Decomposition of multiplace functions in Operations Research

Bernhard von Stengel
Informatik 5, University of the Federal Armed Forces at Munich, D–85577 Neubiberg

In: Operations Research '92, Hrsg. A. Karmann et al., Physica-Verlag, Heidelberg, 145–148.

This extended abstract summarizes the results of a decomposition theory for multiplace functions that generalizes and unifies theories known from a number of areas in Operations Research. The considered decompositions of a multiplace function are representations as terms of functions of fewer variables where variables may be used only once. This restricted "disjoint" functional superposition or "substitution" has been defined independently in switching circuit design, combinatorial optimization over networks and clutters and ordinal and expected utility theory. There, it has led to interesting results on unique "normal form" representations, like additive utility functions. These results have great similarities that are explained by the proposed theory, where the admitted decompositions are characterized set-theoretically: An n-ary operation f on a given set is decomposed into "conditional" functions obtained from f by fixing variables suitably. The following exposition is fairly technical to state results precisely. Proofs are found in [5] and further references in [4][5].

The ranges for the considered functions and their variables are non-empty sets S_i called *coordinate axes*, indexed by *coordinates* i from some index set I. For a finite set $A \subseteq I$ of coordinates, denote by $\langle A\rangle$ the cartesian product $\prod_{i\in A} S_i$ in its usual set-theoretic definition as the set of functions x defined on A with $x(i) \in S_i$ for $i \in A$ (if $S_i = S$ for all $i \in A$, then $\langle A\rangle = S^A$). This allows to postulate for disjoint sets A, B the equality $\langle A \cup B\rangle = \langle A\rangle \times \langle B\rangle$.

A *multiplace function* f is defined on $\langle N\rangle$ for some finite $N \subseteq I$, where $|N|$ is the *arity* of f. The range F of f is some axis S_j, identified with $\langle\{j\}\rangle$. Only total functions $f\colon \langle N\rangle \to F$ are considered. With reference to the domain $\langle N\rangle$ of f, call $i \in N$ the coordinates *of* f, indexing the variables of f, and the sets S_i axes *of* f.

The *substitution* of a multiplace function h into a variable of another multiplace function g is defined if $h\colon \langle A\rangle \to \langle\{i\}\rangle$, $g\colon \langle\{i\} \cup B\rangle \to F$, where $i \notin B$ and $A \cap B = \emptyset$, resulting into the function $f\colon \langle A \cup B\rangle \to F$ given by

$$f(x, y) = g(h(x), y) \qquad \text{for } x \in \langle A\rangle,\ y \in \langle B\rangle, \tag{*}$$

which is abbreviated as $f = g[h]$. (This abbreviation is unique since, in particular, the substituted variable of g is stored in the set-theoretic definition of h as the index i of its range, in some sense the "target" of the substitution of h.) Conversely, (*) is a *decomposition* of f for a suitable subset A of f's coordinate set N, where $B = N - A$. Note that A and B are always disjoint, so variables are never repeated; also, f, g, h are always total functions, and g is fully defined by (*) if h is surjective. Decompositions of f are *iterated* by applying them to g or h, and so on.

Applications of this decomposition are obtained by restricting the admitted functions f, g, h. For example, these are *Boolean* functions if all coordinate axes and ranges of the functions are equal to a two-element set $\{0,1\}$, where (*) denotes a "disjoint decomposition" of f, see [1]. If each axis is a real interval and one considers only continuous functions f, g, h strictly *increasing* in each variable, then the decomposition (*) with an interpretation of f as an (ordinal) *utility function* represents a *preference independence* [3] or "separability" [2] of the set A of "decision attributes". Further applications will be sketched at the end of this text.

Decompositions will be considered in a class Σ of multiplace functions that is closed under substitution. Certain axioms for Σ, presented next, shall serve three goals: decompositions in Σ can be done using "conditional functions" (this approach can be nicely developed just regarding a given multiplace function, which is skipped here for brevity); interesting unique representations hold for any $f \in \Sigma$; and a number of applications are obtained as special cases. Throughout, let $f\colon \langle N \rangle \to F$ be a multiplace function in Σ. Let $A \subseteq N$, $B = N - A$, and y' be any element of $\langle B \rangle$. Then the function $\langle A \rangle \to F$, $x \mapsto f(x, y')$, denoted by $f(\,\boldsymbol{.}\,, y')$, is called an f-*conditional* function, with *reference vector* y'. If this function is unary, with $A = \{i\}$, it is called an f-*translation*. (Thus, an f-conditional function is obtained by fixing the variables indexed by B at $y' \in \langle B \rangle$; each component of this vector y' could be regarded as a constant, defined on $\langle \emptyset \rangle$, substituted into f, but Σ shall *not* contain constants.)

Axiom 1. *$|N| \geq 1$, and for each $i \in N$ there is a bijective f-translation on $\langle \{i\} \rangle$.*

This axiom implies that all coordinate axes S_i, $i \in N$ of the domain of f have the same cardinality as its range F. It holds for Boolean functions f where all variables are *essential* (cannot be dropped) since a nonconstant f-translation $\{0,1\} \to \{0,1\}$ is bijective. By Axiom 1, all unary functions in Σ are bijections.

Axiom 2. *If $|N| = 1$, then the inverse of f belongs to Σ.*

Two multiplace functions $f, g \in \Sigma$ are called *isotopic* if f is obtained from g by composition with unary functions $\phi, \psi_1, \ldots, \psi_n \in \Sigma$ as in $\phi f = g[\psi_1, \ldots, \psi_n]$, where ϕf means $\phi[f]$ and $g[\psi_1, \ldots, \psi_n]$ denotes the "parallel substitution" of all n variables of g by values of the bijections $\psi_1, \ldots, \psi_n$, which can be viewed as transformations of the variables of f ($n = |N|$); if these are all equal to the transformation $\phi\colon F \to G$ of the value of f, then f and g are (in the usual sense) *isomorphic* operations $F^n \to F$ respectively $G^n \to G$ (a function $F^N \to F$ is hereby regarded as an n-ary operation $F^n \to F$, disregarding names of coordinates). Axioms 1 and 2 show that $f \in \Sigma$ is isotopic to an n-ary operation g on F that has the identity on F as a suitable g-translation on each coordinate axis; for a binary operation g ($n = 2$), this means there is a left- and a right-neutral element for g.

Axiom 3. *With f, all f-conditional functions fulfilling Axiom 1 belong to Σ.*

Axiom 3 is fairly natural. For example, f-conditional functions for Boolean functions f are also Boolean. However, some of them may have inessential variables so Axiom 1 is violated; they are not considered in Axiom 3 since otherwise the axiom would be too restrictive for f.

Axiom 4. *All surjective f-translations are bijective.*

Axiom 4'. *(Alternatively.) All injective f-translations are bijective.*

Axiom 4 (or 4') is only necessary if f is not a discrete function (with finite coordinate axes). Essentially, it guarantees a "compatibility" of reference vectors in decompositions, since if $f = g[h]$, an f-translation ϕ on a coordinate axis of h is composed of a g- and a h-translation, and by Axiom 4 or 4' these translations are bijective if ϕ is bijective.

Let Σ be a collection of functions closed under substitution fulfilling Axioms 1, 2, 3 and 4 or 4', called a *function system*. If $f \in \Sigma$ and there are $g, h \in \Sigma$ with $f = g[h]$ and h is defined on $\langle A \rangle$, then A is called an *f-autonomous* set [4] of coordinates of f. Autonomous sets characterize the decomposition possibilities of f within Σ.

Lemma. *For $f \in \Sigma$, defined on $\langle N \rangle$, a set $A \subseteq N$ is f-autonomous iff $f = g[\phi h]$ for suitable f-conditional functions g, h and a bijection ϕ, with h defined on $\langle A \rangle$. Thereby, the reference vectors for g and h are obtained as parts of z' whenever $f(\,\boldsymbol{.}\,, z')$ is a bijective f-translation on $\langle \{i\} \rangle$ for $i \in A$, $z' \in \langle N - \{i\} \rangle$, and g, h, ϕ belong to Σ.*

This central lemma asserts that decompositions of f can always be represented using f-conditional functions, and their reference vectors are recognized from the bijective f-translations of Axiom 1. The lemma allows to apply decompositions of f to f-conditional functions in other decompositions. This permits simple proofs of the "unique decomposition" theorems presented next.

Iterated decompositions should end whenever the arity of the functions can not be reduced. Call $f \in \Sigma$ *prime* (that is, indecomposable) if f is not unary and $f = g[h]$ implies g or h is unary, for $g, h \in \Sigma$.

Jordan-Hölder-Theorem. *Let $f \in \Sigma$ and $f = \big((g_1[g_2]) \cdots\big)[g_k]$. Then if the functions $g_1, \ldots, g_k \in \Sigma$ are prime, they are unique up to their order of substitution and isotopy.*

This theorem is analogous to that for groups. Isotopy is the natural invariance concept here since unary bijections (transformations of variables) followed by their inverses can be interspersed anywhere in the sequence of substitutions. Note that, as for groups, the (isotopy classes of) the prime "factors" $g_1, \ldots, g_k$ usually do not determine f uniquely, so this is weaker than the "decomposition into primes" of natural numbers.

The main representation theorem treats term decompositions that reveal more than iterated substitutions, like $f(x, y, z, t) = g(a(x) \cdot b(y) \cdot c(z), t)$ for functions $f, g, a, b, c \in \Sigma$ with an associative "product" $\cdot$ in Σ showing the autonomy of the coordinate sets A, B, say, that define the variable vectors (x, y) and (y, z). It will be seen that such a product representation can always be found if A and B are autonomous sets that *overlap*, that is, $A - B$, $A \cap B$ and $B - A$ are not empty (note that then these sets, corresponding to the vectors x, y and z, are also autonomous).

Theorem. *Let $f \in \Sigma$, defined on $\langle N \rangle$. Then N and $\{i\}$ are f-autonomous for all $i \in N$. If $A, B \subseteq N$ overlap and are f-autonomous, then so are $A \cup B$, $A \cap B$, $A - B$ and $B - A$. Any collection $\mathcal{C}$ of subsets of N with this property can be represented by a labeled tree $T(\mathcal{C})$ with nodes being those elements of $\mathcal{C}$ not overlapping with any other element of $\mathcal{C}$, root N and leaves $\{i\}$, $i \in N$, where the successors of a node A form a partition $B_1, \ldots, B_m$ of A and the set $\mathcal{C}_A = \{B \in \mathcal{C} \mid B \subseteq A\}$ can be represented as $\mathcal{C}_A = \mathcal{D}_A \cup \mathcal{C}_{B_1} \cup \cdots \cup \mathcal{C}_{B_m}$ with three possibilities for $\mathcal{D}_A$ coded by labels attached to A:*

label "prime": $\mathcal{D}_A = \{A\}$

label "linear": $\mathcal{D}_A = \{\bigcup_{i=l}^{k} B_i \mid 1 \le l < k \le m\}, \quad m \ge 3$

label "full": $\mathcal{D}_A = \{\bigcup_{i \in L} B_i \mid L \subseteq \{1, \ldots, m\},\ |L| \ge 2\}, \quad m \ge 3.$

In this theorem, $\mathcal{C}$ can be reconstructed from the tree $T(\mathcal{C})$ with the labels, where, if a node A is labeled "linear", the order among its successors matters. With $\mathcal{C}$ as the collection of f-autonomous sets, $T(\mathcal{C})$ is called the *composition tree* for f [2][4]. It corresponds to a unique hierarchical term representation of f.

Theorem. *Let $f \in \Sigma$, defined on $\langle N \rangle$, and $\mathcal{C}$ be the set of f-autonomous sets. Then for each node A of $T(\mathcal{C})$ there is a function h_A defined on $\langle A \rangle$ such that $h_N = f$, and, with successors $B_1, \ldots, B_m$ of A if A is not a leaf, $h_A = g_A[h_{B_1}, \ldots, h_{B_m}]$, where if A is labeled "prime": g_A is prime and unique up to isotopy, "linear": $g_A(y_1, \ldots, y_m) = y_1 \circ \cdots \circ y_m$ with an associative operation $\circ$ in Σ that is unique up to isomorphy and not commutative, "full": then the same holds but with commutative $\circ$.*

Note that this theorem implies that f is a hierarchical term of functions g_A for the non-leaf nodes A of the composition tree $T(\mathcal{C})$ and of unary functions $h_{\{i\}}$; the latter can be omitted by substituting them into the next higher functions if these are prime. These functions can be stored with $T(\mathcal{C})$ to represent f. Any decomposition $f = g[h]$ of f is a subterm of this term since any autonomous set B is coded in $T(\mathcal{C})$, where h, defined on $\langle B \rangle$, may be some "subproduct" taken at some stage if B is the union of successors of a node labeled "linear" or "full".

This theorem has well-known representations as special cases, considering different function systems Σ. For Σ as the class of Boolean functions [1], the associative operation $\circ$ is addition $\oplus$ modulo 2 or conjunction $\wedge$ (via isomorphy, taking complements, also equivalence $\equiv$ or disjunction $\vee$). For monotone Boolean functions, $\wedge$ or $\vee$; their decomposition is the well-known substitution decomposition of *clutters* (antichains in the power set of N), see [4]. "Clutter polynomials" in a semiring like $\mathbb{R}_{\geq}, \max, +$, considering $f(x_1, \ldots, x_n)$ as a maximum of sums $\sum_{i \in C} x_i$ for elements C of a given clutter, for example maximal chains C in an order, also form a function system Σ; here $\circ$ is max or $+$. In this example, the decomposition of f corresponds to a decomposition of the order [4], which is series-parallel iff the prime nodes of the composition tree have only two successors.

Multiaffine are those functions $f \colon \mathbb{R}^n \to \mathbb{R}$ whose f-translations are of the form $x \mapsto ax + b$. They form a function system Σ, where, interpreting f as an *expected-utility* function, autonomous sets are called *generalized utility independent*. The operation $\circ$ is $+$ or $\cdot$, corresponding to the well-known additive/multiplicative expected-utility function [3]. For the continuous ordinal utility functions mentioned above, strictly increasing in each variable, $\circ$ is (up to isomorphy with a strictly increasing transformation like the logarithm) addition of reals, yielding an additive utility function [2]. For further applications and references see [4][5].

References

[1] Ashenhurst, R.L. (1959). The decomposition of switching functions. Proc. Int. Symp. Theory of Switching (April 1957), Part I. Ann. Comput. Lab. Harvard U. 29, 74–116.

[2] Gorman, W.M. (1968). The structure of utility functions. Rev. Economic Studies 35, 367–390.

[3] Keeney, R.L. (1974). Multiplicative utility functions. Oper. Res. 22, 22–34.

[4] Möhring, R.H. and F.J. Radermacher (1984). Substitution decomposition for discrete structures and connections with combinatorial optimization. Ann. Discr. Math. 19, 257–356.

[5] von Stengel, B. (1991). Eine Dekompositionstheorie für mehrstellige Funktionen. Mathematical Systems in Economics 123, Anton Hain, Frankfurt.