,\ E_i \cap E_j\ ,\ (W_i \cup W_j) - (L_i \cup L_j)\ \rangle$

The first component consists of those candidates that beat z on at least one of the constraints; since L is dominant in '$\circ$', these must also beat z on the fusion ($L \circ L = L \circ W = L \circ e = L$). The last component consists of those candidates that z beats on at least one of the constraints but that do not beat z on either ($W \circ W = W \circ e = W$). The middle component includes all those candidates that are orderwise indistinguishable from z on *both* constraints (this includes, of course, z itself), falling under the rubric $e \circ e = e$. This immediately leads to the general expression:

(189) $\otimes(C_i^{>}/z) = \langle\ \bigcup L_i, \bigcap E_i, \bigcup W_i - \bigcup L_i \rangle$.

For a concrete sense of how this plays out, consider the following example.

	$C_1^{>}/z$	$C_2^{>}/z$	$(C_1^{>}/z) \circ (C_2^{>}/z)$
L:	**a**,**b**	**a**,**e**	**a**,**b**,**e**
	|	|	|
E:	$\underline{c}$,*d*,$\underline{z}$	**b**,$\underline{c}$,$\underline{z}$	$\underline{c}$,$\underline{z}$
	|	|	|
W:	**e**, *f*	*d*,*f*	*d*,*f*

Suppose we have a hierarchy H in which z is optimal in its candidate set K. Consider the fate of K as we proceed from constraint to constraint down the hierarchy. Construe a constraint as a function that returns its top stratum, with ranking as functional composition: then at each point in the hierarchy only the top stratum is passed on to the next-ranked constraint. The candidate set shrinks around the optima as we descend the hierarchy.

A constraint $\langle L_i, E_i, W_i \rangle$ becomes rankable at the point in RCD where L_i becomes null: where all the candidates in L_i have been eliminated by virtues of their membership in some higher-ranking W_k. (See Samek-Lodovici & Prince 1999 for further development along these general lines). It is clear from expression (189) that when $\otimes C_i$ is rankable under RCD, by virtue of the emptiness of $\bigcup L_i$, all of its component constraints C_i have become rankable as well. This replicates the semantical argument for Prop. 8.3a; the others proceed along similar lines.

A further consequence of algebraic interest follows directly: fusion preserves ranking relations in the MSH. Grossly put, if $|A| \geq |B|$, then $|A \circ X| \geq |B \circ X|$.

(190) **Proposition 8.7**. Let $\mathcal{H}$ be a hierarchy in MSH form containing constraints $\mathbb{C} \cup \{A,B\}$, where $\mathbb{C}$ is a set of constraints, possibly null, disjoint from $\{A,B\}$. Let $\mathcal{H}'$ be a hierarchy in MSH form that contains $\mathbb{C}$ as well as $A \circ X$ and $B \circ X$, for some constraint X, and possibly contains A,B as well.

If $|A| \geq |B|$ in $\mathcal{H}$, then $|A \circ X| \geq |B \circ X|$ in $\mathcal{H}'$.

Pf. Writing $L(A_i)$ for the L component of $A_i = \langle L_i, E_i, W_i \rangle$, we have

(*) $L(A \circ X) = L(A) \cup L(X)$

(**) $L(B \circ X) = L(B) \cup L(X)$.

Let D be the set of constraints in $\mathcal{H}$ dominating A. Since A does not lie among those strata of $\mathcal{H}$ containing members of D, it must be the case that either

(i) A contains no L's and D is empty, or

(ii) The ERC set underlying $\mathcal{H}$ is inconsistent and has no consistent subsets, and therefore $\mathcal{H}$ has only one stratum, with D empty, or

(iii) the distribution of L's in A is such as to lead to domination of A by the constraints in a nonempty D.

Now consider the role of D in $\mathcal{H}'$. Let Σ_D be the set of strata in $\mathcal{H}'$ which contain the members of D. It cannot be that $A \circ X$ lies in Σ_D. This is trivial for cases (i) and (ii), since D is empty. For case (iii) note that $A \circ X$ has all the L's of A and possibly more, as is clear from (*). Similarly for $B \circ X$. If $B \circ X$ were positioned in one of these strata, B would also be positionable among the constraints of D in $\mathcal{H}$: since all the L's of $B \circ X$ would be eliminated from leading position in the hierarchy by constraints in D, so would those of B, by (**).

Now suppose $|B \circ X| > |A \circ X|$ in $\mathcal{H}'$. At the first stratum occupied by $B \circ X$, namely $|B \circ X|$, both L(X) and L(B) must have been eliminated. But since $|B \circ X|$ lies strictly below the strata in Σ_D, L(A) is also eliminated. But then $A \circ X$ must be rankable in $|B \circ X|$ or higher, contrary to assumption. □

The reader might wish to show that the same relation holds for ERCs.

9. The Arithmetic of Optimality Theory

Courage. Weren't strips of heart culture seen
Of late mating two periodicities?
– Empson

Summary. The arithmetic of comparison in Optimality Theory is strictly finitistic.

Optimality Theory can be understood as an extreme case of weighting constraints (Prince & Smolensky 1993, ch. 10), with minimization of the summed weights.

To determine optimality arithmetically by a weighted sum of violations, we might for example assign a numerical score to each candidate: say, $3\cdot C_1+C_2$, with 1 point assessed for each failure. The winner would be the candidate with lowest score, 0 being the best possible. In this system, violations of C_1 count more than violations of C_2 – each C_1 violation weighs in at 3 units, against 1 for violations of C_2. This will imitate optimality theoretic relations over a limited region of violation space, but it lacks the crucial property of strict domination, which forces one violation of a higher-ranked constraint to be worse than *any number* of violations of lower-ranked constraint. Consider the following violation pattern:

(191)

	C_1	C_2
A		****
B	*	

Under OT, the optimum is A, because it bests B on the highest -ranking constraint that distinguishes A and B. But under the weighted-sum model $3\cdot C_1+C_2$, it will be B that obtains the minimal score. Weighted-sum theories equilibrate violations, allowing kinds of trade-off that OT can disallow.

To ensure the strict-domination effect, we'd need to use a formula of the shape $N\cdot C_1+C_2$, where N is strictly greater than the greatest possible number of violations of C_2 incurred by any form. (For example, if 4 is the max, then $5\cdot C_1+C_2$ will work.) Since this relation must hold between every adjacent pair of (crucially interacting) constraints in the hierarchy, we quickly get a multiplicative scheme. If $M\cdot C_2+C_3$ works for the C_2/C_3 relation, then $M\cdot N\cdot C_1+M\cdot C_2+C_3$ will work for all three constraints.[25] If M>N, we can be conservatively sure that $M^2\cdot C_1+ M\cdot C_2+C_3$ will do the job. In general, a successful weighting strategy will surely be obtained if we take as its *base* a number greater than the greatest possible number of violations of any constraint. Then we score violations by the following exponential scheme, starting the sequence with C_0 for convenience and numbering upwards in domination order:

[25] Suppose $C_1(A)=0$ and $C_1(B)=1$. For A to prevail, its score must always remain less than that of B. Now let $C_2(B)=C_3(B)=0$. The worst that A can do, maxing out on the lower ranked constraints, is $C_2(A)=N-1$, $C_3(A)=M-1$. The score assigned to A is therefore $M(N-1)+M-1 = MN-1$. Therefore, the C_1 weight MN is required to push B's single violation over the limit.

(192)

$$|\alpha| = \sum_{k=0}^{n} M^{n-k} C_k(\alpha)$$

But there is typically no principled limit to the number of violations that a given constraint may assess, just as there is no principled limit to the length of a linguistic form. Therefore, to arithmetize OT it would appear that we need to appeal to infinite weights, using the principles of nonstandard arithmetic to calculate with them. The spectre of the infinite has been known to induce discomfiture.

Courage. Calculation throughout this paper has been in terms of only three values: W,L,*e*. We have seen that a logical expression taking values equivalent to these can represent the notion of a strict domination hierarchy, linearly ordered (161) or stratified. (164). Robert Wilson (p.c.) points out that the constituent operators '$\circ$' and '+' can be represented as bivariate polynomials taking values in the field F_3, whose three elements may be understood as the integers mod 3, e.g. as numbers from $\{1,0,-1\}$, with the assignment W=1, e=0, L=-1. From this observation, we can deduce immediately that the logical formulas for domination hierarchies will translate into polynomials taking values from $\{1,0,-1\}$. For P>>Q>>R, for example, we will have the following representation, using lower case for the value assigned to upper case constraints:

(193) **Scoring P>>Q>>R**

$$p + q\,(1-p^2) + r\,(1-p^2)\,(1-q^2)$$

The expressions involving squared variables will detect polar values: $(1-p^2)$ is 0 whenever p=1 or p=-1, but 1 when p=0. Therefore, if p is polar, all terms beyond the first go to 0 and the value of the expression is determined entirely by p. If p=0, then q assumes the determining role; when q is polar, all subsequent terms vanish. And so on. This exactly mirrors the process of evaluation of a strict-domination hierarchy.

The general expression can be written like this, enumerating the constraints in domination order, starting with top-ranked P_1. For convenience, we put p_0=0.

(194)

$$\sum_{k=1}^{n} p_k \cdot \prod_{i=0}^{k-1} (1-p_i^2)$$

A (linearly ordered) hierarchy P_1>>...>>P_n is successful if this expression evaluates to 1 or 0 over every ERC, with 0 occurring only for the degenerate ERC.

The key, of course, is that OT is really only interested in the *difference* in degree of violation between the optimum and its competitors, and then only in the *sign* of the difference. To do arithmetic sensibly, you subtract first, and normalize weights to $\{1,0,-1\}$. In a scheme of infinite weights, all but the greatest really function like 0; here we just use 0. OT arithmetic is then finitistic in the rather extreme sense that it need use integers no larger than 1, no smaller than -1.

Appendix 1. Functional Characterization of Constraints

Let U be the universal set of forms and let an 'OT choice(OTC) function' C be a function $C:\wp(U)\rightarrow\wp(U)$ that meets the following restrictions.

(i) Choice. $C(X)\subseteq X$ for all $X\subseteq U$.
(ii) Forced Choice. $C(X)=\emptyset \Rightarrow X=\emptyset$.
(iii) Contextual Independence of Choice. If $Y\cap C(X) \neq\emptyset$, then $C(Y\cap X)=Y\cap C(X)$.

We show any such function C induces a stratified partial order on each $X\subseteq U$; that C is the 'maximum' function with respect to this order; and conversely, that the maximum function on any stratified partial order defined on U satisfies these three conditions. In addition, we show that a constraint hierarchy, a functional composition of OTC functions, is itself an OTC function.

By a strict partial order, we mean one that is asymmetric, irreflexive, and intransitive. By a stratified order, we mean one in which noncomparability '$\|$' is an equivalence relation. Two elements x,y are noncomparable, $x\|y$, iff neither $x>y$ nor $y>x$.

To avoid a befuddling proliferation of parentheses, commas, and braces, I will use the following aggressive abbreviations:

for	C(X) write	CX
	{x,y,z}	xyz
	{x}	x (with distinction between x and {x} contextually determined).

Given such a C meeting the above description, we define the following order:

Def. Order induced by C. For every $x,y \in U$ such that $x\neq y$, $x>_C y$ iff $Cxy=x$.

Def. max. $x\in\max(X)$ iff $x\in X$ and $\neg\exists y\in X \ni y>x$.

Remark. $x\|y$ iff $Cxy=xy$.

Proposition 1. For any C meeting conditions (i),(ii), (iii), $>_C$ is a stratifed partial order on any $X\subseteq U$.

Proof. First, we show that $>_C$ is a strict partial order: that it is asymmetric, irreflexive, and transitive. Asymmetry is built into the definition, as is irreflexivity. To show transitivity, assume

(a) $Cxy=x$
(b) $Cyz=y$.

We claim that $Cxz=x$.

Suppose $yz\cap Cxyz\neq\emptyset$. Then by (iii), $C(yz\cap xyz) = yz\cap Cxyz = y\cap Cxyz = y$. So $y\in Cxyz$.

Thus $xy\cap Cxyz\neq\emptyset$. But $C(xy\cap xyz)=Cxy=x$ by assumption (a). Contradiction!

Since $yz\cap Cxyz = \emptyset$, $Cxyz=x$, by Forced Choice (ii).

Note that $xz\cap Cxyz$ cannot be null, by Choice (i). So $Cxz = C(xz\cap xyz)=x$.

Next, we show that $>_C$ is stratified. We must show that '$\|$' is symmetric, reflexive, and transitive. Symmetry and reflexivity are immediate from the definition of '$\|$'. To show transitivity, assume:

(a) Cxy=xy

(b) Cyz=yz

We claim that Cxz=xz.

Cxyz has a nonnull intersection with at least one of xy, yz.

If xy, then $Cxy=C(xy\cap xyz)=xy\cap Cxyz=xy$ so $xy \subseteq Cxyz$. And then $yz\cap Cxyz \neq \emptyset$ as well.

If yz, then $Cyz=C(yz\cap xyz)=yz\cap Cxyz=yz$ so $yz\subseteq Cxyz$. And then $xy\cap Cxyz \neq \emptyset$ as well.

So $yz\cap Cxyz \neq\emptyset$ and $xy\cap Cxyz \neq \emptyset$ must both hold, and both of $xy \subseteq Cxyz$ and $yz\subseteq Cxyz$.

Therefore, Cxyz=xyz, so that $xz\cap Cxyz\neq\emptyset$.

Therefore, $Cxz= C(xz\cap xyz) = xz\cap Cxyz=xz\cap xyz=xz$. □

Proposition 2. C is max in its induced order.

Proof.

Say $x\in CX$. Then $x>y$ for any $y\notin CX$. And no $y\in CX$ is such that $y>x$.

Say $x\in \max(X)$. If $x\notin CX$ then $\forall y\in CX$, $y>x$, and $\exists y \in CX$ by Forced Choice. Contradiction! □

Proposition 3. Let > be a stratified partial order on U. Then max for this order is an OT Choice function.

Proof. The function max clearly satisfies Choice and Forced Choice. We show Independence..

Claim: $Y\cap\max(X) \subseteq \max(Y\cap X)$

Suppose $a\in Y\cap\max(X)$. This implies $a\in\max(X)$ and $a\in X$. If $\exists x\in Y\cap X$ with $x>a$ then since $x\in X$, $a\notin\max(X)$. So there is no such x, and $a\in \max(Y\cap X)$.

Claim: $\max(Y\cap X) \subseteq Y\cap\max(X)$.

Suppose $a\in\max(Y\cap X)$. We have $a\in Y$ and $a\in X$; we need $a\in\max(X)$. By assumption, $Y\cap\max(X)\neq\emptyset$. Consider any $c\in Y\cap\max(X)$. By what has just been shown, $c\in\max(Y\cap X)$. If $a\notin\max(X)$, then $c>a$ and $a\notin\max(Y\cap X)$. Contradiction!. So $a\in\max(X)$, as desired. □

Proposition. 4. Composition of OTC's. Let C,D be OTC functions. Then $C\circ D$ is an OTC function.

Proof.

(i) Choice. $DX\subseteq X$ and $C\circ D(X)=C(DX)\subseteq DX\subseteq X$.

(ii) Forced Choice. If $X\neq \emptyset$, then $DX\neq \emptyset$, then $C\circ D(X)=C(DX)\neq\emptyset$.

(iii) Independence. Suppose $Y\cap CDX\neq\emptyset$. Since $CDX\subseteq DX$, $Y\cap DX\neq\emptyset$. So $D(Y\cap X)=Y\cap DX$, and $CD(Y\cap X)=C(Y\cap DX)=Y\cap CDX$.

Appendix 2. Direct Implication Checking and RCD.

The method of implication-checking proposed in §5 involves indirection: we examine not 'φ' but rather '$-\varphi$' and we work through RCD. This raises the question of whether a direct assault on the problem might not have a significantly different character. Here we show that it does not.

Given the charge to determine the validity of $\mathcal{A}\vdash\varphi$, we seek to find a minimal entailing set $\Psi\subseteq\mathcal{A}$, or at least to assure ourselves of its existence. What we need is a subset $\Theta\subseteq\mathcal{A}$ which has the property $f\Theta\vdash\varphi$. The existence of such a subset certifies the desired implication $\mathcal{A}\vdash\varphi$, because $\wedge\Theta\vdash f\Theta$ (Prop. 2.1, (15)). The nonexistence of such a subset certifies the failure of the implication, because any minimal entailing set Ψ for φ has $f\Psi\vdash\varphi$ (corollary to Prop 2.5, (27)).

We cannot be content with grossly examining whether $f\mathcal{A}\vdash\varphi$. If it does, we are done, but if it does not, we still have work to do. To see this, let us consider the conditions under which the entailment can fail.

(195) Conditions leading to failure of the entailment $f\mathcal{A}\vdash\varphi$

I	$[f\mathcal{A}]_k$=W	$[\varphi]_k$=L
II	$[f\mathcal{A}]_k=e$	$[\varphi]_k$=L
III	$[f\mathcal{A}]_k$=W	$[\varphi]_k=e$

Of these, conditions I and II are immediately fatal for the prospects of $\mathcal{A}\vdash\varphi$ as well.

• Condition I. If $f\mathcal{A}\vdash\varphi$, then there is minimal entailing set $\Psi\subseteq\mathcal{A}$ with $f\Psi\vdash\varphi$. But in condition I, it must be that $[f\Psi]_k\neq$L for every $\Psi\subseteq\mathcal{A}$, otherwise we'd have $[f\mathcal{A}]_k$=L. But $[\varphi]_k$=L requires L at the k^{th} coordinate of any entailing ERC. (L may be retracted but not added.) So no such minimal entailing set exists, and it can't be that $\mathcal{A}\vdash\varphi$.

• Condition II. If $[f\mathcal{A}]_k=e$, then $[f\Psi]_k=e$ for any $\Psi\subseteq\mathcal{A}$ and there is no minimal entailing set here either.

Condition III differs from the others in that its defect is potentially remediable. If $[f\mathcal{A}]_k$=W, it's not the case that every subset $\Theta\subseteq\mathcal{A}$ need have the property that $[f\Theta]_k$=W, since a W can arise as the fusion of W's and e's. Therefore a minimal entailing set Ψ, with $f\Psi=e$, may yet lurk inside $\mathcal{A}$. We need only dig it out, or do enough digging to detect its presence.

Under condition III, with $[\varphi]_k=e$, the minimal entailing set Ψ must have $[f\Psi]_k=e$. The virtuous behavior of such a Ψ can be masked by the presence of extraneous spoiler ERCs in $\mathcal{A}$ which have W at their k^{th} coordinate. But these can be removed, and we will be a step closer to Ψ. What we want is a subset $\Theta\subseteq\mathcal{A}$ with the property that $f\Theta=e$. If such a subset exists, it might contain a minimal entailing set. If no such subset exists, there is no entailing set.

If we remove from $\mathcal{A}$ all $\alpha\in\mathcal{A}$ which have W where φ has e, then we will be left with a subset $\mathcal{A}'\subseteq\mathcal{A}$ which is better off with respect to condition III situations. A subtlety now arises. The removal of any such α can change also the situation *for the worse*. Consider the following example:

(196) **Masking of entailment**

	C_1	C_2	C_3	C_4	C_5
α	W	L			
ψ_1		W		L	
ψ_2			W	L	L
φ	*e*	*e*	W	L	*e*

Suppose we are trying to ascertain whether $\{\alpha,\psi_1,\psi_2\}\vdash\varphi$. The obvious point of delicacy is with α at C_1. Removal of α discloses a further problem: the valuable L in C_2 is attached to α; with α gone, we see that $(\psi_1\circ\psi_2)$ suffers from a condition III issue precisely at C_2.

We must therefore iterate the removal procedure and extract ψ_1 from the potential entailing set. In the case at hand, this successfully reveals ψ_2 as a potential (and indeed actual) entailer. In other cases, yet further problems may be revealed and iteration must continue.

Notice that finding a subset Θ free of Condition III problems does not guarantee entailment:

(197) Θ is not enough

ψ	W	L	
φ	W		L

Here ψ doesn't entail φ, nor vice versa, yet no condition III situations are in evidence. The result of condition III elimination is a subset whose fusion must still be tested for the entailment relation.

The disclosure of columns fusng to W is of course the very motif upon which RCD is built, and the iterating removal procedure – like RCD – looks for W-fusing columns. The difference is that condition III repair is not concerned with those columns where φ itself has W.

But if we switch our focus to $-\varphi$, as in the inconsistency-testing procedure of §5, RCD will continue as before to seek out the columns where φ (and $-\varphi$!) have *e*. But RCD no longer wants those columns where φ has W, because *these now fuse to L* when $-\varphi$ is included. The switch to $-\varphi$ thus eliminates one source of extraneous noise from RCD. It has another crucial effect. Columns where $[\varphi]_k$=L are those where $[-\varphi]_k$=W. RCD will wish to stratify any of these that it can. And if $\mathcal{A}\cup-\varphi$ is consistent it will succeed, demonstrating *non*entailment of φ. If $\mathcal{A}\cup$-φ is inconsistent, it will fail, demonstrating $\mathcal{A}\vdash\varphi$.

RCD thus natively includes the entailment check that condition III repair must add on.

Appendix 3. Entailment and Nonentailment between Fusions and Fusands

				W(B) ⊆ P(A)?	W(A) ⊆ P(B)?	L(B) ⊆ L(A)?	L(A) ⊆ L(B)?
	A	**B**	**A∘B**	**A∘B⊢A**	**A∘B⊢B**	**A⊢A∘B**	**B⊢A∘B**
1	(*e*,W,L)	(W,L,*e*)	(W,L,L)	no	YES	no	no
2	(*e*,W,L)	(W,L,L)	(W,L,L)	no	YES	no	YES
3	(*e*,W,L,L)	(W,W,W,L)	(W,W,L,L)	no	YES	YES	no
4	(W,L,*e*)	(W,L,W)	(W,L,W)	no	YES	YES	YES
5	(W,L,*e*,*e*)	(*e*,*e*,W,L)	(W,L,W,L)	no	no	no	no
6	(*e*,*e*,L,W)	(W,L,L,*e*)	(W,L,L,W)	no	no	no	YES
7	(W,L,L,*e*)	(*e*,*e*,L,W)	(W,L,L,W)	no	no	YES	no
8	(W,L,*e*)	(*e*,L,W)	(W,L,W)	no	no	YES	YES
9	(W,L,*e*)	(*e*,W,L)	(W,L,L)	YES	no	no	no
10	(W,W,L,*e*)	(*e*,W,L,L)	(W,W,L,L)	YES	no	no	YES
11	(W,L,L)	(*e*,W,L)	(W,L,L)	YES	no	YES	no
12	(W,L,W)	(W,L,*e*)	(W,L,W)	YES	no	YES	YES
13	(W,L,*e*)	(W,*e*,L)	(W,L,L)	YES	YES	no	no
14	(W,L,*e*)	(W,L,L)	(W,L,L)	YES	YES	no	YES
15	(W,L,L)	(W,L,*e*)	(W,L,L)	YES	YES	YES	no
16	(W,L)	(W,L)	(W,L)	YES	YES	YES	YES

For the conditions on the coordinates cited at the head of the yes/no columns, see (37), p. 19, and (38), p. 20.

Appendix 4. A Kripke-Style Semantics for OT

Dunn 1976 (also ABD:177ff) develops a Kripke or 'world' style semantics for RM, which only uses three truth values. By adapting his construction, we may do the same for OT, i.e. for RM based upon polyvaluations. In this informal overview, I will modify his notation, terminology and definitions rather freely, in the interests of (what I deem to be) accessibility in the present context; the interested reader should turn to the cited pieces for the original details.

The chief ingredients are these: a set of 'worlds' $K=\{w_i\}$, an order relation '$\leq$' on them, and a valuation function $f(p, w_k)$ that assigns values from $\{T,e,F\}$ to prop letters at each world, extended to entire sentences in the usual recursive manner. Crucially, the order relation is required to be linear (strict domination cannot be far in the background). Furthermore, valuation at each world is not entirely independent, but must meet the "hereditary condition".

(198) **Hereditary Condition**. If $w_i \leq w_j$ then $f(p, w_i) = e$ or $f(p, w_i) = f(p, w_j)$.

According to the Hereditary Condition, any legitimate string of world-based valuations will begin with a (possibly null) sequence of *e*'s, which is then possibly followed up by a string of T's, or a string of F's; polar values are therefore 'inherited' intact in successve worlds. As we advance in through the worlds in order, we may start out evaluating any given prop letter to *e* and then possibly move on to one of the polar values, with which we must remain.

Evaluation of complex sentences proceeds at each world by RM3 rules, with the wrinkle that the positive intensional connectives – fusion, fission, arrow – look back in the order to determine their local value. For conciseness, let us write 'rm3[φ]' for the RM3 valuation of φ.

(199) **Arrow** (Dunnean)

$$
\begin{aligned}
f(A \rightarrow B, w_k) &= F \text{ iff } \exists w_i \leq w_k,\ \mathrm{rm3}[f(A, w_i) \rightarrow f(B,w_i)] = F \\
&= e \text{ iff } \forall w_i \leq w_k,\ \mathrm{rm3}[f(A,w_i) \rightarrow f(B,w_i)] = e \\
&= T \text{ otherwise.}
\end{aligned}
$$

In short, $A \rightarrow B$ is F at w_k if its RM3 valuation, based on the value of its consituents, is F anywhere in $w_1,\ldots,w_k$. It is *designated* at w_k iff it is designated throughout $w_1,\ldots,w_k$, and T iff one of those designated values is T.

The passage to OT clears when we observe that the linearity of R, in consort with the Hereditary Condition, imposes very narrow requirements on the progression of values. Instead of having to survey the entirety of the sequence, we need only examine the *first* occasion on which $A \rightarrow B$, or indeed $A+B$, $A \circ B$, assumes a polar value.

(200) **Arrow** (OT style)

- Let w_{dom} be the first world in K such that rm3[$f(A, w_{dom}) \rightarrow f(B, w_{dom})$] is polar. Then $f(A \rightarrow B, w_k)$=rm3[$f(A, w_{dom}) \rightarrow f(B, w_{dom})$] for all $w_k \geq w_{dom}$.
- If there is no such $w_{dom} \in K$, $w_{dom} \leq w_k$, then $f(A \rightarrow B, w_k)=e$.

Adverting to firstness, *i.e.* dominance, eliminates the need for the Hereditary Condition; without it, we find ourselves squarely in the territory of ordered polyvaluations. If we identify each world w_k with a valuation v_k, but assign the world's valuation on the basis of w_{dom}, we are just doing OP semantics without the numbers. Extensional connectives combine the dominant values of their constituents; intensional connectives take as dominant the world where the combination first achieves polarity.

To proceed in a more Dunn-like fashion, let us suppose that we are given a polyvaluation V, and an order R on $V=\{v_1,...,v_n\}$. For convenience, assume that R is given by the order of indices. We put the worlds in the same order $\{w_1,...,w_n\}$ and define the basic valuation function f as follows:

$$f(p,w_1) = v_1(p)$$
$$f(p,w_k) = v_k(p) \text{ if } f(p, w_{k-1})=e,$$
$$\text{else } f(p,w_k) = f(p,w_{k-1})$$

This version of f incorporates the Hereditary Condition, and extension to complex formula then proceeds exactly as before.

The valuation at w_k is exactly the one that would be assigned by the *initial segment* of the ordered polyvaluation running from v_1 to v_k, flattened so as to only distinguish polar from nonpolar and designated from nondesignated values. In essence, a world w_k functions like the ordered subpolyvaluation $V_k=\langle v1>\ldots>vk\rangle$, assigning to φ *e* if $v_i(\varphi)=e$, $i \leq k$, else T if $V_k \Vdash \varphi$ and F if $V_k \Vdash \neg\varphi$.

Appendix 5. Axioms for S and RM

Sobociński gives the following axioms for S (Sobociński1952:23). Parks (1972) shows that S constitutes the implication-negation fragment of RM.

1. $(A\rightarrow B)\rightarrow((B\rightarrow C)\rightarrow(A\rightarrow C))$	Transitivity of '→' (suffixing)
2. $A\rightarrow((A\rightarrow B)\rightarrow B)$	Assertion
3. $((A\rightarrow(A\rightarrow B))\rightarrow(A\rightarrow B)$	Contraction
4. $A\rightarrow(B\rightarrow(\neg B\rightarrow A)$	Double excluded middle
5. $(\neg A\rightarrow\neg B)\rightarrow(B\rightarrow A)$	Contraposition & double negation.

Axioms 1,3 are found in Church 1951ab, along with '$A\rightarrow A$' and '$[A\rightarrow(B\rightarrow C)]\rightarrow[B\rightarrow(A\rightarrow C)]$', which are theorems of S. Therefore, as noted by Sobociński §4.8, p. 54, Church's negation-free 'weak positive implicational calculus' is a subsystem of S. Sobociński deduces a collection of 187 theorems from these axioms. Modus ponens and substitution are the only rules of deduction.

Among the most striking of his meta-results is the finding that no wff of S can be a theorem if it contains a prop letter that occurs only *once* (Sobociński 1932, 1952). From this it follows that $A\rightarrow(B\rightarrow A)$ and the closely related $A\circ B\rightarrow A$ cannot be theorems.

It is also worth noting that only half of the distributive law goes through:

Quasi-distribution in S.

(i) $A\circ(B+C)\rightarrow(A\circ B)+(A\circ C)$

(ii) $(A+B)\circ(A+C)\rightarrow A+(B\circ C)$

These are essentially the same – (ii) follows from (i) by substitution of ¬X for X, contraposition, De Morgan, and double negation; and vice versa. The converse implications are not valid.

A set of axioms for RM is provided in AB:341/394.

R1	$A\rightarrow A$	Identity
R2	$(A\rightarrow B)\rightarrow((B\rightarrow C)\rightarrow(A\rightarrow C))$	Transitivity of →'(suffixing)
R3	$A\rightarrow((A\rightarrow B)\rightarrow B)$	Assertion
R4	$(A\rightarrow(A\rightarrow B))\rightarrow(A\rightarrow B)$	Contraction
R5	$A\wedge B\rightarrow A$	Conjunction Elimination
R6	$A\wedge B\rightarrow B$	Conjunction Elimination
R7	$(A\rightarrow B)\wedge(A\rightarrow C)\rightarrow(A\rightarrow B\wedge C)$	Conjunction Introduction
R11	$A\wedge(B\vee C)\rightarrow(A\wedge B)\vee C$	Distribution
R12	$(A\rightarrow\neg B)\rightarrow(B\rightarrow\neg A)$	Contraposition
R13	$\neg\neg A\rightarrow A$	Double Negation
RM0	$A\rightarrow(A\rightarrow A)$	Mingle

Rules of deduction are modus ponens, conjunction introduction, and substitution.

References

ROA = Rutgers Optimality Archive, http://roa.rutgers.edu
RuCCS-TR = http://ruccs.rutgers.edu/publicationsreports.html
NDJFL = Notre Dame Journal of Formal Logic
JSL= The Journal of Symbolic Logic

Anderson, Alan Ross, and Nuel D. Belnap. 1959. A simple treatment of truth functions. JSL 24:107-111.

Anderson, Alan Ross, and Nuel D. Belnap, Jr. 1975. *Entailment: the logic of relevance and necessity. Volume 1.* Princeton University Press. [AB]

Anderson, Alan Ross, Nuel D. Belnap, Jr., and J. Michael Dunn. 1992. *Entailment: the logic of relevance and necessity. Volume I1.* Princeton University Press. [ABD]

Avron, Arnon. 1987. A constructive analysis of RM. JSL:52.4, 939-951.

Avron, Arnon. 1991. Natural 3-valued logics – characterization and proof theory. JSL:56.1, 276-294.

Belnap, Nuel D., Jr. 1959. *The formalization of entailment*. Yale University doctoral dissertation.

Belnap, Nuel D., Jr. 1960. A formal analysis of entailment. Technical Report No. 7, Contract No. SAR/Nonr-609(l6), Office of Naval Research, New Haven.

Besnard, Ph., G. Fanselow, and T. Schaub. From a linguistic theory to cumulative logics. Handout from presentation at Fifth Workshop on Optimality Theoretic Syntax. Potsdam. December 13-14, 2001.

Białnicki-Birula, A. and H. Rasiowa. 1957. On the representation of quasi-Boolean algebras. *Bulletin de l'académie Polonaise des sciences*, 5: 259-261.

Church, Alonzo. 1951a. The weak positive implicational propositional calculus (abstract). JSL, 16:238.

Church, Alonzo. 1951b. The weak theory of Implication. In A. Menne, A Wilhelmy, and H. Angsil, eds., *Kontrolliertes Denken, Untersuchungen zur Logikkalkül und zur Logik der Einzelwißenschaften*, pp. 22-37. Kommissions-Verlag Karl Alber: München.

Crowhurst, Megan, and M. Hewitt. 1997. Boolean Operations and Constraint Interaction in Optimality Theory. ROA-229.

Dunn, J. Michael. 1966. *The algebra of intensional logics*.University of Pittsburgh doctoral dissertation.

Dunn, J. Michael. 1970. Algebraic Completeness Results for R-mingle and its Extensions. JSL, 35.1:1-13.

Dunn, J. Michael.1976. A Kripke-Style Semantics for R-Mingle Using a Binary Accessibility Relations. *Studia Logica*, 35, 2: 163-172.

Dunn, J. Michael. 1979. R-Mingle and Beneath. Extensions of the Routley-Meyer Semantics for R. NDJFL: 20,2: 369-376.

Dunn, J. Michael. 1986. Relevance Logic and Entailment. In D. Gabbay and F. Guenther, eds., *Handbook of Classical Logic, Volume III: Alternative in Classical Logic*. 117-224.

Fisk, Milton. 1964. *A Modern Formal Logic*.

Grimshaw, Jane. 1999/2001. Optimal Clitic Positions and the Lexicon in Romance Clitic Systems. In Legendre, Géraldine, Jane Grimshaw, and Sten Vikner, eds., *Optimality-Theoretic Syntax*. Also ROA-374.

McCarthy, John and Alan Prince. Generalized Alignment. In *Yearbook of Morphology 1993*, 79-153, ed. Geert Booij and Jaap van Marle. Boston: Kluwer Academic Publishers. ROA-7.

Meyer, Robert K. 1966. *Topics in modal and many-valued logics*.University of Pittsburgh doctoral dissertation.

Meyer, Robert K. 1971. R-mingle and relevant disjunction. Abstract, JSL, 36: 366.
Meyer, Robert K. 1973. On conserving positive logics. NDJFL, 14:224-236.
Meyer, Robert K. 1975. Chapters 29.3 and 29.12 in Anderson & Belnap 1975.
Mortensen, Chris. 1995. *Inconsistent Mathematics*. Kluwer Academic Publishers. Dordrecht.
Parks, R. Zane. 1972. A note on R-mingle and Sobociński's three-valued logic. NDJFL 13:227-228.
Prince, 1997 et seq. Various handouts on Paninian relations. LSA Linguistics Institute, etc.
Prince, Alan. 2000. Comparative Tableaux. ROA-376. Replaces A Proposal for the Reformation of Tableaux (1998), formerly ROA-288.
Prince, Alan & Paul Smolensky. 1993. *Optimality Theory: Constraint Interaction in Generative Grammar*. RuCCS-TR-2, Rutgers Center for Cognitive Science, Piscataway.
Rasiowa, Helena. 1974. *An Algebraic Approach to Non-classical Logics*. North Holland.
Rescher, Nicholas. 1969. *Many-valued Logic*. McGraw-Hill.
Restall, Greg. 2000. *An Introduction to Substructural Logics.*Routledge: London and New York.
Samek-Lodovici, Vieri & Alan Prince. 1999. Optima. ROA-363. RuCCs-TR-57, Rutgers Center for Cognitive Science, Piscataway. [SLP]
Samek-Lodovici, Vieri & Alan Prince. 2001. Properties of Harmonic Bounding. Ms. Rutgers University and UC London.
Słupecki, J. 1930. Pełny trójwartośiowy rachunek zdań. Comptes Rendues des séances de la Société des sciences et des lettrees de Varsovie, XXIII, Classe III. Warsaw.
Sobociński, Bolesław. 1932. A badań nad teorią deducji. Przeglad Filozoficzny, XXXV, §4. Warsaw.
Sobociński, Bolesław. 1952. Axiomatization of a partial system of three-valued calculus of propositions. *The Journal of Computing Systems*, 1:23-55.
Sugihara, T. 1955. Strict implication free from implicational paradoxes. *Memoirs of the faculty of Arts, Fukui University*, Series I, Humanities and social sciences, 5, 55-59.
Tesar, B., & Smolensky, P. (1994). The Learnability of Optimality Theory. In R. Aranovich, W. Byrne, S. Preuss, and M. Senturia (eds.), *The Thirteenth West Coast Conference on Formal Linguistics,* 122-137. CSLI.
Tesar, Bruce. 1995. *Computational Optimality Theory.* Ph.D. Dissertation, Univ. of Colorado. ROA-90.
Tesar, Bruce and Paul Smolensky. 1998. Learnability in Optimality Theory. *Linguistic Inquiry* 29.2, 229-268. ROA- 155, ROA-156.
Tesar, Bruce and Paul Smolensky. 2000. *Learnability in Optimality Theory*. MIT Press: Cambridge.
Woodruff, Peter. 1969. *Foundations of three-valued logic.* University of Pittsburgh doctoral dissertation.