Chapter 27
Formal Investigations of Value

Sven Ove Hansson

Abstract We can express values in three major ways: in terms of classification ("good", "bad", "best", etc.), comparison ("better", "at least as good", "equal in value"), and quantity (numbers are assigned). The interrelations among these three types of value expressions are surveyed, with a particular emphasis on relations of interdefinability. Furthermore, interrelations between value terms and terms expressing norms or choices are explored. Several of these connections have been surprisingly little studied, and further investigations may possibly lead to the discovery of additional connections among the different formal representations of value and value-related concepts.

27.1 Introduction

Example 1
CUSTOMER: Can you say something about the quality of these two wines, the Argentinian and the South-African one?
WAITER: Well the Argentinian wine is quite good but the South-African one is better.
CUSTOMER: So the South-African wine is the best of the two?
WAITER: No, that is not what I said. The Argentian wine is best of the two.
CUSTOMER: I am sorry but I cannot make sense of what you are saying.

Example 2
"I need to buy a new car. There are three options that I choose between, a Volkswagen, a Volvo, and a Peugot. I compared the first two and found the Volkswagen to be better than the Volvo. Then I compared the Volkswagen to the Peugot and concluded that the Peugot was better than the Volkswagen. But then I started to think about the Volvo again, and I couldn't avoid the

S. O. Hansson (✉)
Division of Philosophy, Royal Institute of Technology (KTH), Stockholm, Sweden
e-mail: soh@kth.se

S. O. Hansson, V. F. Hendricks (eds.), *Introduction to Formal Philosophy*, Springer Undergraduate Texts in Philosophy, https://doi.org/10.1007/978-3-319-77434-3_27

conclusion that the Volvo is better than the Peugot. So which car should I buy? I just can't make up my mind."

Example 3

UNHAPPY WIFE: Now that I have told you about all the problems in my marriage, do you recommend me to divorce?

MARRIAGE COUNSELLOR: No, my advice is to stay with your husband in spite of his faults.

UNHAPPY WIFE: So you think that it would be better for me to stay with him than to apply for a divorce.

MARRIAGE COUNSELLOR: No, it would be worse. But I nevertheless think that it is what you ought to do.

As these examples show, we have expectations that our value statements should cohere with each other. The second example also shows that we expect a rational person's choices to cohere with her values, and the third that we expect her norms and her values to form a coherent whole. These coherence issues are also important in moral philosophy. As one example of this, utilitarians and deontologists have different views on the exact nature of the required coherence between norms and values.

Formal representation has turned out to be indispensable if we wish to account in a precise way for coherence in issues such as these. However, it must be emphasized that the values held by a human being are inseparably connected with other components of her mind, such as her beliefs and her emotions. The very process of isolating her values from the rest of her mind involves a considerable idealization, and when these isolated values are expressed in a formal language we take the idealization one step further. Therefore, we should not expect to find a single, correct formalization. Instead, we should expect different formalizations to be suitable for capturing various features of what may be called the value-component of her state of mind.

This being said, there are a number of well-established representations, in particular preference relations, value functions, and choice functions, that have turned out to be useful for a wide variety of purposes. These devices are used not only in philosophy but even more in economics, psychology, and the decision sciences. Their most common use is to express what rationality demands of a person's values (and similarly of her norms and her choices). This chapter will provide an overview of these and some other representations, with a strong emphasis on how they relate to each other and in particular on whether they can be defined in terms of each other.

27.2 Values, Facts, and Norms

The separation of facts from values, and the principle that no "ought" can be derived from an "is", belong to the standard messages of elementary philosophy teaching. This exemplifies a general type of logical issues that can be raised for any two categories of statements. We can ask whether two such categories are logically

separable, so that no element of one of them can be logically derived from elements of the other. Contrariwise, we may ask whether the two categories are interdefinable, so that for any element of one of them there is a logically equivalent element of the other. Obviously, it is also possible for such definability to go only in one direction.

Two categories for which rather subtle issues of this nature arise are those of norms and values. A normative expression such as "You ought to exercise two hours every day" prompts or commends some course of action. An evaluative expression such as "The best you can do is to exercise two hours every day" does not prompt or commend. Doing the best may for instance be a too demanding recommendation. An evaluative sentence may contextually imply advice or requirements, but that is not part of what it inherently means [11], [13, p. 143].

Terminological ambiguity often makes it difficult to uphold these distinctions. The terms "norm" and "normative" are sometimes used to cover both types of moral expressions. There is nothing wrong with such a terminological practice, as long as the distinction is made by some other linguistic means.[1]

Once this distinction has been made, it will be seen that the fact−value and is−ought delimitations do not coincide. Although "fact" and "is" denote the same category, "ought" refers to norms which is a separate category from that denoted by "value". Interesting issues of logical relationships arise among all three of these categories.

27.3 Varieties of Definability

George Edward Moore [20, pp. 172–173] pointed out that in spite of being different in meaning, a normative and an evaluative expression may be extensionally equivalent. In particular, a moral theory may imply a specific connection between values and moral requirements. However, it is important to distinguish between those relationships among concepts that hold according to a particular moral theory and those that hold conceptually. Moral standpoints may be supported by different kinds of arguments, but we should not expect substantial normative conclusions to follow from the structure of our concepts. In this chapter, the focus will be on conceptual connections that do not depend on the types of standpoints that tend to differ among moral theories. However, even on that general level it is important to distinguish between what we can call *definability* and *determinability*. The difference is that definability requires intensional equivalence whereas determinability only requires extensional equivalence. The word "bachelor" is definable in terms of "married" and "man" since there is an expression with these two words whose meaning coincides with that of "bachelor". The word "Stockholm" is determinable, but not definable in terms of "capital of Sweden" since these two expressions only have the same reference (for contingent reasons) but not the same meaning.

[1] The notion of supererogation, i.e. doing more good than what is morally required, is a particularly interesting case. It appears to be a composite concept that cannot be adequately explained without reference to both values and norms [15].

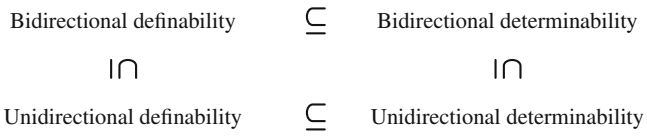

Fig. 27.1 The different types of definability and determinability referred to in the text

Since we are concerned with sets of expressions, such as value expressions, normative expressions etc., this distinction will have to be explicated for such sets. Let A and B be two sets of expressions. Then A is definable in B if and only if for every expression in A there is some expression in B that has the same meaning. Furthermore, A is determinable in B if and only if for every expression in A there is some expression in B that has the same reference. Obviously, definability implies determinability, but not the other way around.

For a simple example, let A be the English insect names and B the (scientific) Latin insect names. Then (if we disregard some minor ambiguities) A is definable in terms of B since for every English insect name there is a Latin insect name with the same meaning. However, the relationship does not go in the other direction since quite a few Latin insect names lack an English equivalent. This is a case of *unidirectional* definability. By *bidirectional* definability is meant that each of two sets of expressions is definable in terms of the other. A similar distinction can be drawn between unidirectional and bidirectional determinability.

Obviously, interdefinability is the ideal, and one might well ask whether definitions not complying with that standard should at all be considered. However, a connection between two categories of expressions (such as norms and values) can be philosophically interesting and/or practically useful although it is not derivable from purely conceptual knowledge, or works only in one direction. Therefore, all the four types of relations of definability and determinability specified in Fig. 27.1 are useful in philosophical investigations.

This chapter has its focus on value statements. After some basic specifications have been introduced in Sect. 27.4, three major types of value statements will be introduced in Sect. 27.5 in the form of a "value triangle". The three types are discussed in somewhat more detail in Sects. 27.6, 27.7, 27.8 and their logical interrelations are investigated in Sects. 27.9, 27.10, 27.11, and 27.12. After that, the logical interrelations between value expressions (of all three types) and statements about choice are studied in Sect. 27.13, and their interrelations with statements about norms are treated in Sect. 27.14. The general picture that emerges from these investigations will be briefly summarized in Sect. 27.15.

27.4 Three Basic Specifications

Most of our value statements in everyday life are ambiguous or at least unspecified in several ways. If I tell you that salmon is the best food fish, you may have to ask several questions in order to find out what I mean: Best for whom? Best in comparison to what? (Among the fish we can buy in our local store, or among those that are available anywhere in the world?) Best from what point of view? (Taste, nutritional value, etc.) Let us have a closer look at these three types of specifications.

The subject. Values can be related to persons in at least two ways. We may refer either to what is good or better *according to* a person or to what is good or better *for* that person. The distinction is not always made with sufficient clarity, but it is crucial in many contexts. One example is medical ethics where increasing emphasis on the patient's autonomy has led to a shift from arguments based on what is good for the patient to arguments based on what is good according to her. Both modes of speaking can also be applied to collective agents. In addition, value terms can be used in an impersonal way (that may at least sometimes be interpreted as "good for everyone").

Instead of saying "This is better according to him" we can say simply: "He prefers this." Logicians have often used the term "logic of betterness" when referring to values that are impersonal or assumed to hold *for* a person. The more common term "logic of preference" usually refers to values held *by* persons. However, no logical or otherwise structural differences seem to have been detected between the two types of connections between betterness and a person. A major reason for this is that the logical discourse on preferences does not usually refer to the preferences that actual people have but to the preferences of (idealized) rational agents. This is also the type of preferences that is usually discussed for instance in economics and decision theory.

This practice should be understood against the background that it would be difficult to identify any structural properties of the preferences (or other values) of agents who do not satisfy at least minimal requirements of rationality. We can assume that a rational agent does not both claim that Wagner's music is better than Verdi's and that the music of Verdi is better than that of Wagner. However, irrational agents can be expected to violate this and presumably any other structural requirement that we may wish to impose. This makes the values of (idealized) rational agents much more interesting than those of actual agents. Of course, we need not assume that agents are rational in all respects, only that they have reflected enough on their value statements to avoid certain structural features that further deliberation would show to be untenable.

The objects of evaluation. Most value statements have an (at least implicit) comparison class. It is one thing to say that Emma is a very good sprinter when you are discussing members of the local running club, but quite another thing to say so when discussing who should represent her country in the upcoming Olympic Games. Both in formal and informal accounts of values we need to keep track of the

comparison class (also called alternative set). Quite a few pseudoparadoxes in value theory have their background in unmentioned shifts in the comparison class [12]. But on the other hand, carefully performed and described such shifts can be used to account for changes in values, and as we will see such shifts can also be used as a mechanism for interdefinability between different types of value statements.

Comparison classes can have interesting structural properties. A particularly important such property is *mutual exclusivity*. By this is meant that no two elements can be combined. The comparison class:

{dog owner, cat owner}

does not satisfy mutual exclusivity since it is possible to have both a cat and a dog. The more precisely described comparison class:

{dog owner but not cat owner, cat owner but not dog owner, both a dog owner and a cat owner, neither a dog owner nor a cat owner}

satisfies mutual exclusivity. It also satisfies *exhaustiveness*, i.e. it covers all possibilities. For most purposes, formal work is simplified by the use of exhaustive and mutually exclusive comparison classes.

We also need to determine what types of entities the comparison class consists of. Two approaches are common in the philosophical literature. One is to regard the elements as primitive, which means that they have no structural connections with each other. The other is to assume that they are sentences. This is often convenient since sentences representing states of affairs provide a highly versatile representation of both philosophical and mundane subject matter. In what follows, the letters x, y, z will be used to represent elements of the comparison class if they are taken as primitive. When these elements are assumed to be sentences they will instead be denoted by the letters p, q, r.

The evaluative viewpoint. Value statements can be made from different points of view, and they are therefore always ambiguous to the extent that the point of view has not been specified. The best car on sale is not necessarily the best car for me to buy. A good philosopher may be a bad mother, etc. There are at least three major ways in which such standards can be specified.

Many such specifications can be interpreted as positing a goal, such that the value terms refer to the achievement of that goal. We can for instance say that something is good from an ethical, economical, environmental, or aesthetic point of view. Something is "morally good" if it is good for satisfying our moral commitments and aspirations, "economically good" if it is good for achieving economic goals, etc. Such goals can be specified to different degrees and in different directions. The best car from the viewpoint of fuel economy may not be best from the viewpoint of total cost per kilometre.

Another way to disambiguate an evaluation is to mention one of the categories to which the evaluated object belongs. I have a friend who can be described as a good pianist but a bad driver. The two expressions refer to the same person, but evaluated

according to our criteria for different categories that she belongs to, namely those of pianists respectively drivers. Such, category-specified value statements are quite common, and they are precise to the extent that we have determinate criteria for the categories in question [14].

As a limiting case, value statements may be intended to include all aspects, i.e., represent an evaluation that takes everything into account ("synoptic" values). It is contentious whether moral values and synoptic values coincide or whether the synoptic values are a broader or more over-arching category that includes non-moral values as well.

If vacillation between value criteria is allowed, then counter-examples can be constructed against any structural condition for value terms that we may think of. ("Rocky is the best saddle horse in the village, but not the best workhorse." Therefore, something may be both best and not best at the same time.) For formal analysis to be meaningful, we have to assume *criterial constancy*, i.e. the viewpoint of evaluation should be the same for all evaluations under consideration.

27.5 The Triangle of Value Concepts

In his book about the logic of probability, Rudolf Carnap distinguished between three major types of empirical descriptive terms. A *classificatory* concept such as "warm" divides objects into mutually exclusive classes. A *comparative* concept such as "warmer" compares two objects to each other. Finally, a *quantitative* concept such as "temperature" characterizes objects by assigning numerical values to them [3], [cf. 17]. The same three categories can be used to classify the value terms.

Among the classificatory value expressions we find those articulated with terms such as "good", "very bad", "almost worst", "fairly good", and "worst", all of which have a single referent that they identify as element of a class. In the formal language they are represented by monadic (one-place) predicates, such as G for "good" and B for "bad". The formula Gx means "x is good", and Bx means "x is bad".

Comparative value expressions such as "better", "worse", and "equal in value to" describe the relation between two referents. In the formal language they are expressed with dyadic (two-place) predicates. In what follows we will use $>$ for "better", $\sim$ for "equally good as", and $\geq$ for "at least as good as". Thus $x > y$ means that x is better than y, and $y \sim z$ that y and z are equally good. (A common alternative notation uses P instead of $>$, I instead of $\sim$, and R instead of $\geq$.)

Quantitative value expressions represent a referent's amount of value in numerical terms, i.e. in numbers saying "how good" something is. Quantitative value is expressed by a numerical function v that takes us from objects of evaluation to real numbers. Thus $v(x) = 3$ means that x has the value represented by the number 3.

In everyday life, moral statements are usually expressed with classificatory or comparative expressions. In moral theory, quantitative valuations are important, primarily since they are required in utilitarianism.

Table 27.1 The three major types of value expressions and their formal representations

Type of value expression	Formal representation	Examples
Classificatory	Monadic predicate	Good, bad, best
Comparative	Dyadic predicate	Better than, equally god as, at least as good as
Quantitative	Numerical function	Utility

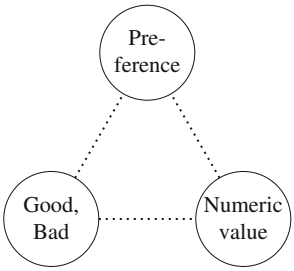

Fig. 27.2 The value triangle, representing the three major types of value statements

The three types of value terms are summarized in Table 27.1 and in the value triangle depicted in Fig. 27.2. The next three sections are devoted to the structural properties of each of these three types. For expository reasons we will begin with the comparative terms.

27.6 Comparative Value Concepts

Preference logic, the logic of the dyadic value predicates, is the most well-developed part of the logic of value concepts.[2] It has a long history. Aristotle discussed structural properties of preferences in Book III of his *Topics*. Representations in modern logical language were developed by Sören Halldén [9] and Georg Henrik von Wright [28].

The two fundamental comparative value concepts are "better" ($>$, strict preference) and "equal in value to" ($\sim$, value equality). The former of these represents both betterness and converse worseness, hence $x > y$ is taken to mean both "x is better than y" and "y is worse than x".

The relation $\geq$, "at least as good as" (weak preference) can be defined in terms of the two fundamental concepts:

$x \geq y$ if and only if either $x > y$ or $x \sim y$.

The three expressions "x is better than y", "y is better than x", and "x is equal in value to y" are usually taken to be mutually exclusive, i.e. no two of them can hold at the same time. It is also assumed that everything is equal in value to itself and

[2]For a more detailed exposition, see Chap. 29.

that equality in value always works in both directions. These assumptions add up to the following four constitutive properties of the comparative notions:

$x > y \rightarrow \neg(y > x)$ (*asymmetry of preference*)
$x \sim y \rightarrow y \sim x$ (*symmetry of indifference*)
$x \sim x$ (*reflexivity of indifference*)
$x > y \rightarrow \neg(x \sim y)$ (*incompatibility of preference and indifference*)

A much more controversial principle is completeness, according to which it holds for any two objects of the comparison class that either one of them is better than the other, or else they are equal in value. This property can be expressed in either of the following two equivalent ways:

$x > y \ \vee \ x \sim y \ \vee \ y > x$ (*completeness*)
$x \geq y \ \vee \ y \geq x$ (*completeness, alternative formulation*)

By far the most discussed postulate for comparative value is transitivity, according to which two steps of weak preference can be combined into one:

$x \geq y \geq z \ \rightarrow \ x \geq z$ (*transitivity*)

(To simplify the notation, we contract series of dyadic predicate expressions, thus writing $x \geq y \geq z$ for $x \geq y \ \& \ y \geq z$.)

Transitivity is often regarded as an essential rationality criterion.[3] The same applies to various weakened versions of it, such as:

$x > y > z \ \rightarrow \ x > z$ (*quasi-transitivity*)
$x_1 > x_2 > \ldots > x_n \ \rightarrow \neg(x_n > x_1)$ (*acyclicity*)

27.7 Classificatory Value Concepts

There is a wide variety of classifying value predicates: “good”, “best”, “bad”, “very good” etc. Here, the focus will be on “good” and “bad” that are denoted by G respectively B. “Good” and “bad” are usually taken to be mutually exclusive, i.e. they cannot consistently both be applied to one and the same object of evaluation. If someone says that a particular novel is both good and bad, then this is perceived as paradoxical. We expect a resolution that typically assigns different evaluation criteria to the two statements, for instance: “The plot is good, but the language is bad”. Due to our assumption of criterial constancy we can presume that goodness and badness are mutually exclusive:

$\neg(Gx \ \& \ Bx)$ (*mutual exclusiveness*)

If the objects of evaluation (elements of the comparison class) are represented by sentences, then additional logical principles can be introduced. In particular, we can express the intuition that a state of affairs and its negation are not (from the

[3] See Chaps. 29 and 31.

same point of view) both good or both bad. If you say "It is good to be married, and it is also good to be unmarried", then you typically mean that matrimony and bachelorhood are good in different respects or according to different criteria. Something similar can be said about the dismal pronouncement "It is bad to be married, and it is also bad to be unmarried." Such equivocations are excluded by the following principles:

$\neg(Gp \,\&\, G\neg p)$ (*non-duplicity of good*)
$\neg(Bp \,\&\, B\neg p)$ (*non-duplicity of bad*)

Two other potential postulates are $Gp \rightarrow B\neg p$ and (symmetrically) $Bp \rightarrow G\neg p$. (For both of them to hold it is sufficient and necessary that $B\neg p \leftrightarrow Gp$ holds.) However, it is easy to show that neither of them is a plausible postulate.

> My uncle is a great music lover. It would be good if I give him a recording of *Das Wohltemperierte Klavier* for his birthday. However, it would not be bad if I do not give him such a recording. This is because not doing so is compatible with giving him some other nice present that he will appreciate.

> Maria is an alcoholic who consumes different brands of whiskey every evening. It is bad that she drank Hazelburn whiskey yesterday. However, it would not have been much of a good thing if she had not done so, since then she would in all probability have taken some other whiskey instead.

Without further devices it seems difficult to obtain any plausible postulates for "good" and "bad" in addition to mutual exclusiveness and non-duplicity. There are at least two devices that we can use to obtain further postulates: shifts in the comparison class and the insertion of "good" and "bad" into a language that also contains a preference relation.

It is easy to find examples in which our usage of "good" and "bad" depends on the context. Jennifer and Robert are both members of the local chess club. Jennifer is one of its best players, but Robert seldom wins a game. When discussing members of the club it would be reasonable to say "Jennifer is a good player, but Robert is not". Suppose that they both join a large competition with several thousand participants, most of whom neither Jennifer nor Robert has much of a chance to defeat. In such a context it would be more natural to count neither Jennifer nor Robert as a good player.

To express this in the formal language we will use capital letters such as A and D to denote comparison classes. These letters can be attached as indices to the monadic value predicates G and B. Thus $G_A x$ means that x is good among the elements of A and $B_A x$ that x is bad among the elements of A. Johan van Benthem [26] has proposed the following postulates for such indexed monadic value predicates:

> If $G_A x \;\&\; \neg G_A y$, then there is no comparison class D such that $G_D y \;\&\; \neg G_D x$ (*non-reversal of good*)
> If $B_A x \;\&\; \neg B_A y$, then there is no comparison class D such that $B_D y \;\&\; \neg B_D x$ (*non-reversal of bad*)

G differentiates between x and *y* in *A* if and only if either $G_A x$ & $\neg G_A y$ or $G_A y$ & $\neg G_A x$. Furthermore, *G differentiates within A* if and only if there are $x, y \in A$ such that *G* differentiates between *x* and *y* in *A*. The corresponding definitions apply to the badness predicate *B*. With these definitions, the following postulates, also proposed by van Benthem, can be introduced:

If *G* differentiates between *x* and *y* in *D*, and $\{x, y\} \subseteq A \subseteq D$, then *G* differentiates between *x* and *y* in *A*. (*downward difference of good*)
If *B* differentiates between *x* and *y* in *D*, and $\{x, y\} \subseteq A \subseteq D$, then *B* differentiates between *x* and *y* in *A*. (*downward difference of bad*)
If $A \subseteq D$ and *G* differentiates within *A*, then it differentiates within *D*. (*upward difference of good*)
If $A \subseteq D$ and *B* differentiates within *A*, then it differentiates within *D*. (*upward difference of bad*)

The other device for obtaining postulates for "good" and "bad" is to include the dyadic and monadic value predicates in one and the same framework. This is intuitively plausible, since our classificatory and comparative concepts appear to be closely connected to each other. This was implicitly recognized already by Aristotle, when he said that "if one thing exceeds while the other falls short of the same standard of good, the one which exceeds is the more desirable" (*Topics*, III:3), which can be interpreted as a statement that:

$Gx \ \& \ \neg Gy \rightarrow x > y$ (*negation-sensitivity of good*)

Other, at least seemingly plausible, connections between the monadic and dyadic predicates include:

$\neg Bx \ \& \ By \rightarrow x > y$ (*negation-sensitivity of bad*)
$Gx \ \& \ By \rightarrow x > y$ (*bivalent sensitivity*)
$x > y \rightarrow Gx \vee By$ (*closeness*)
$Gx \ \& \ y \geq x \rightarrow Gy$ (*positivity of good*)
$Bx \ \& \ x \geq y \rightarrow By$ (*negativity of bad*)
$Gx \ \& \ Gz \ \& \ x \geq y \geq z \rightarrow Gy$ (*continuity of good*)
$Bx \ \& \ Bz \ \& \ x \geq y \geq z \rightarrow By$ (*continuity of bad*)
$Gx \ \& \ x \sim y \rightarrow Gy$ (*indifference-sensitivity of good*)
$Bx \ \& \ x \sim y \rightarrow By$ (*indifference-sensitivity of bad*)

27.8 Quantitative Value Concepts

A numerical function is any function that takes us from some objects to real numbers. In measurement theory, numerical functions are classified according to how much information they carry. Football teams have shirts with numbers on them.

A function that assigns to each football player the number on his or her shirt, for instance $v(\text{Ronaldinho}) = 10$, carries no other information than any other label that could be used for the same purpose. It is called a *nominal* function. Such functions have no use in the representation of values.

Other numerical functions represent an order or rank, so that something can be learnt from which of two objects is assigned the highest value. These are called *ordinal scales*. The ranking of tennis players is an example. The player ranked number 1 is presumably better than that ranked number 2, etc., but the differences on the scale have no significance. Thus the difference between the 1st and the 2nd player cannot be inferred, and it may be very different from that between the 200th and the 201st.

An *interval scale* has uniform differences. A common temperature scale (°C or °F) exemplifies this. The difference between 4 and 5 °C is the same as that between 40 and 41 °C. However, 10 °C is not ten times hotter than 1 °C. Ratios on an interval scale do not carry any meaningful information.[4]

Finally, on a *ratio scale* ratios are also meaningful. Length is measured on a ratio scale. Thus, 10 mm is ten times longer than 1 mm. These lengths stand in the same proportion to each other as 10 to 1 km (which is useful to know when reading a map with the scale 1:1,000,000). The scientific temperature scale is also a ratio scale, thus 300 °K (27 °C) is twice as hot as 150 °K (−123 °C).

The requirements on a numerical function that represents values depends on its intended use. For the purposes of a utilitarian moral theory a ratio scale will be necessary. This makes it possible to add values and to compare the values of aggregated wholes to each other. Other types of moral theories may be less demanding on the value function.

With these definitions in place we can now investigate interdefinabilities among the three categories of value statements. We will begin with the left side of the triangle of Fig. 27.2.

27.9 From Comparative to Classificatory Value

Several proposals have been put forward that define "good" and "bad" in terms of the dyadic predicates. The first such proposal was made by Albert P. Brogan [2], according to whom "good" means "better than its negation" and "bad" means "worse than its negation".

$Gp \leftrightarrow p > \neg p$ (negation-related good)
$Bp \leftrightarrow \neg p > p$ (negation-related bad)

[4]More precisely: The information that we can extract from knowing the exact values of ratios coincides with the information we can extract from just knowing for each ratio whether it is higher than, equal to, or less than 1.

This definition has a strong intuitive appeal, but of course it only works for relata that have a sentential structure so that they can be negated. Another disadvantage is that if G and B are defined in this way, then they do not always satisfy positivity, respectively negativity. For an example, let $\neg q \sim q \sim p > \neg p$. Then Gp, $q \geq p$ and $\neg Gq$, contrary to positivity.

Another major tradition is based on the identification of some neutral object or group of objects. Then "good" can be defined as "better than something neutral" and "bad" as "worse than something neutral". As a general recipe this works for non-sentential as well as sentential objects of comparison:

$Gx \leftrightarrow x > n$ (neutrality-related good)
$Bx \leftrightarrow n > x$ (neutrality-related bad)

Several proposals have been made on how to specify the neutral object(s). Most of these proposals require the objects to be represented by sentences. Some authors have recommended that the neutral propositions should be tautologies [6, p. 37] or contradictions [29, p. 164]. Writing $\top$ for an arbitrary tautology and $\bot$ for an arbitrary contradiction we then have:

$Gp \leftrightarrow p > \top$ (tautology-related good)
$Bp \leftrightarrow \top > p$ (tautology-related bad)
$Gp \leftrightarrow p > \bot$ (contradiction-related good)
$Bp \leftrightarrow \bot > p$ (contradiction-related bad)

However, it is difficult to make sense of a statement saying that something is better or worse than a tautology or a contradiction. If we wish to base our identification of the neutral elements on evaluative comparisons that we can actually make, then the solution must be sought elsewhere.

The most influential identification of neutral elements was proposed by Roderick Chisholm and Ernest Sosa [4]. They defined "good" as "better than something that is equal in value to its negation" and "bad" as "worse than something that is equal in value to its negation". For instance, let us assume that it is (morally) neither good nor bad for a person to read crime fiction. According to this definition, any action that is (morally) better than reading crime novels is a good action. Since Chisholm and Sosa used the term "indifferent" for "equal in value to its own negation", these can be called the "indifference-related" versions of "good" and "bad":

$Gp \leftrightarrow (\exists q)(p > q \sim \neg q)$ (indifference-related good)
$Bp \leftrightarrow (\exists q)(\neg q \sim q > p)$ (indifference-related bad)

Although this pair of definitions is conceptually related to Brogan's negation-related good and bad, the two pairs of definitions do not coincide unless rather strict demands are put on the structure of the preference relation [10]. The two pairs of definitions share the disadvantage of sometimes giving rise to predicates for "good" and "bad" that do not satisfy positivity respectively negativity. The indifference-related definitions also have the additional disadvantage of sometimes giving rise to predicates for "good" and "bad" that do not satisfy the even more elementary postulates mutual exclusiveness and non-duplicity [13, pp. 123–124].

The following definitions were introduced in order to obtain predicates for "good" and "bad" that satisfy these postulates for a wider category of preference relations [10]:

$Gp \leftrightarrow (\forall q)(q \geq^{\star} p \rightarrow q > \neg q)$ (canonical good)
$Bp \leftrightarrow (\forall q)(p \geq^{\star} q \rightarrow \neg q > q)$ (canonical bad)

Here, $\geq^{\star}$ stands for the ancestral of $\geq$. This means that $p \geq^{\star} q$ holds if and only if either $p \geq q$ or there is a series $r_1, \ldots, r_n$ of sentences such that $p \geq r_1 \geq \ldots \geq r_n \geq q$.

Whenever $\geq$ satisfies reflexivity, canonical good and bad satisfy the required postulates for a plausible interpretation of "good" and "bad" (including mutual exclusivity, closeness, non-duplicity of both predicates, positivity of "good", and negativity of "bad").[5] Furthermore, this pair of predicates is a generalization of negation-related good and bad in the following sense: If the preference relation is such that negation-related good satisfies positivity and negation-related bad satisfies negativity, then these negation-related predicates coincide with the canonical ones [13, p. 123].

In summary, we have well-functioning methods for defining the classificatory value terms "good" and "bad" from the comparative ones. We will now turn to the much less discussed issue of defining the comparative terms from the classificatory ones.

27.10 From Classificatory to Comparative Value

The philosophical significance of the above-mentioned definitions of classificatory values in terms of comparative ones has sometimes been put to question. To the extent that natural language can tell us anything about the structure of concepts, it points in the direction of treating classificatory rather than comparative notions as the primitive concepts from which others should be defined. There does not seem to be any natural language in which the classificatory terms are derived from the comparative ones. Instead, derivation in the opposite direction seems to be a universal pattern. As examples of this, the English "better" is believed to originate from a comparative form of a Proto-Indo-European adjective meaning "good", and the French "meilleur" from a comparative form of a Proto-Indo-European word meaning "strong" [25].

According to Henry Kyburg [17, p. 382] "[t]o apply a classificatory term is often to invoke an implicit comparison." But this only holds subject to two important

[5] An even weaker property than reflexivity, namely ancestral reflexivity ($p \geq^{\star} p$), is sufficient for this result.

provisos. First, classifying statements have to be available about more than one object. This is why only the second of the following two statements has comparative implications:

The fish is good in this restaurant.
The fish is good in this restaurant but the meat is not.

Secondly, the comparative implications may depend on the context. Consider again the example in Sect. 27.7 about the two chessplayers, Jennifer and Robert. In the context of the large tournament it is reasonable to say that neither of them is a good player. In the context of the local club we tend to describe Jennifer but not Robert as a good player. If we want to derive a comparison between two objects from classificatory statements about them, then we have to determine the context of these classificatory statements. As the chess-player example illustrates, a smaller context tends to yield more nuances than some of the larger contexts. This gives us a reason to choose the smallest possible context in which classificatory statements about both objects can be made, i.e. the context containing only these two objects.

Using this insight, Johan van Benthem [26] defined comparative concepts in terms of the corresponding classificatory ones as follows:

> x is α-er than y if and only if: In the context $\{x, y\}$, x is α while y is not α [26, p. 195].

This is a general recipe that can be applied to concept pairs such as tall/taller, rich/richer etc. In preference logic we take it for granted that worseness is nothing else than converse betterness. Therefore, this recipe can be interpreted in two ways depending on whether we read $x > y$ as "x is better than y" or as "y is worse than x":

$x > y$ if and only if $G_{\{x,y\}}x$ & $\neg G_{\{x,y\}}y$ (goodness-based preference)
$x > y$ if and only if $B_{\{x,y\}}y$ & $\neg B_{\{x,y\}}x$ (badness-based preference)

It is easy to see that these two definitions are not equivalent and also that neither of them is plausible in all cases. Let x be good and not bad in the context $\{x, y\}$, and let y be neither good nor bad in the same context. Then $x > y$ holds according to first definition but $\neg(x > y)$ according to the second. This seems to be speak in favour of the first definition since we would expect $x > y$ to hold in this case. But next, let x be neither good nor bad in the context $\{x, y\}$, and let y be bad but not good in the same context. Then $\neg(x > y)$ holds according to the first definition but $x > y$ according to the second, which seems to speak in favour of the second definition.

To solve this problem we can replace the goodness- and badness-based definitions by the following one that takes both goodness and badness into account [16]:

$x > y$ if and only if either $G_{\{x,y\}}x$ & $\neg G_{\{x,y\}}y$ or $B_{\{x,y\}}y$ & $\neg B_{\{x,y\}}x$
(bivalently based preference)

Indifference and weak preference can be defined in the same vein:

$x \sim y$ if and only if $G_{\{x,y\}}x \leftrightarrow G_{\{x,y\}}y$ and $B_{\{x,y\}}x \leftrightarrow B_{\{x,y\}}y$

$x \geq y$ if and only if either: (i) $G_{\{x,y\}}x$, (ii) $B_{\{x,y\}}y$, or
(iii) $\neg G_{\{x,y\}}x \ \& \ \neg G_{\{x,y\}}y \ \& \ \neg B_{\{x,y\}}x \ \& \ \neg B_{\{x,y\}}y$

With these definitions we obtain the standard relationship between weak preference, strict preference, and indifference, i.e. $x \geq y \leftrightarrow x > y \ \vee \ x \sim y$. Furthermore, $\geq$ satisfies completeness. If G and B satisfy five of the conditions mentioned in Sect. 27.7, namely mutual exclusiveness, non-reversal of both good and bad, upward difference, and downward difference, then $\geq$ satisfies transitivity [16].

In summary, with this focus on the minimal comparison class it is possible to define comparative values in terms of classificatory ones. Since we have already seen that definitions in the opposite direction are available, this means that the two classes of value terms are definable in terms of each other. However, the two directions of these definitions form a rather disharmonious pair. When we go from the classificatory to the comparative terms we need to have context indices on the classificatory predicates that we start with, but the comparative predicates that we obtain do not come with such indices. When we go in the other direction, from comparative to classificatory predicates, no context indices are obtained for the latter. It remains an open question whether a framework can be constructed in which comparative and classificatory value terms are fully interdefinable.

27.11 Between Quantitative and Comparative Values

We will now turn to the right-hand side of the value triangle, namely that which connects comparative and quantitative value expressions. One direction, namely that from quantitative to comparative values, is easily obtained. For any value function v we can equate preference or betterness ($>$) with having higher value and indifference ($\sim$) with having the same value:

Exact numerical representation:
$x > y$ if and only if $v(x) > v(y)$
$x \sim y$ if and only if $v(x) = v(y)$

It follows directly that the preference relation defined in this way will be complete and transitive.

The derivation of quantitative from comparative value is a somewhat more intricate matter. Suppose that we have an (admittedly strange) preference relation $\geq$ such that $x > y$, $y > z$, and $z > x$. A value function corresponding to this relation would have to be such that $v(x) > v(y)$, $v(y) > v(z)$, and $v(z) > v(x)$, which is clearly impossible. It has in fact been shown that if the comparison class is countable, then a preference relation $\geq$ is reconstructible in terms of a numerical function if and only if it satisfies both completeness and transitivity [21].

One of the most discussed mechanisms for intransitivity is indiscernibility. Consider John who prefers coffee with as little sugar as possible. However, his

ability to taste the difference between cups of coffee with different amounts of sugar is limited. He can only taste the difference if it is more than 0.15 grams. If we present him with the three cups x, with 2 grams of sugar, y with 2.1 grams and z with 2.2 grams, then he is able to taste the difference between x and z, but neither that between x and y nor that between y and z. This will yield a preference relation such that $x \sim y$, $y \sim z$, and $x > z$. Such a preference relation is not representable with a numerical function if we use the exact value representation introduced above. However, it can be represented if we include the limit of discrimination, in this case 0.15 grams, into the representation as follows:

Constant-threshold numerical representation:
$x > y$ if and only if $v(x) - v(y) > \delta$, where δ is a positive real number.

In other contexts, δ in this formula can be interpreted as a limit distinguishing those differences in value that are worthy of consideration from those that are negligible (even though they may be discernible). When interpreted as a discrimination limit, δ is often called a "just noticable difference" (JND). It has been shown that a preference relation over a finite comparison class can be numerically represented with a constant threshold if and only if it satisfies completeness and the following two properties [23]:

$x > y > z \rightarrow (x > w) \vee (w > z)$ (semi-transitivity)
$(x > y)$ & $(z > w) \rightarrow (x > w) \vee (z > y)$ (interval order property)

This construction can be generalized. The most general numerical structure that is still intuitively reasonable is arguably that in which the threshold is allowed to depend on both objects under comparison:

Doubly-variable-threshold numerical representation:
$x > y$ if and only if $v(x) - v(y) > \sigma(x, y)$, where σ is a function such that $\sigma(x, y) > 0$ for all x and y.

It has been shown that a preference relation over a finite comparison class can be numerically represented with a doubly variable threshold if and only if it satisfies acyclicity [1].

In summary, we can easily go from numerical to comparative values, and we can also go in the opposite direction, provided that the preference relation is acyclic and that we use the extra device of a threshold that can be interpreted as a limit of discrimination or of negligibility.

27.12 Between Quantitative and Classificatory Values

Let us now look at the bottom side of the value triangle, and begin with the right-to-left direction, i.e. the issue whether classificatory values can be defined in terms of quantitative ones. Such definitions can be modelled on the definitions in terms of comparative values that were discussed in Sect. 27.8. Thus, the negation-related

definition of "good" and "bad" can be transferred to a quantitative framework as follows [18]:

$$Gp \leftrightarrow v(p) > v(\neg p)$$
$$Bp \leftrightarrow v(\neg p) > v(p)$$

These definitions have the advantage that G and B will always satisfy non-duplicity and mutual exclusivity. However, without restrictions on v they will not in general satisfy positivity respectively negativity.[6] A further problem with this pair of definitions, if unaided by restrictions on v, is that it allows violations of the postulate *bivalent sensitivity* ($Gp \,\&\, Bq \rightarrow p > q$).[7]

The indifference-related definitions of "good" and "bad" can also be transferred to a quantitative framework. We can gain in lucidity (without losing in generality) by assuming that value assignments have been calibrated so that indifferent things have the value zero. Then "good" and "bad" can be defined as follows:

$$Gx \leftrightarrow v(x) > 0$$
$$Bx \leftrightarrow 0 > v(x)$$

This definition is in even greater need than the negation-related one of support from requirements on the structure of v. In particular, unless we disallow v from assigning positive values to both a statement and its negation, or negative values to them both, non-duplicity of G and B will not be satisfied.[8]

In the opposite direction, from classificatory to quantitative values, a simple construction is available provided that G and B satisfy mutual exclusiveness ($\neg(Gx \,\&\, Bx)$). We can then define the numerical function v as follows:

If Gx then $v(x) = 1$
If $\neg Gx \,\&\, \neg Bx$ then $v(x) = 0$
If Bx then $v(x) = -1$

The results of these deliberations are summarized in Fig. 27.3. We have found that definitions are obtainable in both directions along all three sides in the value triangle, although in some cases we needed "tricks" in the form of extensions of the formal apparatus such as context indices. In the rest of this chapter we will consider the connections between these value terms and two other categories of statements that they have often been associated with.

[6] To see that positivity of G does not follow, let $v(p) = v(q) = v(\neg q) = 0$ and $v(\neg p) = -1$. Then Gp and $q \geq p$ but $\neg Gq$.

[7] This can be seen from an example such that $v(\neg q) = 1$, $v(p) = v(q) = 0$, and $v(\neg p) = -1$.

[8] This can be seen from an example such that $v(p) = v(\neg p) = 1$ and $v(q) = v(\neg q) = -1$.

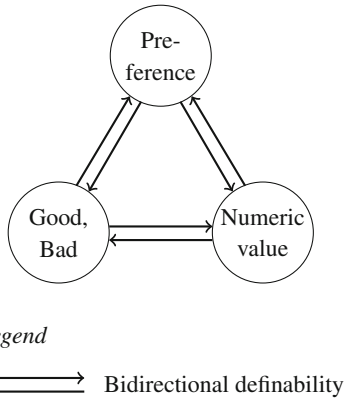

Fig. 27.3 Definability relations in the value triangle

27.13 Choices and Values

Statements about choices refer to actions, whereas statements about values (such as preferences) refer to states of mind. The difference comes out clearly if we consider states of affairs that we cannot choose. I prefer winning € 10,000 in a fair lottery to winning € 5,000 in the same fair lottery, but it is impossible for me to choose winning € 10,000 in this lottery, since if I could make such a choice then the lottery would not be fair.

Admittedly, if we adopt a behaviourist stance according to which states of mind do not exist other than as propensities to act, then preferences can be equated with hypothetical choices, and choices with actualized preferences. But this is a problematic metaphysical standpoint that should not be taken for granted in a formal analysis. Therefore it is advisable to treat choices as belonging to another category than values, from which follows that they are not interdefinable.

But lack of interdefinability does not mean lack of interconnections. We expect rational choices to be guided by preferences. There is something strange in choosing $\neg p$ while preferring p to $\neg p$. Of course there may be reasons to do so, for instance that the preferences in question do not include all the choice-relevant aspects of the alternatives. But some kind of justification is needed in cases like these, and it is certainly worth investigating what it means for choices to be guided by preferences. We should expect preference-guided choices to be restricted by the preferences, perhaps even derivable from them. We may ask what consequences it may have for the structure of choices that they have such connections with preferences, and conversely we may ask what structure preferences should have in order to ensure that their guidance gives rise to choices with desirable structural properties. All these are questions that we can (and should) ask without blurring the distinction between the different categories that choices and preferences belong to.

In order to perform such studies we need a formal representation of (hypothetical) choices.[9] The standard approach is to use choice functions for that purpose:

> C is a *choice function* for a set A if and only if it is a function such that for all B:
>
> (1) If $\emptyset \neq B \subseteq A$, then $\emptyset \neq C(B) \subseteq B$.
> (2) Otherwise, $C(B)$ is undefined.

Various rationality principles for choice functions have been proposed, such as the following:

> If $B_1 \subseteq B_2$ then $B_1 \cap C(B_2) \subseteq C(B_1)$ (Property α, the Chernoff property)

The most obvious way to construct a choice function out of a preference relation $\geq$ is to have the function always choose those elements that are at least as good as anything else that could have been chosen:

$$C(B) = \{x \in B \mid (\forall y \in B)(x \geq y)\}$$

Conversely, from a given choice function C we can construct a preference relation, based on choices from two-member sets:

$$x \geq y \text{ if and only if } x \in C(\{x, y\})$$

The interrelations between choices and preferences that can be obtained with these two definitions have been studied in considerable detail. It turns out that if a preference relation satisfies standard rationality criteria, then the choice function that it gives rise to will satisfy the major rationality criteria for choices (such as the above-mentioned Property α and others in the same style). Conversely, if a choice function satisfies these principles, then the preference relation that it gives rise to will in its turn satisfy the standard rationality criteria for preferences. Furthermore, retrievability holds in both directions: If we use these definitions to go from preferences to choices and then from choices back to preferences, then we regain the preference relations that we started with. Similarly, if we go from choices to preferences and then back to choices, then the original choice function will be regained. We therefore have pathways yielding full interdeterminability between preferences satisfying reasonable (but contestable) rationality criteria and choices guided by these preferences. But as already indicated, it is important to recognize that these are relations of interdeterminability, not interdefinability, since choices and preferences belong to different conceptual categories between which extensional but not intensional equivalence is possible.

[9] See Chap. 29 for additional information on choice functions and their connections with preference relations.

27.14 Norms and Values

In deontic logic, the logic of norms, it is generally recognized that there are three major groups of normative expressions in ordinary language, namely prescriptive, prohibitive, and permissive expressions.[10] In the formal language, they are represented by the corresponding three types of predicates. Here, prescriptive predicates such as "ought", "obligatory", and "morally required" will be denoted by O. Permissive predicates such as "permitted" and "allowed" will be denoted by P, and prohibitive ones such as "forbidden", "prohibited", and "morally wrong" by F. The three types of predicates are standardly and sensibly assumed to be interdefinable in the following way:

$Pp \leftrightarrow \neg O \neg p$ and $Pp \leftrightarrow \neg Fp$
$Fp \leftrightarrow O \neg p$ and $Fp \leftrightarrow \neg Pp$
$Op \leftrightarrow \neg P \neg p$ and $Op \leftrightarrow F \neg p$

The three categories of normative statements form a "norm triangle" with much more simple and direct definitions than those that we needed in our investigations of the value triangle. (See Fig. 27.4.)

As we saw in Sect. 27.2, normative and evaluative expressions belong to different categories in terms of their meanings, and therefore they cannot be interdefinable. But they are nevertheless strongly connected, and we expect them to cohere in some way or other. Therefore it is meaningful to search for possible relations of determinability between the two categories.

The most common proposal for a connection between predicates for norms and values is to identify what ought to be done with the best. This may be called the best-ought connection. It has strong support in the utilitarian camp. G.E. Moore, in a *locus classicus*, identified the assertion "I am morally bound to perform this

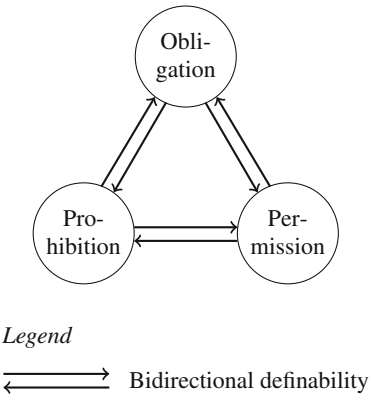

Fig. 27.4 The norm triangle with its interdefinabilities

[10] See Chap. 32 for more information.

action" with the assertion "This action will produce the greatest possible amount of good in the Universe" [19, p. 147]. Another proposal, put forward by Gupta [8] and von Kutschera [27], equates "ought" with "good". This may be called the good-ought connection. Both these proposals equate a prescriptive predicate with a value predicate that satisfies positivity (i.e. a predicate H such that $Hq \,\&\, p \geq q \;\rightarrow\; Hp$). However, all proposals of this kind are highly problematic since they are threatened by counter-examples with the following structure [12]:

(1) p and q are mutually exclusive.
(2) $O(p \vee q)$
(3) $\neg Op$
(4) $\neg Oq$
(5) Either $p \geq (p \vee q)$ or $q \geq (p \vee q)$.

It follows straight-forwardly that if an example of this type can be found for a prescriptive predicate O, then that predicate cannot be equivalent with any positive value predicate. Such examples can indeed readily be found. One way to construct them is to let p and q represent two jointly exhaustive ways to satisfy the same moral requirement, and such that the difference between p and q is morally irrelevant. For instance, p may signify that I pay my debt to Adam by letting Simone bring my money to him, and q that I pay the debt in any other way.

Examples such as this, and others that can be constructed with the same structure, make a negative conclusion inevitable: No prescriptive predicate is extensionally equivalent with any value predicate that satisfies positivity. However, there are other ways to connect norm and value predicates to each other. Two other interesting options are to connect a permissive predicate P to a positive value predicate and to connect a prohibitive predicate F to a negative predicate (i.e. a predicate H such that $Hp \,\&\, p \geq q \;\rightarrow\; Hq$). Interestingly enough, these two options are equivalent. Let O, P, and F be three norm predicates that are interdefinable as explained above. It is easy to show that the following three properties of their connections with a preference relation $\geq$ are equivalent:

(1) P satisfies positivity ($Pq \,\&\, p \geq q \;\rightarrow\; Pp$),
(2) F satisfies negativity ($Fp \,\&\, p \geq q \;\rightarrow\; Fq$), and
(3) O satisfies contranegativity ($Op \,\&\, (\neg p \geq \neg q) \rightarrow Oq$).

One way to make this concrete is to connect a prohibitive term such as "forbidden" or "wrong" with the negative value term "bad" (the bad-wrong connection). Then an action is wrong if and only if it is bad. It ought to be performed if and only if it is bad not to perform it, and it is allowed if and only if it is not bad not to perform it. This definition assigns what seem to be suitable logical properties to the normative terms, but the norm–value interface that it provides is not flawless. A major reason for this is that the words "bad" and "wrong" do not necessarily have exactly the strengths necessary for exact interdeterminability. As was pointed out in another context by Chisholm and Sosa [5, p. 326], there are actions of "permissive ill-doing", i.e. "minor acts of discourtesy which most of us feel we have a right to perform (e.g. taking too long in the restaurant when others are known to be waiting)." Such

acts are arguably morally bad but not morally forbidden. Therefore the bad-wrong connection should only be seen as a very rough approximation. It is, however, an interesting approximation since it provides us with full interdeterminability (though not interdefinability).

27.15 Conclusion

The major conclusions from these considerations are summarized in Fig. 27.5. Interdefinability holds internally among the three different kinds of normative predicates, and also – with some tailoring of the formal structure – among the three types of value terms. There are also connections of interdeterminability both between norms and values and between values and actions. Several of the connections indicated in the diagram, in particular those involving monadic value predicates, have been surprisingly little studied. Further investigations may possibly lead to the discovery of improved definability relations among the different formal representations of values and norms.

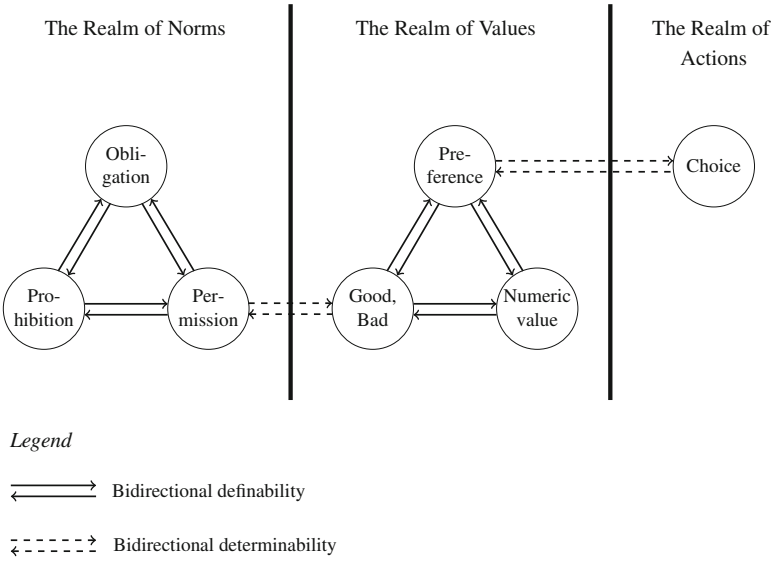

Fig. 27.5 Interdefinabilities and interdeterminabilities among statements belonging to the three realms of norms, values, and actions

References and Recommended Readings

1. Abbas, M. (1995). Any complete preference structure without circuit admits an interval representation. *Theory and Decision, 39*, 115–126.
2. Brogan, A. P. (1919). The fundamental value universal. *Journal of Philosophy, Psychology and Scientific Methods, 16*, 96–104.
3. Carnap, R. (1950). *The logical foundations of probability*. Chicago: University of Chicago Press.
4. Chisholm, R. M., & Sosa, E. (1966). On the logic of 'Intrinsically Better'. *American Philosophical Quarterly, 3*, 244–249.
5. Chisholm, R. M., & Sosa, E. (1966). Intrinsic preferability and the problem of supererogation. *Synthese, 16*, 321–331.
6. Danielsson, S. (1968). *Preference and obligation*. Uppsala, Sweden: Filosofiska Föreningen.
7. ∗ Gabbay, D., Horty, J., Parent, X., van der Meyden, R., & van der Torre, L. (Eds.). (2013) *Handbook of deontic logic and normative systems* (Vol. 1). London: College publications. [Covers most aspects of the logic of normative predicates.]
8. Gupta, R. K. (1959). Good, duty and imperatives. *Methodos, 11*, 161–167.
9. Halldén, S. (1957). *On the logic of 'Better'*. Lund: Gleerup.
10. Hansson, S. O. (1990). Defining 'good' and 'bad' in terms of 'better'. *Notre Dame Journal of Formal Logic, 31*, 136–149.
11. Hansson, S. O. (1991). Norms and values. *Crítica, 23*, 3–13.
12. Hansson, S. O. (1999). Preferences and alternatives. *Crítica, 31*, 53–66.
13. ∗Hansson, S. O. (2001). *The structure of values and norms*. Cambridge: Cambridge University Press. [Detailed discussion of the relationships among preferences, monadic values, and norms.]
14. Hansson, S. O. (2006). Category-specified value statements. *Synthese, 148*, 425–432.
15. Hansson, S. O. (2015). Representing supererogation. *Journal of Logic and Computation, 25*(2), 443–451.
16. Hansson, S. O., & Liu, F. (2014). From good to better. Using contextual shifts to define preference in terms of monadic value. In A. Baltag & S. Smets (Eds.), *Johan van Benthem on logical dynamics* (pp. 729–747). Cham: Springer.
17. Kyburg, H. E. (1997). Quantities, magnitudes, and numbers. *Philosophy of Science, 64*, 377–410.
18. Lenzen, W. (1983). On the representation of classificatory value structures. *Theory and Decision, 15*, 349–369.
19. Moore, G. E. ([1903] 1951). *Principia Ethica*. Cambridge: Cambridge University Press.
20. Moore, G. E. (1912). *Ethics*. London: Oxford University Press.
21. Roberts, F. S. (1979). *Measurement theory*. In G.-C. Rota (Ed.), *Encyclopedia of mathematics and its applications* (Vol. 7). Reading: Addison-Wesley.
22. ∗Rønnow-Rasmussen, T., & Zimmerman, M. J. (Eds.). (2005). *Recent work on intrinsic value*. Dordrecht: Springer. [Highly useful collection of papers on formal and informal value theory.]
23. Scott, D., & Suppes, P. (1958). Foundational aspects of theories of measurement. *Journal of Symbolic Logic, 23*, 113–128.
24. ∗Sen, A. (1970). *Collective choice and social welfare*. San Francisco: Holden-Day. [Unusually clarifying presentations of the relationships between choice and value.]
25. Ultan, R. (1972). Some features of basic comparative constructions. *Stanford Working Papers on Language Universals, 9*, 117–162.
26. van Benthem, J. (1982). Later than late: On the logical origin of the temporal order. *Pacific Philosophical Quarterly, 63*, 193–203.
27. von Kutschera, F. (1975). Semantic analyses of normative concepts. *Erkenntnis, 9*, 195–218.
28. von Wright, G. H. (1963). *The logic of preference*. Edinburgh: Edinburgh University Press.
29. von Wright, G. H. (1972). The logic of preference reconsidered. *Theory and Decision, 3*, 140–169.