Journal of the American Philosophical Association (2015) 545–563 © American Philosophical Association
doi: 10.1017/apa.2014.15

Intrinsic Utility's Compositionality

abstract: *To compare the options in a decision problem, a common method evaluates for each option the world that would result if the option were realized. This paper argues that one evaluation of an option's world, intrinsic utility, is compositional given a division of an option's world according to the option's consequences and other events. The argument first justifies the norm that an ideal agent should be intrinsically indifferent between two options' worlds given that she is intrinsically indifferent between the options' consequences. Then it uses this norm and the existence of intrinsic utilities respecting intrinsic indifference to establish intrinsic utility's compositionality. The results regulate human agents when they approximate ideal agents in pertinent respects. The paper begins with a general explanation of compositionality; the related phenomena of interchangeability, complementarity, and independence; and the effect on compositionality of context and arrangement of a composite's parts. After arguing for intrinsic utility's compositionality, the paper explains its role in decision theory.*

keywords: complementarity, compositionality, decision theory, interchange of equivalents, utility

The structural properties of utility support methods of inferring and forming utilities. This paper presents the structural property of compositionality and uses norms of rationality to argue that a type of utility is compositional. Its compositionality assists inference of an agent's preference ranking of the options in a decision problem and, in cases where the preference ranking is incomplete, guides the agent's completion of the ranking.

The paper defines compositionality in a general way and then applies that definition to utility. It uses features of the general definition to explain features of the definition's application to utility. The general definition attends to the range of composites and their division into parts, and its application to utility also attends to these matters.

Compositionality for a property and a range of composites asserts the existence of a function that obtains a property's application to a composite from the property's application to the composite's parts. The function is addition in the case of weight but need not be addition in the case of other compositional properties. It is not addition for sentence meaning, for example, because meanings

I am grateful to anonymous referees for helpful comments.

are not quantitative. Additive compositionality assumes a quantitative property and a scale of measurement for the property. Whether a quantitative property's compositionality is additive may depend on conventions for measuring the property. Multiplicative compositionality becomes additive if a convention uses logarithms to rescale quantities.

Utility's compositionality conflicts with a type of value holism that denies that the utility of a composite is a function of the utilities of its parts, no matter what the type of utility, the type of composite, and its division into parts. However, utility's compositionality does not challenge value holism's opposition to utility's additivity, separability, atomism, invariantism, or nonrelativism. The function supporting utility's compositionality need not be addition and need not have separable variables for the utilities of a composite's parts. The parts need not be indivisible, and the utilities of parts need not be the same in all contexts. A part's utility may be relative to an agent and to background considerations as well as to features of context.

This paper assumes causal decision theory's method of comparing options using the options' worlds. It assumes that an option's world is the world that would be realized if the option were realized. The subjunctive conditional preserves causal relations as much as possible when supposing the option's realization. For simplicity, following Gibbard and Harper (1981), this paper assumes that given an option's realization, a unique world is nearest and serves as the option's world. Also, for simplicity's sake, in this paper I treat mainly cases in which an agent is certain of an option's world and the option's consequences. For cases of uncertainty, I adopt causal decision theory's definition of an option's expected utility as a probability-weighted sum of the utilities of the possible worlds that might be the option's world. However, this leaves open whether an option's evaluation should use its unconditional expected utility and whether an agent should choose an option that maximizes unconditional expected utility. Versions of causal decision theory differ according to their answers to these questions.

Sections 1–6 introduce utility's compositionality without specifying a type of utility, and section 7 specifies the type of utility that appears in the paper's principle of compositionality. The general account of compositionality applies to utility attaching to objects as well as to utility attaching to propositions, but here my principle of compositionality treats a type of utility attaching to propositions, which may represent possession of objects.

Some decision theorists, such as Savage ([1954] 1972), define probability and utility as functions that represent an agent's preferences as following expected utilities. Accordingly, probability and utility are constructed from preferences and do not represent attitudes toward propositions that are definitionally independent of preferences. Joyce (1999) extends this interpretation of probability and utility to causal decision theory. Within this school, the utility of a possible world (taken as a proposition specifying for each goal an agent has whether the agent attains it) is not compositional for any interesting division of the world into parts. Such compositionality also fails for alternative interpretations of utility that take it to represent a propositional attitude definitionally independent of preferences, assuming that a proposition's utility comprehensively evaluates the proposition's realization. In this paper I show that a world's *intrinsic utility*, a type of utility that evaluates a proposition noncomprehensively, displays a type of compositionality that simplifies evaluation of the options in a decision problem.

Some measurement theorists, such as Krantz and others (1971: chap. 6) seek a compositional quantitative representation of an order of composites, such as a preference order of commodity bundles. They assume a type of compositionality and derive a quantitative representation. In contrast, in this paper I consider a quantitative property attaching to components and to composites and ask whether it has a type of compositionality. The quantity comes first, and the issue of its compositionality is second. The objective is to justify a type of compositionality rather than simply assume it for some purpose. This strategy is appropriate for utility if its definition does not assume the type of compositionality to be justified. Section 7 introduces intrinsic utility, which need not, but may, be defined by assuming a type of compositionality that divides an option into prospects for various possible worlds; in that section I then argue that intrinsic utility exhibits another type of compositionality that divides an option's world into an option's consequences and other events.

1. A Property's Compositionality

Some properties apply to both an object and its parts. For example, weight attaches to a wall and the bricks that constitute the wall. Meaning attaches to a sentence and the words that constitute the sentence. A property's application to a composite is *compositional*, that is, it decomposes into its applications to the composite's parts if and only if the property's application to the composite is a function of the property's application to its parts, taken in their order if the composite orders them. The weight of a whole is an additive function of the weights of its parts. Granting Frege's theory of meaning, the meaning of a sentence is a function, not additive, of the meanings of its parts taken in their order. The weight of a composite and the meaning of a sentence are each compositional.

Although many types of compositionality are possible—for example, types involving different properties applied to a whole and its parts—in this paper I deal only with composition of a property's application to a whole based on the same property's application to the parts of the whole. Utility's compositionality is my target, but I consider compositionality of properties besides utility for general insights concerning compositionality and for methods of establishing it. Besides characterizing compositionality in general, in the first few sections I note assorted salient features of compositionality to enrich its characterization.

Compositionality governs a property, such as weight, whose application varies from object to object. A function represents the property by specifying for an object the result of the property's application to the object, that is, the property's value for the object. For weight, the function may specify for an object its weight in grams (with instead of the property being weight in grams and the property's value for an object being a number, the property being weight and the property's value for an object being a number-scale pair specifying a number on the scale). A quantitative property's application to an object yields the quantity of that property that the object possesses; for example, weight's application to an object yields the object's weight. However, compositionality may extend to a nonquantitative property whose application to an object yields a nonquantitative value of the property. The nonquantitative property of meaning applied to an expression yields a particular meaning; for example, the meaning of the expression 'snow' is snow. For simplicity's sake, in this paper I often identify a property with the function from objects to values that represents the property. Accordingly, the property of weight is a function from physical objects to numbers on a scale for weight. The property of meaning is a function from linguistic expressions to meanings.

The function that given a compositional property's applications to a composite's parts yields its application to the composite differs from the function that represents the property. For weight, the function addition, not the function weight, grounds compositionality. Adding the weights of a composite's parts yields the composite's weight. Assuming meaning's compositionality, some function combines the meanings of a sentence's words to obtain the sentence's meaning. This function awaits discovery even though for each sentence the meanings of the sentence and its words are known. Showing the function's existence requires showing that a move to a sentence's meaning from its words' meanings may use nothing but the words' meanings in their order and not other features of the words such as their syllables.

Whether a property's application to a whole is a function of the property's application to the parts of the whole depends on the nature of the parts. Compositionality always holds if the whole's only part is the whole itself but may fail for finer divisions. Consider the compositionality of meaning. The meaning of a sentence, even if it is compositional taking words as a sentence's parts, is not compositional when its parts are taken to be letters, spaces, and punctuation marks, because not all those elements have meanings. A finite set's cardinality is compositional when a partition of the set is taken to yield its parts; the set's cardinality equals the sum of the cardinalities of the partition's elements. However compositionality fails if the set's parts are its members and they lack cardinality.

According to a simple statement of compositionality for utility, the utility of a whole is a function of the utilities of its parts. This simple statement does not specify the parts of a whole. However, utility's compositionality may not hold for commodity bundles unless the commodities are dated, that is, given times of use or consumption. Compositionality of an option's utility depends on a division of the option into parts. An option's utility may be compositional dividing the option into possible outcomes taken as worlds that might be realized if the option were realized; the disjunction of the worlds is equivalent to the option. In contrast, its utility may not be compositional dividing the option into possible amounts of money that might be procured if the option were realized; the disjunction of procurements of the possible amounts is equivalent only to the option's monetary prospects. An evaluation of the option using only its possible monetary consequences may omit factors, such as risk, that influence the option's utility.

Because compositionality may hold for one division but not for another division of a whole into parts, a complete claim of compositionality specifies not only a property but also a division of a whole into parts. A complete claim explicitly states the division into parts that a property's compositionality assumes if context does not settle the division.

A function that grounds compositionality is general. It processes a property's applications to the parts of each combination in a class of multiple combinations possessing the property. If the function governed just a single case, it would exist trivially. It would have the property's applications to the parts as arguments and the property's application to their combination as value. For a single composite, the property's value for the whole is bound to be some function of its values for the parts because the function governs only one and not also any other composite. Every property's application to a composite is some single-case function of the property's application to the composite's parts, assuming that the property applies to the parts of the composite.

Consider utility. For any pair of objects of utility, some function goes from the utilities of the elements of the pair to the pair's utility. Nonetheless, for some collections of pairs, no general function goes from the utilities of a pair's elements to the pair's utility. Take two pairs made from one first element x and two second elements y_1 and y_2. It is possible that the pairs' utilities are not any general function of the elements' utilities. Assume that the utilities of y_1 and y_2 are equal and that the utilities of (x, y_1) and (x, y_2) differ. For any F, $F(U(x), U(y_1)) = F(U(x), U(y_2))$ because $U(y_1) = U(y_2)$. Because $U(x, y_1) \neq U(x, y_2)$, either $U(x, y_1) \neq F(U(x), U(y_1))$, or $U(x, y_2) \neq F(U(x), U(y_2))$. So F does not yield the pairs' utilities. Utility's compositionality requires a general function that applies to both pairs.[1]

A utility function goes from a pair's elements to the utilities of the pair's elements. Suppose that some general function goes from the utilities of the pair's elements to the pair's utility. Then some general composite function applying first the utility function and then the general function goes from a pair's elements to the pair's utility. This composite function does not establish utility's compositionality, however. Compositionality asserts the existence of a general function stating a property's application to a whole given the property's application to the parts of the whole, not the existence of a general function stating the property's application to the whole given its parts. The latter function may exist although the former does not. The meaning of a word is a function of its letters but not of the meanings of its letters (because its letters do not have meanings). A composite's utility may be a function of its components without being a function of its components' utilities. For example, a conjunction's utility is a function of its conjuncts. However, in cases of complementarity between propositions x and y, $U(x \& y) \neq U(x) + U(y)$, so a conjunction's utility is not in general the sum of its conjuncts' utilities. Features of the conjuncts besides their utilities may influence the conjunction's utility. If $U(y) = U(z)$, then $U(x) + U(y) = U(x) + U(z)$. Nonetheless, it may not be the case that $U(x \& y) = U(x \& z)$ because y, but not z, complements x. A conjunction's

1 A special case of compositionality occurs if a pair's first element settles the pair's utility and a fortiori the utilities of the pair's elements settle the pair's utility. A pair's utility is then a function of its elements' utilities even if a pair's utility does not change when only its second element and its second element's utility change because the first element swamps the second.

utility may not be a sum or any general function of its parts' utilities. The general function that for each combination selects the trivial function from the property's application to the parts to the property's application to the combination goes from a combination's parts to the property's application to the combination. It does not go from the property's application to the parts to its application to the combination and so does not ground the property's compositionality.

Notice that for variables x and y, if $U(x, y) = F(U(x), U(y))$, then $U(x, y) = F'(x, U(y))$ for some F'. Because x settles $U(x)$, it can replace $U(x)$ in a transformation of F that yields (x, y)'s utility. Two values of x may yield the same utility, but if $U(x)$ settles the composite's utility, then x settles it also. Compositionality involving components' utilities warrants such moves to compositionality involving some components directly.

The function establishing weight's compositionality has to accommodate combinations of an arbitrary number of objects. It may go from the number of objects in a combination to a subsidiary function with that many places. Or, it may be an infinite-placed function whose application to n objects takes their weights for its first n arguments and zero weight for all remaining arguments. The function for sentence meaning also has to accommodate sentences with an arbitrary number of words. It may similarly go from the number of words in a sentence to a subsidiary function with that number of places. Or the function may be an infinite-placed function in which the words of an n-word sentence furnish the first n arguments and the remaining arguments have a null value.

Some types of compositionality are not of interest. The weight of a brick of constant weight and a twelve-hour duration has a decomposition into weights of the brick's twelve consecutive hour-long stages. The brick's weight during its existence equals the value of the twelve-place function that applied to the weights of the brick's stages yields their common weight. This type of compositionality lacks interest. Interesting types of compositionality for weight divide objects spatially rather than temporally.

A function grounding a property's compositionality may have multiple specifications. In some cases a property's compositionality is of interest only if the function grounding it has a certain specification. The debate about meaning's compositionality concerns the existence of a simple, learnable description of a function that yields a sentence's meaning given the meanings of the words making up the sentence. In general, a property's compositionality is of interest only given a suitable specification of a function stating the property's application to a whole given the property's application to the parts of the whole. The theoretical framework for investigations of a property's compositionality settles a specification's suitability.

A type of compositionality specifies a property, the objects to which the property applies, and the components of composite objects to which the property applies. Weight applies to all material objects. The range of application of meaning is less clear; therefore, for definiteness one may specify that it applies to words or linguistic expressions composed of words (but not to paintings). Utility may apply to a material object or to an abstract proposition. In the later sections of this paper I treat utility's application to a proposition (given a way of understanding it) and let a principle of compositionality specify a method of dividing a proposition into parts.

Addition of weights works for all combinations of material objects. Compositionality of meaning may restrict itself to combinations of words that constitute sentences. A function establishing a property's compositionality under a restriction has a domain with the same restriction. Probability theory decomposes the probability of a disjunction of incompatible propositions into the probabilities of its disjuncts, but it does not decompose the probability of a conjunction of propositions into the probabilities of its conjuncts. In probability theory, a decomposition of an exclusive disjunction's probability is of interest even if the decomposition does not extend to all propositions. Restricted forms of utility's compositionality are also of interest in utility theory.

2. Interchange

Nontrivial compositionality treats a range of composites that includes composites generated by interchanging the parts of these composites. It treats cases in which some composites interchange parts that are equivalent with respect to the property. If no parts are equivalent and interchangeable, compositionality holds trivially, combining trivial functions for single cases.

According to the principle of *interchange of equivalents* for a property and range of composites, replacing a composite's part with another of the same value yields a new composite with the same value as the original, provided that the new part does not interact with remaining parts to generate an additional part. A property's compositionality entails interchangeability of equivalent parts, provided that the interchange does not create a composite outside the targeted range of composites. That is, given that a property's application to a whole is a function of the property's application to the parts of the whole, a substitution may introduce a part of equal value with respect to the property and not change the value of the whole with respect to the property. Suppose that for a property f and a function F that establishes f's compositionality $f(a_1, a_2, \ldots, a_n) = F(f(a_1), f(a_2), \ldots, f(a_n))$. If $f(a_i) = f(a_{i'})$, then $F(f(a_1), \mathrm{f}(a_2), \ldots, f(a_i), \ldots, f(a_n)) = F(f(a_1), f(a_2), \ldots, f(a_{i'}), \ldots, f(a_n))$. Hence, by compositionality, $f(a_1, a_2, \ldots, a_i, \ldots, a_n) = f(a_1, a_2, \ldots, a_{i'}, \ldots, a_n)$, as the principle of interchange of equivalents requires.

Interchangeability is equivalent to compositionality, that is, it is necessary and sufficient for compositionality, given a common range of composites. The previous paragraph shows that it is necessary. It is sufficient because interchange of equivalents for a property ensures that two composites whose components are the same with respect to the property are themselves the same with respect to the property. If interchange of equivalents moves from one composite to another, the principle of interchange of equivalents requires that the property have the same value for both composites. Consequently, some function states the property's value for a composite given its value for the composite's parts.

To establish this point, suppose that interchangeability holds but compositionality does not for a property f applied to pairs and their elements.

https://doi.org/10.1017/apa.2014.15 Published online by Cambridge University Press

Because interchangeability holds, for all x, x', and y, $f(x, y) = f(x', y)$ if $f(x) = f(x')$, and for all x, y, and y', $f(x, y) = f(x, y')$ if $f(y) = f(y')$. Because by supposition compositionality fails, no function F exists such that $f(x, y) = F(f(x), f(y))$ for all x and y. Hence, for some x, y, x', and y', $f(x) = f(x')$ and $f(y) = f(y')$, but $f(x, y) \neq f(x', y')$. However, by interchange, for those x, y, x', and y', $f(x, y) = f(x', y)$ and $f(x', y) = f(x', y')$ so that $f(x, y) = f(x', y')$. This contradiction following from compositionality's failure shows that given interchangeability, compositionality holds.

The demonstration for pairs extends to n-tuples in general. Suppose that the utility of an n-tuple is not a function of the utilities of its components. Then two n-tuples exist that differ in utility although their elements, in their order, agree in utility. However, the first n-tuple yields the second by iterated interchange of equivalents. According to interchangeability, the n-tuples have the same utility. Hence, given interchangeability, the utility of an n-tuple is a function of its elements' utilities; compositionality holds.

If compositionality fails for composites with n parts lacking an order, then for every order of their parts it fails for some pair of composites having parts that agree in utility according to that order. If it fails for composites of an unequal number of parts, then it fails for some pair of composites with the same number of parts. Hence, the result for n-tuples generalizes. Interchangeability implies compositionality (Westerståhl and Pagin [2011] define compositionality and note its equivalence with interchange of equivalents, given observance of domain restrictions).

Applications of the principle of interchange of equivalents often assume that for a property and combinations of factors the property's value for a factor is the same in all combinations containing the factor. Interchangeability for weight, for example, assumes that an object's weight does not change from one combination of objects to another. The assumption for interchangeability carries over to compositionality. For utility, interchangeability is equivalent to compositionality, assuming the existence of utilities.

Because interchangeability is equivalent to compositionality, the evidence may support compositionality without identifying the function that grounds compositionality. The substitution of equivalent sentence-parts that preserve a sentence's meaning is evidence of meaning's compositionality although the evidence does not identify the general function that generates a sentence's meaning from the meanings of its parts.

3. Complementarity

A property's application to a composite fails to be compositional according to a function just in case the function does not yield the property's application to the composite given its application to the composite's parts. In that case, with respect to the property and function, the composite's parts are *complementary*. Complementarity opposes compositionality, but because it holds among a composite's parts rather than in a set of composites, only a general form of

https://doi.org/10.1017/apa.2014.15 Published online by Cambridge University Press

complementarity rules out compositionality. Given a property, if for every function, the parts of some composite are complementary, the property is not compositional.

Consider utilities attached to economic goods (or to propositions concerning their possession). Take pairs of gloves. The utility of a matched pair of gloves is not the sum of the utility of the left glove and the utility of the right glove, taking each glove in isolation. Accordingly, the gloves are complementary with respect to the property of utility, understood to apply to items in isolation, and with respect to the addition function. In fact, the utility of a matched pair of gloves is not any general function of the utilities of individual gloves. Suppose that such a general function were to state the utilities for pairs of gloves. Then a left glove whose utility alone is the same as the utility alone of the right glove in a matched pair may replace the right glove, and the utility of the resultant unmatched pair of gloves would be the same as the utility of the original matched pair. However, this is not the case. In a matched pair the right glove complements the left glove with respect to the property of utility and any general aggregation function for utilities. The utility of a pair of gloves x and y is not a general function of the utilities of gloves, each taken alone. In a pair, the gloves' utilities do not settle the pair's utility because it matters whether the gloves match. A pair x and y's complementarity with respect to a general function implies that $U(x, y)$ does not come from $U(x)$ and $U(y)$ according to the function. The complementarity of some pair of gloves for each general function refutes the compositionality of the pairs' utilities. In some cases, the utility of a combination of economic goods is not the sum of the utilities of the goods when each is taken alone. The utilities of bundles of economic goods are not in general compositional.

4. Context

A property's compositionality may hold relative to a context, and even if a property's compositionality is independent of context, its value for a component and that value's effect on its value for a composite may depend on context, including the composite's other components. Can attention to context resurrect utility's compositionality despite combinations with complementary components?

Utilities of economic goods are context-sensitive. Ice cream on a hot day has more utility than ice cream on a cold day. Can contextual effects preserve compositionality for pairs of gloves? A defense of compositionality may hold that the utility of a right glove in the context of a matched pair is greater than the utility of that glove alone. Although a left glove alone may have the same value as the right glove alone, by itself alone it does not have the same value as the right glove in the context of the matched pair. Hence its utility does not substitute for the right glove's utility as an argument of a general function that yields the matched pair's utility.

Economics rejects this defense of compositionality because, where economic interest is concerned, compositionality needs practical value. Fine-grained decomposition appealing to contextual effects does not have practical value. No economic consequences follow if the utility of the right glove in a matched pair combines with the utility of the left glove in the matched pair to yield the matched pair's utility. For the purposes of economics, each glove must have the same utility in multiple contexts. A glove's utility is not interesting if it attaches to the glove in a particular context only. Economics attaches utilities to goods that may appear in many contexts so that it may measure utilities in one context and use them for evaluation of commodity bundles in another context. It has no use for a context-sensitive utility of a right glove. No general function with economic interest grounds decomposition of the utility of a matched pair of gloves into the utilities of the gloves making up the pair.

Philosophy has a theoretical interest in contextually dependent forms of compositionality. When context affects a property's application to objects, compositionality specifies contextual features that are assumed to be constant for combinations and their parts. Compositionality holds for a property's application to certain combinations of items under a general function with respect to certain constant contextual features. That is, compositionality licenses substitutions for an item that preserve the property's application to the item, assuming that the substitutions preserve the contextual features compositionality specifies. The utility of possession of a pair of gloves in a context may decompose into the utilities of possession of each glove in the context.

Finely individuating a property's objects may replace specification of constant contextual features. A common type of fine graining takes a utility's object to be a proposition, in fact, a proposition understood a certain way. The fine individuation may put contextual features in the objects themselves. Then compositionality may hold without direct reliance on context. For example, if a utility attaches to the possession of a right glove as part of possession of a matched pair, so that the context is built into the utility's object, then substitution of a left glove that has the same utility alone as the right glove has alone is not substitution of an equivalent utility; for an equivalent utility the substitution needs a glove that makes the same contribution to a matched pair, that is, another right glove.

Attention to the productivity of combinations may also replace context as a means of preserving compositionality. The elements of a pair may interact to produce a good without making the components of a combination containing the pair complementary. If the combination contains the pair and the good it produces, then its components need not be complementary. Compositionality may hold for combinations that include goods their parts generate. The combination of a right glove and a matching left glove may be productive. Each glove may have a constant utility, alone or in pairs, but a pair's utility may be a sum of the gloves' utilities and the utility of their combination's product, call it a match. The pair's utility may be a sum of the utilities of its parts if the match counts as a part of the pair.

Compositionality's reliance on context, fine-graining, and productivity move it toward the trivial dependence a property's application to a whole has on the features of the whole. Philosophically interesting compositionality stops short of that trivial dependence because it characterizes a whole's parts and the general manner in which the property's application to them settles the application of the property to the whole. Such a characterization has explanatory power. Moreover, an interesting dependence on context may specify the relevant features of context, and thus not every difference between cases counts as a difference in relevant context.

5. Arrangements

A property's compositionality entails that the property's application to the composite is a function of the property's application to the parts of the composite. A function's arguments have an order. A composite's order of parts, if the composite gives them an order, provides the order of the function's arguments. A function for the meaning of a sequence of words uses their order in the sequence to assign their meanings to the function's argument places. The order of words in a sentence settles the order of arguments in the function that yields the sentence's meaning. 'John kicked Bill' does not mean the same as 'Bill kicked John' because word order differs in the two sentences and thus changes the order of arguments in the function that goes from the meanings of the words to the meanings of the sentences the words compose.

For a kind of utility U applied to an ordered composite $(x_1, x_2, \ldots, x_n)$, compositionality claims that $U(x_1, x_2, \ldots, x_n) = F(U(x_1), U(x_2), \ldots U(x_n))$. This equation does not assert that the utilities of a whole's parts settle the utility of the whole. It recognizes that the order of the parts may contribute to the utility of the whole. The composite may present events in their temporal order. The events' temporal order may affect the composite's utility. Taking an introductory logic course before an advanced logic course is better than taking the courses in the reverse order. The order of courses in a sequence settles the order of arguments in the function that obtains the sequence's utility from the utilities of the courses. If composites order parts, compositionality uses a function that incorporates their order.

A common objection to a property's compositionality holds that the property's value for a whole depends not only on its value for the parts but also on the order of the parts. A literary critic may make this point about the aesthetic value of a novel and its parts. The objection that compositionality fails when the order of parts matters is invalid because compositionality regards the order of the parts. Their order yields the order of arguments for an aggregation function. Compositionality does not claim that a composite's value is independent of the order of its parts. Given a property's compositionality, the property's application to the whole's parts in the order that the whole imposes settles the property's application to the whole.

In some cases of compositionality, the order of parts does not matter. A treatment of utility's compositionality may take a composite as a shopping basket of goods that an ordered n-tuple represents. If the order of goods does not affect the basket's utility, the order of elements in the n-tuple is insignificant, and an unordered set may represent the basket of goods. However, if the order matters, compositionality may use it.

To make prominent the order of parts, an application of compositionality may identify composites using an order of their parts. For example, the ordered pair (x, y) identifies a composite whose components are x and y in this order. The formula $U(x, y)$ uses the ordered pair (x, y) to stand for the composite that x and y form. It uses the components in an order to represent the composite. Compositionality's equation $U(x, y) = F(U(x), U(y))$ uses the order of variables in the pair (x, y) to establish an order of utilities for the application of F. According to compositionality, the utilities of a composite's components, in the order the composite gives the components, settle the utility of the composite.

For utility, arrangements besides temporal order may matter, such as one event occurring near another event. Then an ordered n-tuple does not adequately represent a composite. It neglects arrangements more complex than an order. Compositionality may fail when those complex arrangements affect the value of a composite. Generalizations of compositionality may obtain a property's application to a whole from the property's application to its parts in the arrangement they have in the whole, or they may weaken compositionality so that instead of a function it advances a relation that restricts a property's application to a whole given the property's application to the parts of that whole.

In this paper I do not pursue such generalizations of compositionality because attention to a part's context may substitute for them. A function going from the utilities of events to the utility of their combination may use temporal order to assign events to argument places and then use other arrangements of the events as a context to settle their utilities. Context enriches compositionality's resources. Compositionality may hold with respect to a property of variable value for a range of wholes given a division of wholes into parts and given fixed contextual features. According to compositionality for utility, given a context, the utilities of a composite's components in their order settle the utility of the composite.

A possible world is a composite that carries its own context. Its parts' utilities may take account of the parts' arrangement. A world's parts may include the complex arrangement of the world's parts. The function for a world's utility then does not need the parts' arrangement in addition to the parts' utilities. It may derive a world's utility from the utilities of its parts alone.

6. Independence

Opposing a property's complementarity relative to a function is a type of independence relative to a function and, as for complementarity, holding among objects rather than among the property's applications to them. Showing this type of independence generally shows compositionality because the absence of complementarity entails compositionality, as noted above in section 3. To illustrate, consider utility for pairs. Objects x and y are *independent* with respect to utility and the function F if and only if $U(x, y) = F(U(x), U(y))$. This type of independence for all pairs with respect to the same function F entails utility's compositionality for pairs. Independence for all pairs entails complementarity for no pairs and therefore compositionality for pairs.

Conditional utility offers a means of defining another type of independence for utility. For propositions x and y, let $U(x \mid y)$ stand for the utility of x given y, that is, the utility of x under supposition of y (see Weirich [2015: sec. 2.6] for a description of this type of conditional utility). The utility of x is *propositionally independent* of y if and only if $U(x) = U(x \mid y)$. The pair (x, y) has propositionally independent elements if and only if $U(x) = U(x \mid y)$ and $U(y) = U(y \mid x)$. Given this type of independence for all values of propositional variables x and y, the values of x and

https://doi.org/10.1017/apa.2014.15 Published online by Cambridge University Press

the values of y have constant utilities in all their combinations. Such independence justifies making coarse-grained the propositions to which utility attaches. For example, if this independence holds for propositions expressing possession of apples forming pairs, utility may dispense with considering possession of an apple in a context. Possession of an apple's utility in isolation is the same as its utility as part of possession of any pair of apples.

Utility's compositionality for pairs (x, y) entails that $U(x)$ and $U(y)$ in this order settle $U(x, y)$ so that no variation in conditions that does not affect $U(x)$ or $U(y)$ affects $U(x, y)$. Hence for utility values u_1 and u_2, $U((x, y) \mid (z \,\&\, U(x) = u_1 \,\&\, U(y) = u_2))$ is constant for all values of a propositional variable z that yield a consistent condition for (x, y)'s utility. A pair's utility is independent of everything that does not change its elements' utilities. Showing this independence supports compositionality.

Suppose that $U(x, y) = F(U(x \mid y), U(y))$ for all x and y. This is a weak form of compositionality. Then imagine that propositional independence holds for every pair (x, y) so that $U(x) = U(x \mid y)$ for all x. Consequently, $U(x, y) = F(U(x), U(y))$ for all x and y, that is, compositionality holds. Hence, compositionality follows from the weak form of compositionality given propositional independence.

7. Intrinsic Utility

This section introduces compositionality for intrinsic utility of options' worlds, dividing an option's world into the option's consequences and other events. After explaining this application of compositionality's definition, I will argue for intrinsic utility's compositionality. The argument takes intrinsic utility's compositionality as fundamental rather than as derived from another property's compositionality. Its first step establishes a norm of intrinsic indifference, and its second step uses the norm to establish compositionality. Principles of rationality, the nature of intrinsic utility, and the definition of consequences ground the argument.

Intrinsic utility evaluates a proposition's realization using only the proposition's implications. In this sense, it considers the proposition's realization in isolation. A proposition's *intrinsic utility* for a rational ideal agent—that is, one who has no cognitive limits—equals the strength of the agent's intrinsic desire for the proposition's realization. Norms of rationality for intrinsic desires impose constraints on intrinsic utilities.

Intrinsic utility registers intrinsic desires and aversions. Ordinary, comprehensive utility registers both intrinsic and extrinsic desires and aversions and so differs from intrinsic utility. Whereas a proposition p's comprehensive utility $U(p)$ evaluates p's world, a proposition p's intrinsic utility $IU(p)$ evaluates p's a priori implications only. Compare intrinsic utility and comprehensive utility for a typical agent. $U(\text{wealth}) > IU(\text{wealth})$, and $U(\text{smoking}) < IU(\text{smoking})$, because wealth is desired mainly as a means to other things whereas smoking is desired mainly for pleasure and despite its adverse consequences. Comprehensive and intrinsic utility follow different principles. Comprehensive, but not intrinsic, utility follows the principle of expected utility. Hence, U (happiness) nearly equals U (a 99 percent chance for

https://doi.org/10.1017/apa.2014.15 Published online by Cambridge University Press

happiness) whereas *IU* (happiness) is much greater than *IU* (a 99 percent chance for happiness). The chance's intrinsic utility evaluates just what the chance entails. A bet's comprehensive utility typically varies with information about the prospects of winning, but such information typically does not affect the bet's intrinsic utility.

A proposition may represent a possible world by specifying events in the world that matter to an agent. Given finite cares, which I assume, the specification is finite. A world's intrinsic utility equals its ordinary, comprehensive utility because a world's propositional representation implies all the world's relevant features, that is, its realizations of desires and aversions. *Intrinsic difference* between two propositions' realizations considers only the propositions' implications and, for two propositions representing worlds, agrees in a rational ideal agent with ordinary, comprehensive indifference.

Because a world's intrinsic utility equals its comprehensive utility, equality of intrinsic utilities of two worlds entails equality of their comprehensive utilities. Consequently, equality of the intrinsic utilities of two options' worlds justifies comprehensive indifference between the two options. Intrinsic indifference and intrinsic utility regulate, respectively, comprehensive indifference and comprehensive utility for options because the latter follow, respectively, intrinsic indifference and intrinsic utility of options' worlds. This brief account of intrinsic utility suffices for my points here about the structure of intrinsic utility.[2]

Gibbard and Harper (1981: 167–68) define an option's consequences using subjunctive conditionals such as ($o \Box\rightarrow c$), read 'if *o* were the case, then *c* would be the case', and taken to be true just in case *c* is true in the nearest world in which *o* is true. The definition stipulates that an event is an option's *consequence* just in case it would occur if the option were realized and is avoidable by realizing another option: an event *c* is a consequence of an option *o* if and only if $o \Box\rightarrow c$ and $\sim(o' \Box\rightarrow c)$ for some option o'. According to the definition, an option's consequence is a counterfactually dependent event that does not follow from every option's realization. The definition selects as an option's consequences the avoidable events in the option's world, and it excludes events that occur regardless of the option realized but allows for two options having a common consequence. Because an option is avoidable and would occur if it were realized, the definition counts an option as a consequence of itself.

Some accounts of an option's consequences restrict these to certain types of events, such as monetary gains or losses. A restricted account of consequences allows for many options to share consequences so that one may infer a consequence's utility from comparisons of options that yield the consequence. Unless these accounts of consequences treat only cases without complementarity, they undermine compositionality because they do not recognize consequences that

2 The account of intrinsic utility is realist as opposed to constructivist, but the paper's points about compositionality also apply given a constructivist account that defines intrinsic utilities of composites and parts using their ranking according to evaluations of their implications. Buchak (2013: sec. 3.1) reviews the distinction between realist and constructivist (or interpretationist) accounts of utility. Psychology and financial management often use a realist account of utility, whereas economics favors a constructivist account. Theoretical fruitfulness supports a definition of a technical term such as intrinsic utility. Weirich (2015: chap. 3) contains a fuller account of intrinsic utility.

occur because of complementarity. For example, they consider only the monetary consequences of an investment and not whether the investment combined with investments already placed reduces risk. I define consequences comprehensively so that they include events that occur because of complementarity. Sometimes the literature says that this definition finely individuates consequences. It does not count the consequences of a new investment as the same with or without previous investments; it distinguishes the new investment's consequences according to the investment's context and does so finely.

Compositionality for a property holds that the property's application to a composite is a function of its application to the composite's parts. Assuming that composites and parts may be equivalent with respect to the property, compositionality holds if and only if interchange of equivalents holds: composites that differ only in parts that are equivalent are themselves equivalent. If the definition of equivalence uses intrinsic indifference rather than equal intrinsic utilities, then equivalence and interchangeability may hold despite the absence of intrinsic utilities and do so without intrinsic utility's compositionality, but interchangeability and compositionality are equivalent given intrinsic utilities.

A property's compositionality holds for a range of composites. Intrinsic utility's compositionality holds in a decision problem for options' worlds with each option's world divided into the option's consequences and other events, the other events being the same for all options' worlds. An interchange of equivalent sets of consequences moves from one option's world to another option's world. Only such interchanges stay within the range of composites, and such interchanges never generate additional parts. Each world has two parts and nothing more; each option's world contains the option's consequences in the world, the other events in the world, and nothing more.

The argument for intrinsic utility's compositionality uses a model with a cognitively ideal agent in an ideal decision problem with stable grounds for the comparison of options. For intrinsic indifference, the argument advances a requirement of rationality, with rationality understood so that violation of the requirement is inexcusable and blameworthy. According to the requirement, *an ideal agent must be intrinsically indifferent between options' worlds if she is intrinsically indifferent between their consequences*. This norm of intrinsic indifference mandates interchangeability of consequences between which intrinsic indifference holds. Given the satisfaction of this requirement, if options' worlds and options' consequences have intrinsic utilities, intrinsic utility is compositional because it agrees with intrinsic indifference. Intrinsic utility's compositionality requires only that if the consequences of two options have the same intrinsic utilities, then the options' worlds also have the same intrinsic utilities. The norm of intrinsic indifference establishes intrinsic utility's compositionality given the existence of intrinsic utilities.

The argument for the norm of intrinsic indifference uses intrinsic indifference's dependence on a comparison of the implications of propositions and the definition of an option's consequences to ward off counterexamples to the norm. A counterexample identifies two options such that a rational ideal agent is intrinsically indifferent between their consequences but not between their worlds. The argument shows that no counterexample challenges the norm.

Suppose that a dean is deciding whether to hold a graduation ceremony. He is intrinsically indifferent between the consequences of holding the ceremony and the consequences of not holding the ceremony. However, he is not intrinsically indifferent between the world with the ceremony and the world without the ceremony because in the first world but not the second a college tradition continues. A counterexample along these lines is implausible. Because continuing a tradition is a consequence of holding the ceremony, if it makes the dean not intrinsically indifferent between the world with the ceremony and the world without the ceremony, it should also make the dean not intrinsically indifferent between the consequences of holding the ceremony and the consequences of not holding the ceremony, contrary to the example's assumption of the dean's indifference between these consequences. The example makes the dean incoherent and so irrational; it is not a counterexample to the norm of intrinsic indifference.

For an ideal agent, who considers options' worlds, only a difference in their implications justifies not being intrinsically indifferent between the worlds. Suppose that such a difference justifies intrinsic indifference between the worlds. Because the difference entails a difference in the options' consequences, it also justifies not being indifferent between the options' consequences. A plausible counterexample to the norm of intrinsic indifference requires, when intrinsic indifference does not hold between two options' worlds, a difference in the worlds' relevant implications. A plausible counterexample presents a difference in the two options' worlds that is not a difference in the two options' consequences. By the definition of options' consequences, no such difference exists. Any difference in the two options' worlds is also a difference in the two options' consequences. Hence no plausible counterexample arises.

Suppose that $(p \,\&\, q)$ represents an option's world, with q representing the option's consequences and p representing the other events, and that $(p \,\&\, q')$ represents another option's world, with q' representing the option's consequences and p again representing the other events. If the pair of options generates a counterexample to the norm of intrinsic indifference, intrinsic indifference does not hold between $(p \,\&\, q)$ and $(p \,\&\, q')$ despite holding between q and q'. A reason for not being intrinsically indifferent between the options' worlds cannot be a difference in relevant implications of the options' consequences (the agent is intrinsically indifferent between the option's consequences) or a difference in relevant implications of the other events (they are the same for both options); therefore, it must be a difference in relevant implications of the options' worlds that is not a difference in relevant implications of the options' consequences or other events. However, by the definition of an option's consequences, a difference in the implications of $(p \,\&\, q)$ and $(p \,\&\, q')$ is a difference in the options' consequences and so a difference in the implications of q and q'. A counterexample needs a difference in the relevant implications of the options' worlds to explain a failure of intrinsic indifference between the worlds despite intrinsic indifference between the options' consequences. However, the definition of an option's consequences makes any such difference also a difference in the options' consequences. Thus, no plausible counterexample exists.

For an instructive comparison, consider aesthetic value. The aesthetic value of a pair of paintings is not a function of the aesthetic values of the paintings. Two paintings that have the same values, respectively, as two others may make a better pair because they harmonize, whereas the others do not. Their harmony is not a feature of either painting, and their evaluations omit it. Compositionality fails for aesthetic value because nothing assigns the two paintings' harmony to either painting. In contrast, compositionality holds for intrinsic utility because a relevant event of an option's world belongs to the option's consequences if it does not belong to the world's other events. The intrinsic utilities of an option's consequences and other events overlook no relevant event in the option's world.

Some theorists may object that giving an option's consequences a comprehensive definition trivializes the norm of intrinsic indifference because fabricated consequences ensure compliance in all cases. However, suppose that an ideal agent who cares only about money is not indifferent between the option of gaining a one-dollar bill and the option of gaining instead another one-dollar bill (this does not entail that she prefers one option to the other); she is not intrinsically indifferent between the options' worlds. Because the agent is intrinsically indifferent between the options' consequences, the agent violates the norm of intrinsic indifference. What fabricated consequence discredits the violation? The options' monetary consequences are the same, and their nonmonetary consequences are not relevant, so the violation stands. Because only a difference in the two options' relevant consequences discredits the agent's intrinsic indifference between their consequences, satisfying the norm is not trivial.[3]

8. Benefits

The norm of intrinsic indifference and intrinsic utility's compositionality bring two benefits. First, intrinsic utility's compositionality assists an outsider's inference of an agent's utility assignment. Second, the norm of intrinsic indifference justifies for an agent an efficient general method of settling indifference between two options.

Suppose that an outsider knows an agent's intrinsic utilities are compositional and that the agent is intrinsically indifferent between the consequences of two options. The outsider then infers that the agent is indifferent between the options' worlds and so between the options, and therefore assigns the options equal comprehensive utilities. Compositionality grounds the outsider's conclusions about the agent's comparison of the options.

Norms of rationality for ideal agents entail compositionality for intrinsic utility. Whether intrinsic desire is compositional for humans is an empirical matter. Huang and Bargh (2014) study motivational structures involving goals that resemble intrinsic desires. In some cases, intrinsic desire may be approximately

3 Weirich (2015: chap. 6) argues that addition is the function that establishes intrinsic utility's compositionality for an option's world divided into the option's consequences and other events.

compositional. Then it assists inferences about a person's assignment of comprehensive utilities.

Next, consider indifference. How may an agent compare two options with the purpose of either adopting or rejecting indifference between them? The agent may compare the options' worlds, but a simpler form of deliberation uses the options' consequences. This simple deliberation recommends indifference between the options if intrinsic indifference holds between their relevant consequences. The norm of intrinsic indifference justifies this procedure. In fact, it requires indifference between options given intrinsic indifference between their consequences because the latter requires intrinsic indifference and thus indifference between the options' worlds.

Nonideal agents following the procedure may irrationally become indifferent between two options. If such an agent should not be indifferent between the two options, then the agent should not have been intrinsically indifferent between their consequences. The norm of intrinsic indifference imposes a requirement, not of nonconditional rationality, but of conditional rationality, granting the condition of intrinsic indifference between the options' consequences. For an ideal agent, who is rational except possibly in the formation of indifference between the two options, rationality given the condition is, however, the same as nonconditional rationality.

The norm of intrinsic indifference regulates formation of intrinsic indifference among option's worlds, and then indifference among options in cases where comparisons of options are incomplete. An agent becomes indifferent between two options because she has intrinsic indifference between their consequences. Her shortcut, bypassing options' worlds, reproduces the results of more extensive deliberation. If she were to form intrinsic utilities for the options' worlds, the options' consequences, and other events, then her intrinsic-utility assignment would be compositional, given that she is a rational ideal agent or imitates such an agent. Therefore, respecting intrinsic indifference, the intrinsic utilities of the options' consequences would be equal and by interchange the intrinsic utilities of the options' worlds would be equal; intrinsic indifference would hold between the options' worlds so that indifference holds between the options. Foreseeing these hypothetical results, the agent may become indifferent between the options without forming intrinsic utilities.[4]

The shortcut is especially valuable when ignorance impedes utility assignments to worlds. An agent comparing two options may know the options' relevant consequences but not know the other events in the options' worlds. Despite ignorance of the options' worlds, the agent may still rationally form indifference between the options using only her indifference between the options' consequences.

An agent's comparison of options may use a narrow evaluation that ignores events aside from consequences. For more efficiency, an evaluation may ignore all but basic consequences, granting that their intrinsic utilities generate the intrinsic utilities of nonbasic consequences. If in a special case the options have common

4 If a rational ideal agent's information is sparse so that deliberation would lead to a set of intrinsic utility assignments rather than a single assignment, each assignment would be compositional and agree with the application of the norm of intrinsic indifference.

https://doi.org/10.1017/apa.2014.15 Published online by Cambridge University Press

consequences, a maximally efficient evaluation ignores them too (see Weirich [2015: chap. 6] for an account of an option's basic and nonbasic consequences).

Intrinsic indifference between options' consequences may settle some decision problems. If only two options are reasonable candidates, and they differ only in equivalent consequences, then they are equivalent and both are rational choices. If all options in a decision problem are equivalent, then each is a solution.

As this section shows, the norm of intrinsic indifference and intrinsic utility's compositionality make significant contributions to decision theory. They enrich utility theory and solve some decision problems.

PAUL WEIRICH
UNIVERSITY OF MISSOURI
weirichp@missouri.edu

References

Buchak, L. (2013) *Rationality and Risk*. Oxford: Oxford University Press.

Joyce, J. (1999) *The Foundations of Causal Decision Theory*. Cambridge: Cambridge University Press.

Krantz, D., R. D. Luce, P. Suppes, and A. Tversk. (1971) *The Foundations of Measurement*, vol. 1. New York: Academic Press.

Gibbard, A., and W. Harper. (1981) 'Counterfactuals and Two Kinds of Expected Utility'. In W. Harper, R. Stalnaker, and G. Pearce (eds.), *Ifs: Conditional, Belief, Decision, Chance, and Time* (Dordrecht: Reidel), 153–90.

Huang, J., and J. Bargh. (2014) 'The Selfish Goal: Autonomously Operating Motivational Structures as the Proximate Cause of Human Judgment and Behavior'. *Behavioral and Brain Sciences*, 37, 121–35.

Savage, L. ([1954] 1972) *The Foundations of Statistics*. New York: Dover.

Weirich, P. (2015) *Models of Decision-Making: Simplifying Choices*. Cambridge: Cambridge University Press.

Westerståhl, D., and P. Pagin. (2011) 'Compositionality'. In C. Maienborn, K. von Heusinger, and P. Portner (eds.), *Semantics: An International Handbook of Natural Language Meaning*, vol. 1. (Berlin: De Gruyter Mouton), 96–123.