SYNTHESE LIBRARY / VOLUME 116

RAIMO TUOMELA

HUMAN ACTION AND ITS EXPLANATION

A Study on the Philosophical Foundations of Psychology

D. REIDEL PUBLISHING COMPANY
DORDRECHT - HOLLAND / BOSTON - U.S.A.

HUMAN ACTION AND ITS EXPLANATION

A Study on the Philosophical Foundations of Psychology

RAIMO TUOMELA
Department of Philosophy, University of Helsinki

D. REIDEL PUBLISHING COMPANY
DORDRECHT–HOLLAND/BOSTON–U.S.A.

Library of Congress Cataloging in Publication Data

Tuomela, Raimo.
Human action and its explanation.

(Synthese library; v. 116)
Bibliography: p.
Including indexes.
1. Psychology—Philosophy. I. Title.
BF38.T86 150'.1 77–24522

ISBN-13: 978-94-010-1244-7 e-ISBN-13: 978-94-010-1242-3
DOI: 10.1007/978-94-010-1242-3

Published by D. Reidel Publishing Company,
P.O. Box 17, Dordrecht, Holland

Sold and distributed in the U.S.A., Canada, and Mexico
by D. Reidel Publishing Company, Inc.
Lincoln Building, 160 Old Derby Street, Hingham,
Mass. 02043, U.S.A.

CONTENTS

To Sirkka

PREFACE

This book presents a unified and systematic philosophical account of human actions and their explanation, and it does it in the spirit of scientific realism. In addition, various other related topics, such as psychological concept formation and the nature of mental events and states, are discussed. This is due to the fact that the key problems in the philosophy of psychology are interconnected to a high degree. This interwovenness has affected the discussion of these problems in that often the same topic is discussed in several contexts in the book. I hope the reader does not find this too frustrating.

The theory of action developed in this book, especially in its latter half, is a causalist one. In a sense it can be regarded as an explication and refinement of a typical common sense view of actions and the mental episodes causally responsible for them. It has, of course, not been possible to discuss all the relevant philosophical problems in great detail, even if I have regarded it as necessary to give a brief treatment of relatively many problems. Rather, I have concentrated on some key issues and hope that future research will help to clarify the rest.

A remark on my syntactic conventions is due here. Generally, symbols may be used autonymously whenever this cannot be expected to cause confusion. As usual, the single quote-operation is used to form names of expressions. Numbered formulas and statements are referred to only by their numbers, when they occur in the same chapter (e.g. formula (*18*)); otherwise they are referred to by the number of the chapter and the formula number (e.g. by (*7.18*) when formula (*18*) of Chapter 7 is meant).

I would especially like to thank Professors Robert Audi, Richard Rorty and Zoltan Domotor, as well as Dr. Lars Hertzberg for discussions concerning some of the topics dealt with in this book. Professor Ilkka Niiniluoto made some helpful remarks on Chapter 11 when I presented the material in my seminar. I wish to thank Mr. Farrell Delman for checking the English of most of this book.

I am very much indebted to the Academy of Finland for research grants which made it possible for me to be on leave from my professorship during 1974 and the latter half of 1976 to work on this project.

With appropriate permission, I have in this book used some passages of the following papers of mine:

'Psychological Concepts and Functionalism', in Hintikka. J. (ed.): 1976, *Essays On Wittgenstein*, *Acta Philosophica Fennica*, vol. 28, North-Holland, Amsterdam, pp. 364–393

'Causality and Action', forthcoming in Hintikka. J. and R. Butts (eds.), *Logic, Methodology and Philosophy of Science V*, Reidel, Dordrecht

'Purposive Causation of Action', forthcoming in Cohen, R. and M. Wartofsky (eds.), *Boston Studies in the Philosophy of Science*, vol. 31, Reidel, Dordrecht

'Dispositions, Realism, and Explanation', *Synthese* **34** (1977), 457–478.

Helsinki, December 1976

Raimo Tuomela

CHAPTER 1

SCIENTIFIC REALISM AND PSYCHOLOGY

Wittgenstein once said that in psychology there are experimental methods and conceptual confusion. This still seems largely true. As evidence for the claim about conceptual confusion it suffices to take a look at almost any psychological work on cognitive psychology (perhaps excluding some of the most recent).[1] It is hoped that this book will bring at least some clarification into the conceptual framework of cognitive and motivation psychology.

Our basic aim in this book is to develop a causal account of complex actions (cf. especially Chapter 10). We shall also discuss the related matter of how to explain actions causally. This latter task again presupposes that we have a relatively good grip on explanatory psychological concepts (e.g., proattitudes, intentions, beliefs). Thus, we shall devote relatively much space to the investigation of the nature of inner psychological states and events, especially those to be called conceptual states and events.

Our development of a causal philosophical theory of complex actions will naturally involve a conceptual clarification of various types of action and of various psychological concepts for inner psychological states and episodes, and most importantly, of the various interrelations between these groups of concepts. The resulting system should give a kind of philosophical and methodological foundation for cognitive and motivation psychology, i.e. roughly for all psychological theories which deal with human actions and their explanation by reference to wants, intentions, beliefs, and related factors. Thus, what we are going to do should be of interest not only to philosophers and methodologists but to psychologists as well. Philosophy of psychology should, and can, be seen as intimately related to the science of psychology.

The general philosophical view that this book espouses could be called *critical scientific realism* (see Tuomela (1973a) for a brief exposition). One way of viewing our present work is then to regard it as an attempt to try and see how critical scientific realism applies in detail to some of the central philosophical problems of psychology (and other related social sciences, too). Even if most of our arguments and developments will be more or less independent of any particular realist assumptions, perhaps at

least some kind of "spirit of realism" is likely to be felt throughout the book. Especially our view of the nature of explanatory psychological concepts as theoretico-reportive concepts (Chapter 4) and our treatment of causality (Chapters 9, 12) will be seen to have a clearly realist flavor.

Our version of critical scientific realism primarily involves a certain view of what there is and of how knowledge of what there is can be obtained. Roughly and concisely speaking, critical scientific realism says that science is about a reality that exists independently of observers. In principle, and contra Kant, this reality is knowable, though normally only in a symbolic, partial and distorted way.

Knowledge about this reality is primarily obtained by means of scientific theorizing and observation, and this knowledge always remains corrigible. A central feature of this view is that there exist objects and properties which are not observable or experiencible at all. We need special theoretical concepts and theoretical discourse to speak about this non-sensible (non-empirical) part of reality.

Our scientific realism can be partially and crudely summarized by the following concise epistemic-ontological thesis:[2]

(R) All non-sentient physical objects and all sentient beings have exactly the constituents and properties ascribed to them by the theoretical scientific terms that are required for the best scientific explanations of the behavior of these entities.

What (R) really comprises can of course only be seen after its philosophical key terms have been explicated. Especially much depends on what the 'best scientific explanation' involves and on how much content is given to 'behavior'. What we understand by the first phrase will roughly become clear later in this book (also see Tuomela (1973a)). 'Behavior' is to be taken in an inclusive sense in our idealized (R). In the case of *persons* (or agents), our primary objects of investigation in this book, this term will be taken to cover, in addition to actions and other overt behavior, also mental states and episodes. Thus the "sensuous" features of the world (e.g. the qualitative features involved in seeing a red rose and in sensing the smell of the rose) will also become included. (This is something, e.g., Sellars (1963) has strongly emphasized.)

Thesis (R) involves the idea that what explains best describes best. As it is scientific theories which explain best we can almost take as our general motto: *scientia mensura*. As Sellars puts it, "in the dimension of describing and explaining the world, science is the measure of all things, of

what is that it is, and of what is not that it is not" (Sellars (1963), p. 173). In all interesting cases the best explaining theories (R) speaks about are yet to be found.

Sometimes and for some purpose it is useful to philosophize by means of highly idealized concepts ("ideal types"). We shall occasionally below employ Sellars' idealized notions "manifest image" and "scientific image" when discussing broader vistas (cf. Sellars (1963)). The manifest image of the world consists, roughly, of "Strawsonian" entities such as tables, trees, stones, colors, sizes, and so on, and of the framework of persons or agents capable of knowing, believing, wanting, hoping, saying, promising, acting, obeying norms, building societies, and, generally, bringing about cultural products. The manifest image can be characterized as the conceptual framework within which (Western) man became aware of himself. Moreover, man is essentially to be regarded as that being which conceives of itself in terms of the manifest image.

In addition to the above epistemic-ontological criterion one can also characterize the manifest image epistemically by saying that within it only correlational inductive methods (say, Mill's methods and many ordinary statistical multivariate methods of data analysis) are used. No postulation of unobserved objects nor of properties "lying behind" the observable ones and explaining their behavior is involved. An example of the use of the postulational method is given by the explanation of the macro-properties of gases (e.g., the Boyle–Charles law) by means of the kinetic theory of gases.

The scientific image can be characterized as consisting of those objects and properties which scientific theories truly postulate to exist. Epistemically (or methodologically) the scientific image can be described by saying that it makes extensive use of the postulational method.

Which of these images is primary? Obviously all scientific realists will say that ontologically or in "the order of being" the scientific image is primary. How about the "order of conceiving", viz. from the point of view of concept formation and gathering knowledge? Here scientific realists differ. Radical ones want to make the scientific image primary even in the order of conceiving, whereas, e.g., Sellars considers the manifest image primary in this respect. When discussing psychology and the social sciences in this book we shall agree with Sellars' view of the primacy of the manifest image in the order of conceiving. However, we emphasize the social and conventional nature of this primacy and the possibility of changing linguistic conventions and habits.

We said earlier that Sellars' notions of the manifest and the scientific image are really ideal types. In fact, it is easy to show that they have several unrealistic features.

First of all, these notions are vague. Concerning the manifest image, we may ask whether, e.g., a distant star which can be seen only through a powerful telescope belongs to the manifest image. How about bacteria seen only through microscopes? They are certainly not postulated, however.

Next, it would be easy to argue that there are several incompatible but "equally possible" refinements of the manifest image (cf. e.g. the incompatibilist versus the reconciliationist explications of the freedom of the will). So, we may ask, what is the use of the definite article in the connection of this notion.

Still another criticism against Sellars' notion of the manifest image is that theorizing within the social sciences in fact typically employs the postulational method even when it is otherwise within the conceptual framework of the manifest image (e.g. statistical factor analysis is often used in a postulational way).

The scientific image is supposed to be an asymptotic unification of the scientific pictures different scientific theories give us at the limit. Now this Peircean "limit science" which is supposed to tell us the final truth about the world is of course something mythical. We do not a priori know how it will look like (and how the final scientific image will be like); we do not even know whether science will grow towards any such limit at all. Furthermore, our present ununified science probably gives a very bad estimate of the final scientific image.

I would also like to point out that there are other ways of forming theories in science than by starting from observable entities and by postulating unobservable entities to explain the behavior of these observable ones. Many physical theories simply do not speak about any observable entities at all and hence do not have any observational content (cf. Tuomela (1973a)).

What my above criticisms show is at least that the notions of the manifest image and the scientific image are problematic and cannot be used in serious philosophical argumentation without qualifications. Still we are going to occasionally employ the truths involved in these notions. Our basic setting will be the following.

We think that psychological theorizing (at least within cognitive and motivation psychology) has to start within the manifest image and to

essentially employ the concepts of the framework of persons. However, on our way to fusing the manifest and the scientific image into a "three-dimensional" picture, we employ the postulational method to introduce inner psychological entities (states, episodes, etc.) to explain overt actions. Future science (presumably neurophysiology and neuropsychology) will then tell us more about these inner psychological entities. It should also be emphasized that this method makes the manifest and the scientific image in part *conceptually* connected. Within a setting like this, the above criticisms against the vagueness and ambiguity of the notions of the manifest and the scientific image can be avoided, it seems. (If necessary, we will, so to speak, relativize these images to research paradigms and programmes, as will in effect be done in Chapters 3 and 4.)

As mentioned earlier, we shall in this book try to create a causal theory of action which applies to arbitrarily complex actions. The various notions of action that we are going to deal with are typical elements of a refined framework of persons (agents). But in our treatment it will turn out that all an agent ever does (in our achievement sense of doing) is to move his body, while the rest is up to nature. Movements of the body seem naturalistic enough not to present great obstacles for unifying the manifest image with the scientific image.

Our causal theory of actions basically claims that intentional action-events are causally brought about by the agent's effective intendings, and, more generally, any action token conceptually involves reference to some kind of antecedent causal mental event or state (see Chapter 10).

It will, not surprisingly, turn out that the acceptability of our causal theory of action is going to depend strongly on the acceptability of a certain kind of *functionalist* account of inner mental events. In general, if one wants to construe an action theory so that it at least does not conflict with an attempt to fuse the manifest and the scientific images, special attention must be paid to the conceptual construal of mental states and episodes. Thus, in fusing the manifest and the scientific images together into a "stereoscopic" view one must give an account of both how "raw feels" (non-conceptual mental events and episodes such as sense impressions, sensations, and feelings) and how "thoughts" (i.e. conceptual events and episodes like perceivings and desirings) are to be characterized. (We assume here that raw feels and thoughts are really understood as inner episodes, etc., and that they are taken to involve some categorical features over and above their various overt-dispositional and overt-categorical ones.)

We are not really going to discuss the philosophical problems related to raw feels in this book. But thoughts (conceptual events and episodes, inner actualizations of propositional attitudes) must be discussed in some detail because of their importance. As said earlier, we are going to construe the concepts representing wantings, believings, perceivings, etc., as theoretical (or theoretico-reportive) concepts. That they can be so construed will be shown in Chapter 4. One essential condition for our construal is that we reject the Myth of the Given (cf. Sellars (1956)). According to this myth words for mental entities are somehow causally and independently of learning given their meanings, due to the impact of some suitable extra-linguistic mental entities (cf. Sellars (1956)). Therefore mental entities (sense data, etc.) would be (externally) fixed and they could not be theoretical entities in any sense.

We are going to construe thoughts (thinkings) *analogically* on the basis of overt speech (cf. Chapters 3 and 4). This account, if successful, will also account for the *intentionality* or intentional aboutness of thoughts. For one can then claim that intentionality is reducible to the *semantical* metalinguistic discourse (for semantic "aboutness") pertaining to a language for overt public entities. This accounts for the conceptual side of intentionality. In the real world (*in rerum natura*) there will be very few exactly statable "criteria", such as necessary conditions, for the non-metaphorical application of the categories of intentionality and thinking. One such elementary criterion is that any "entity" to which one can adequately ascribe thoughts must have some suitable representational sensory mechanism for "picturing" the "facts" of the world; furthermore, it must be capable of self-controlled activity (even if it is hard to be very exact about this).

There is another criterion for intentionality and thinking, which, though vague, is central. It is that the entity can be taken to be "one of us" and to share our intentions and norms and be capable of fulfilling and obeying them. This idea also involves the important insight that what is peculiar to and distinctive about persons is their "rule-following" nature, which can only be accounted for by means of a *normative* or prescriptive language, and it is in this sense that social science and the humanities most basically differ from the physical sciences. On the other hand, in the dimension of describing science can in principle tell us everything there is to be told about persons, societies and cultures. It is roughly in this sense that a nomological social science is possible.

NOTES

[1] To support our claim consider this recent statement:

"Our conceptual framework has grown out of our own research in the attitude area. It emphasizes the necessity of distinguishing among beliefs, attitudes, intentions and behavior; four distinct variables that have often been used *interchangeably* in the past." (Fishbein and Ajzen (1975) p. vi, my italics).

[2] The kind of scientific realism espoused by this work on the whole comes close to the views of Wilfrid Sellars (see, e.g., Sellars (1963), (1967), (1968), and (1974). The thesis (R) below would probably be acceptable to Sellars, too, although I perhaps want to emphasize its *regulative* and *idealized* character more than Sellars does.

My (R) can be interpreted to amount to the *Sellarsian scientific realism* of Cornman (1971). Note that (R) should be taken to concern only *contingent* descriptive properties.

CHAPTER 2

HUMAN ACTION

1. Actions as achievements

As we all know human behavior can be described and classified in a variety of ways. We speak about activities, actions, achievements, habitual and automated behavior, reflexes, and so on. It is plausible to think that our explanations of behavior are strongly linked with our ways of describing behavior. We shall in fact argue later that one's general views concerning the nature of man and the sources of man's activity strongly affect one's way of conceptualizing and explaining behavior.

In this book we shall concentrate on behavior as action and on some of the various ways one can explain actions. This chapter is in part devoted to a preliminary clarification of some key concepts related to action.

Among actions we customarily include behavior such as the following: raising one's arm, opening a door, crossing a street, paying a debt, killing somebody, saying or promising something, refusing or forbearing to do something, and so on. These actions can be done rationally, intentionally, voluntarily, deliberately, etc.

Actions must of course be distinguished from bodily movements (and reflexes), even if normal actions in a sense consist of bodily movements. Paying a debt by signing a check, say, certainly involves more than my hand and body moving in a certain way. But what is this more? This is one of our basic problems in this book.

Generally speaking, actions can be primarily characterized either in terms of their *antecedents* or in terms of their *consequences*. In the first type of characterizing action, reference can be made to such things as the agent's intentions, acts of will, purposes, reasons, wants, desires, beliefs, emotions and feelings. For instance, in our view an action may be conceived as behavior caused or brought about by effective intentions (states or episodes of intending).

When actions are characterized in terms of their consequences the emphasis often is on the idea of an agent intervening with and changing the external world. Or, with a slightly different twist, an action can be considered as a kind of "response" to *task* specifications in such a way

that they essentially involve a public element—a result by which the correctness of the action performed can be judged.

In any case, to act (in this achievement or task sense) is to bring about public *results* or changes (or to prevent these change-events from occurring) in the world, which, so to speak, contains an opportunity for this type of intervention. Thus, an agent's opening the window presupposes that the window was not open before his acting and that it did not open without the agent's intervention. The change in the world brought about by the agent's action is, of course, the window becoming open. Furthermore, such results, which are conceptually involved in this concept of action, may themselves have various intended and unintended causal and other consequences (in this case, e.g., the ventilation of the room and the agent getting rid of his headache).[1]

In our view, to be developed later in detail, to give an adequate account of the notion of action one has to refer both to the antecedents and the consequences of a behavior process. The primary antecedents will in our approach be effective intentions (intendings) and beliefs, while wants, desires, duties and obligations, etc. are to be construed as secondary intention-forming antecedents.

We distinguish between action tokens, which are singular events, and action types, which correspondingly are generic events (universals). We shall later define and classify both action tokens and action types in detail.

A singular action token is a complex event brought about by an agent. It is complex in the sense of being process-like: an action token consists of a sequence of events (see especially Chapter 10). (We are going to technically call actions events, even if in many cases it would accord more with ordinary usage to call them processes.)

When discussing events seriously, one has to give account of how they are to be described or otherwise characterized and "picked out". This means that one must give a criterion of identity for events (and especially actions). Only then can one speak of "redescriptions" of actions. To get a glimpse of some of the problems involved let us quote a well-known passage from Davidson: "I flip the switch, turn on the light, and illuminate the room. Unbeknownst to me I also alert a prowler to the fact that I am home. Here I do not do four things, but only one, of which four descriptions have been given" (Davidson (1963), p. 686). A rather similar view of the identity of actions is held by Anscombe (1957) and von Wright (1971). Against these "radical *unifiers*" we have the "*fine grain*" approach held, for example, by Goldman (1970). According to this

approach we can say that there are four actions in question here. Intermediate positions are possible as well (see next section).

Let us draw an "action-tree" for Davidson's example in order to see what is involved. To make the example a little more interesting assume in addition that the agent, call him *A*, hurts his finger when flipping the switch and that the lighting of the bulb, due to a causal mechanism unknown to *A*, explodes a bomb. We get the simple diagram in Figure 2.1.

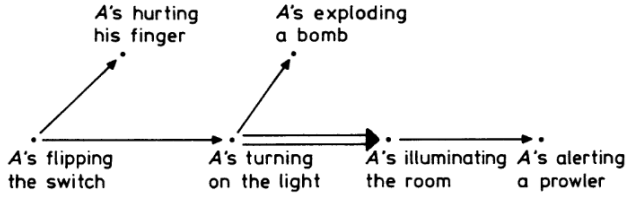

Figure 2.1

In this "action tree" the nodes represents different actions (for a fine grain theorist) or different descriptions of the same action (for a unifier). Single arrows represent a causal (or, more generally, factual) nomic relationship whereas the double arrows represent a "conceptual" relationship (relationship based on semantical or conceptual connections). In this example all the actions (or action descriptions) are (or have to do with) singular actions. Thus we can read from the diagram, for instance, that the singular action of *A*'s flipping the switch causally or *factually* generated *A*'s turning on the light. This is based on the fact that the flipping of the switch caused the light turning on. We may understand that in our present case the light's turning on (semantically) *means* the room becoming illuminated; thus *A*'s turning on the light *conceptually* generates *A*'s illuminating the room. Notice also that *A*'s flipping the switch in this tree generates two otherwise independent actions. The possibility of branching gives our diagram its tree-like character (or, more generally, a graph-structure).

For radical unifiers our action tree still represents only one action, whereas for a fine grain theorist there are now six different actions represented. Our fine grain theorist might perhaps concede that all the actions on the tree are the same or equivalent in the sense of being generated by the same primitive or basic action of flipping the switch, though these

actions are not strictly identical. One reason why they might not be considered strictly identical is simply that they hold different "generational" relationships to each other; another reason is that they have different results, as we shall argue later.

A difficulty for the fine grain theorist is this: the action tree is presumably supposed to represent what the agent did in a certain situation; but it is not restricted to representing previously intended actions, and may well include unintended or non-intentional actions; but there are, then, for a fine grain theorist an indefinite number of actions an agent performed in that given situation. One may invent all kinds of causally and conceptually generated actions to enlarge the tree: *A* flips the switch forcefully, *A* flips the switch forcefully by his left thumb, *A* flips the switch forcefully with his left thumb by standing firmly on his feet, *A* causes a bomb to explode through a mechanism connected to the switch.

It seems that in the debate between the radical unifiers and the fine grain theorists the distinction between actions (as extralinguistic events) and action descriptions has not been kept clear enough. The same holds for the notions of a ("bare") singular event and an exemplification of an action type. Let us now make a brief digression into this difficult topic, thereby shaping the groundfloor of our conceptual framework.

2. Actions and events

Our discussion has clearly suggested that one important task in characterizing actions is to give them philosophically illuminating criteria of identity: When do two actions count as the same? Or better (as we, strictly speaking, must answer the previous question by "never"): When is an identity statement '$a = b$', representing the identity of two actions a and b, true? For a number of reasons, it is important that such a criterion of identity be given to *singular* or tokened actions in the first place. One reason for this is that when we describe what an agent actually did, we of course have to start with action tokens. Another reason is that actions are so context-dependent that what is said about the relationship between action tokens very often does not hold true for the corresponding action types. This is true especially of the generational relationships between actions; and we think that any theory of action must give an account of them.

We claimed that actions are events; and events are here understood in a broad sense (including, e.g., short-term states). Before discussing identity

criteria for events, we must obviously say something about the nature of events. Our emphasis in this book will be on *singular* events as contrasted with *generic* ones. Singular events can be construed either as *unstructured* or as *structured* particulars. Furthermore, events may be regarded as either *extensional* or *intensional*, although we shall be less concerned with this distinction here. We shall in this book adopt a view according to which events can in a certain sense be treated as "bare" particulars while in another sense they are structured. In the sense in which they are bare entities, they are extensional, and in the sense in which they are structured, they are intensional.

Before discussing events in detail a few words have to be said about the reasons for considering events an ontological category on a par with, e.g., "substances". Let us note here that in this book the class of singular events (including, e.g., explosions, switch flippings, rememberings) will be taken as a rather broad category so as to include what might be called short-term states (i.e., perceivings and wantings). Thus, when we later in Chapter 3 speak about Jonesian states and episodes, they (except for process-episodes) will fall under the broad category of events as discussed here and elsewhere in this book.

What kind of reasons are there for taking singular events seriously? One important reason comes from explanation and causality. We often explain singular events and speak of causation between singular events (see Chapter 9). This also makes it natural to construe actions as events, as will be seen later in this book.

A closely related reason is that we want to speak about redescriptions of events (e.g., actions) and to distinguish between events and their different descriptions. (A satisfactory account of actions cannot be given without speaking about redescriptions of actions, and for this it is best to require that actions be construed as events.)

There are also more direct reasons, based on linguistic considerations, for the necessity of an autonomous category of singular concrete events. Davidson (1969) forcefully argues that the logical forms of some sentences require event-ontology as strongly as substances are required by some other statements. Thus, 'Jones buttered a piece of toast in the bathroom' seems to require event-ontology in its logical form in order to account in the *most economical* way for the fact that it entails 'Jones buttered a piece of toast' (see next section). A good example of a sentence which requires event-ontology is the so-called numerically restrictive sentence (see Wallace (1971) and next section).

Let us now go on to consider the event of somebody's raising his arm. We may regard it as a concrete event. But is it to be treated as a "bare" event or as a structured event? We can conceptually and epistemically think of this event as a "substratum" or a "bracketed" event. But from an ontological point of view, one has to be careful in speaking of substrata and unstructured events. Substrata without any qualities or properties are only *conceptual abstractions* or "bracketings" of real events. Any event that occurs in reality has factual properties. The point we are trying to make is indeed that all singular events do exemplify universals; but still a datable singular event is not to be regarded essentialistically as something which is somehow *constituted* by its instantiating a certain property (universal) rather than another one. Singular events are not reducible to universals.

What we perceive is an event conceptualized in a certain way. In our example we may perceive, for instance, a bodily movement of the arm going up or, alternatively, an intentional arm raising by the agent. These conceptualized events are structured, for they are events seen as exemplifying some specific universal (i.e., some generic events or properties).

It is unstructured events (i.e., events whose structure has not been displayed) that we are going to quantify over in our formal developments. Such events can be regarded as extensional entities and discussable within standard predicate logic. My conception of unstructured events seems to correspond fairly closely to Davidson's view of singular events (cf. Davidson (1969)), although I do not know exactly how far the agreement goes.

Structured events as we shall discuss them in this book are events as conceptualized within the "manifest image" (see Chapter 1). My view of conceptualized events (events with explicitly displayed structure) can be explained partly by reference to the rather well-known account by Kim (see, e.g., Kim (1973)). Kim thinks of a singular event as a concrete object (or n-tuple of objects) exemplifying a property (or n-adic relation) at a time. This exemplified property is to be intensionally construed. To take a simple dyadic example, the event of John's greeting at time t_0 is represented by [⟨John, t_0⟩, greeting], or, more formally, by [⟨a, t_0⟩, P]. The property P is the *constitutive* property of this event. That is, each singular event is assumed to be uniquely associated with a certain property which determines the *generic* event (or event type) that the singular event exemplifies. (The constitutive property P is, presumably, not really an *element* of the event, but it is a *fact* that the event exemplifies P.) The

generic event in question can be represented by $[\langle x, t\rangle, P]$ where x and t are variables ranging, respectively, over objects (including agents) and times. Each singular event can, however, "*merely exemplify*" a number of different properties. Thus, the event 'John's greeting' may exemplify the property 'uttering the sound *hello*' but also 'standing in front of the university'.

An identity condition for monadic events might now be formulated as follows (cf. Kim (1973), p. 223):

(K) Two singular events $[\langle a, t\rangle, P]$ and $[\langle b, t'\rangle, Q]$ are identical if and only if $a = b$, $t = t'$, and $P = Q$.

Identity conditions for n-adic events become slightly more complicated, although they, of course, are based on the above idea, too (cf. Kim (1973)).

Goldman (1970) has applied the above Kimian kind of analysis of events to action contexts. In the above definition of identity, we may take a and b to be agents and P and Q to be action properties. Thus, two singular actions are said to be identical if and only if they involve the same agents, the same action properties, and the same times (cf. Goldman (1970), p. 10). What results is the (or a) fine grain approach to identifying actions we discussed earlier. There are, however, difficulties with Kim's and Goldman's views, and some of them must be taken up here.

The constitutive properties in Kim's account are assumed to represent generic events. Therefore singular events are really analyzed as being merely exemplifications of generic events. But now the problem arises what the set of constitutive properties (or generic events) really is, and that question Kim does not answer. Nor has anybody else answered it (fully) satisfactorily. Still, it is clear that this approach is a fine grain approach, and too much so, I think. Thus, for Goldman and Kim individual stabbings can never be killings, nor can singular arm wavings be greetings.

A related problem is what the identity conditions for properties exactly are. Linguistic criteria have often been tried. Thus, for example, Goldman (1970) suggests that two properties P and Q are identical if and only if they are designated by synonymous predicates. But that at least presupposes an analysis of synonymity which is independent of property identity. That seems hard to give. What is more, the suggested criterion seems clearly false. Let me indicate why.

Goldman's criterion is this: "if there are phrases in a language which express [properties] ϕ and ϕ', then their being synonymous is a necessary

and sufficient condition for their expressing the same property" (Goldman (1970), p. 12). But as even Goldman himself notices one can use non-synonymous descriptions to express one and the same property. Thus, 'blue' and 'the color of the sky' gives one such pair and 'temperature' and 'mean kinetic energy' serves as another. Thus, synonymity is not necessary for property identity.

Synonymity is not even sufficient for property identity. I do not have a very good scientific example to show this, but in any case my argument goes as follows. Assume as true that theoretical terms in science are often "observationally" open (cf. Tuomela (1973a) for arguments). To use our later terminology, if $T(\lambda \cup \mu)$ is a theory with λ as its set of observational predicates and μ its set of theoretical predicates (see Chapter 4 and Tuomela (1973a) for this dichotomy), then the model-theoretic interpretations of the μ-predicates are not, normally at least, completely fixed by the interpretations of the λ-predicates. Assume, furthermore, as acceptable the *semantic* principle of *Conceptual Functionalism* (see (CF) of Chapter 4) according to which, roughly, the λ-predicates determine the meanings of the μ-predicates. This principle seems good and tenable at least for cognitive and motivation psychology. Then it can be argued that, in general, the μ-predicates *multiply designate*, i.e., one predicate will simultaneously designate several properties (which model-theoretically are represented by different and nonisomorphic extensions). Thus, going back to Goldman's criterion, 'ego strength' could stand for the two synonymous property-designating expressions (which are identical!) in a personality theory. Still 'ego strength' designates a number of different properties. This result is in good agreement with what has been found out in personality psychology by means of experimental work, although no sufficiently exact personality theories exist to really prove this point in full detail.

As we have noticed, Kim's approach is faced with the difficult and so far unsolved problem of property identity. Before more can be said about this problem his account cannot be seriously evaluated. A related feature about Kim's approach that we noticed is that it is essentially a fine grain approach: each singular structured event is associated with exactly one constitutive property. How fine grained the approach really is of course depends on how the "final" set of generic events will look.

A. Rosenberg (1974) argues that in face of problems like the above Kim's approach has to be modified by allowing for more than one constitutive property for each singular event (call this criterion (K′)). This

modification perhaps makes the approach a little less fine grained, but it is still faced with the difficulties of selecting the total set of constitutive properties and the set of constitutive properties for each particular event and of giving identity criteria for properties. Furthermore, Rosenberg claims that Kim's modified approach now is equivalent to Davidson's approach (as far as event identities are concerned), given a suitable backing law account of singular causation. If so, it could be said that Kim's modified criterion of event identity gives the *meaning* of identity statements while Davidson's causalistic approach (see below) gives (the) *epistemic* grounds for them. But Rosenberg's claim seems incorrect. It is instructive to see why.

Davidson's (1969) well known criterion is as follows:

(D) Two singular events e_1 and e_2 are identical, viz. $e_1 = e_2$, if and only if

(e) (e caused $e_1 \rightarrow e$ caused e_2) and
(e) (e_1 caused $e \rightarrow e_2$ caused e)

Let us postpone a discussion of the adequacy of (D) for a while in order to discuss immediately Rosenberg's thesis.

The type of backing law account of causation Rosenberg (and presumably Kim) have in mind is simply this:

(CC) e_1 caused e_2 if and only if the constitutive properties of e_1 and e_2 are "constantly conjoined" or lawfully correlated.

Presumably (CC) is to be understood so that for e_1's causing e_2 it suffices that *one* of e_1's constitutive properties be lawfully correlated with *one* of e_2's constitutive properties. We shall, later in this book, indicate that, as it stands, (CC) is inadequate in several respects, and we are going to replace it by a much stronger condition (see Chapter 9). But, for the sake of argument, let us here consider only (CC).

It is obvious that if two events e_1 and e_2 are identical in the sense of (K′) then, as their constitutive properties coincide, these events have got to satisfy the right hand side of (D). But how about the converse? It seems that it does not have to be true. Assume that e_1 is associated with some finite set $\{P_1, \ldots, P_m\}$ of constitutive properties and e_2 with the set $\{Q_1, \ldots, Q_n\}$ (if the sets really turn out to be identical then of course $m = n$). If we now go through all the events and find that the right hand side of (D) comes out true, what do we have? The idea here would be to

try and guarantee that each P_i, $i = 1, \ldots, m$, could be correlated with a Q_j, $j = 1, \ldots, n$, so that P_i and Q_j are found to play exactly the same causal roles with respect to every other property. But this is not entailed by the fulfilment of the right hand side of (D) without further qualification. The qualification needed is just that we really have to go through *all* the properties (whatever their set really is). If we had worked with Kim's original view, which associates a unique constitutive property with each event, that problem obviously would not have arisen.

Perhaps we may accept the above qualification, after all. There is, however, still another difficulty here. For, as said, in order for (D) to entail (K′) what we essentially need is the truth of this criterion:

(D*) Two properties P and Q are identical, viz. $P = Q$, if and only if P and Q have identical causal relations with respect to every other property.

However, even if the entailment goes from right to left, the converse implication can hardly be accepted as true (irrespective of whether only some kind of scientifically "good" and "nonartificial" properties are admitted in the class of all properties here).

First, I think non-causal nomic connections should be accepted as affecting the identity of properties, and criterion (D*) ignores them. (An example of a non-causal law would be, "The formation of the respiratory system in a human embryo always succeeds the formation of the circulatory system".)

Secondly, it seems that we cannot rule out the possibility that even if P and Q play identical causal (or nomic) roles they may at the same time have different, either "accidental" or conceptual connections to some properties. As an example of a conceptual connection, which should matter when discussing property identity, consider "Extending one's arm when cycling means signalling for a turn". (When discussing examples like this we admittedly run into problems concerning the ontological status of universals like properties and types, but we cannot go into those problems here.)

We have thus found that Kim's modified criterion (K′) and Davidson's criterion (D) are different. Furthermore, in our approach neither of these criteria will be considered acceptable as such. We have already discussed the criteria (K) and (K′). Now a few words about (D) are in order.

The most basic fault with (D) is that it is *circular* as a meaning analysis of event identity. This is simply due to the fact that the right hand side of

(D) makes use of quantifiers which are intended to range over events (there is no way of interpreting (D) without admitting events into the range of the quantifiers). But this clearly requires that these events have already been identified (in some sufficiently clear sense). Therefore I would not take (D) as a meaning criterion but rather as an *epistemic* criterion.

There is also a technical fault in (D). It is due to the use of the material implication $\rightarrow$ in it. Suppose that e_1' is an event which has no causes and no effects. Then if we substitute e_1' for e_1 in (D) we obtain the result that e_1' is identical with any arbitrary event. This fault can be removed by using another explicate for the if-then relation in (D). For instance Lewis' variably strict conditional $\Box\!\!\rightarrow$ might be considered as a good candidate here (cf. Lewis (1973a)).

We shall, in addition, use our causal predicate $C(-, -)$ (to be defined later in Chapter 9) to explicate singular causation. Let us call the resulting criterion (D'). (D') can be used as an, or perhaps "the", epistemological criterion for event identity in "normal" contexts, i.e., where events like e_1' above are not concerned. For even (D') does not give us anything constructive and useful with respect to such "random" events, although it technically removes the mentioned paradox connected with (D).

In the case of unstructured singular events (i.e. events whose structure has not been displayed) there is yet another epistemological criterion of identity which must be taken up here. If singular events are viewed as such unstructured events, then intuitively they are spatiotemporal "worms", or "chunks", or "slices", or even sequences of such. This suggests that they can be given a *spatiotemporal* criterion of identity.

It is clear that neither same place nor same time alone gives more than a necessary condition for identity. There can clearly be events that happen at the same place at different times, and there can be simultaneous events occupying different places. But jointly these factors give an identity criterion for two singular events e_1 and e_2:

(PT) $e_1 = e_2$ if and only if e_1 and e_2 occur at the same time and at the same place.

This criterion may seem either too obviously true or too obviously false. It seems obviously false for conceptualized events (cf., e.g., arm-movement and arm-raising or the ball warming example below).

In the case of nonconceptualized events it again may seem so obviously true as to be unilluminating. Still, for instance, Davidson seems sceptical about (PT). Let us consider why. First, we have to make an agreement

concerning location. We take as the location of an event the location of the smallest material object a change in which is identical with the event in question. (Otherwise it can be argued on the basis of the Frege–Quine–Davidson theorem that every simultaneous event takes place everywhere; cf. Davidson (1969).) Given this convention for location (accepted by Davidson) we consider Davidson's example: "If a metal ball becomes warmer during a certain minute, and during the same minute rotates through 35 degrees, must we say these are the same event?" (Davidson (1969), p. 230). Davidson himself suggests the possibility that the warming of the ball during this minute is identical with the sum of the motions of the particles that constitute the ball during this time; and let us construe rotation similarly. This (or something like this) must indeed be said here. But it should be emphasized that when making statements (like the above) about this situation we of course cannot help introducing and using conceptualizations of the situation. This is what makes us think that two events took place. But, I claim that we are here speaking of the same event which, however, was conceptualized in two different ways.

One trouble with (PT) is that the location of mental events is problematic. In the case of wantings, believings, etc. the person presumably is the substance which determines the location (cf. our convention). But how about, e. g., sensations and impressions? What about aches in phantom limbs? It seems to me that the *adverbial* account of these cases works best, and hence the location again will be the whole person, irrespective of how unilluminating that may sound (cf. Chapter 4).

To sum up, we have not found damaging counterexamples against (PT). It seems acceptable as an epistemological criterion for (unstructured) events. I do not think, however, that (PT) qualifies as a meaning analysis of identity statements, for that in effect would require going through all the descriptive information there is in principle to be had about the entities in question.

How is the meaning of identity statements then to be analyzed, if neither (K), (K′), (D′) nor (PT) qualify in this task? What comes closest to a meaning criterion of identity is simply Leibniz's criterion, of course:

(L) $e_1 = e_2$ if and only if $(F)[e_1$ has F if and only if e_2 has $F]$.

In (L) F ranges over properties. It should be emphasized that the events e_1 and e_2 here should be taken as nonstructured events (in my sense). Otherwise, difficulties connected to the independent semantical ascribability of properties to e_1 and e_2 will arise.

As we have indicated there is something important in Kim's type of account. However, as we have seen, it connects singular events to generic events somewhat "too closely" (no matter whether (K) or (K′) is accepted). We prefer to approach the matter as follows.

Singular events' having structure can be regarded as a matter of one's pragmatic interests. Consider thus, for instance, a monadic singular event $e = [\langle a, t\rangle, P]$. Here a represents the "locus" and t the time of occurrence of e. P can be termed the aspect property of the event e. It should be emphasized however that P should be taken to represent the aspect from which we, so to speak, view the particular in question. It reflects and represents our pragmatic interests rather than being an a priori *essential* property of the event in something like the sense of traditional essentialism.

There is no prefixed set of aspect properties for a given singular event. In principle it can be viewed from any aspect. Thus the structured event $e = [\langle a, t\rangle, P]$ can be taken to be just an "unstructured" spatiotemporal event viewed from the aspect P. In other words 'e' and 'e as a P' represent one and the same event. (Any event can have many names and descriptions, of course.)

The aspect properties are universals. We often speak about types as well. Thus, we shall frequently speak of singular actions u as X'ings, where X is an action type. In this book we shall not make any strong assumptions about the precise metaphysical status and nature of properties and types (and other universals).

In our language (which may be a first order extensional language or a more complex one) we may pick out a singular event e by singular terms. Perhaps the event has a proper name (individual constant) 'e' by which the event can be registered. Below we shall in fact for convenience assume that for every singular event e we pick out for consideration there is available a singular term 'e' suitable for this referring task. Because of this correspondence it does not matter much whether we speak in the material mode or the linguistic mode about the event. We shall prefer the former way of speaking unless confusions are to be expected. Following common practice we will then not pay much attention to this difference, which of course is important in principle.

Similarly, as with singular events, let us assume that the predicate 'P' *designates* the aspect property P. While 'e' suffices to pick out e, it does not in itself contain any *descriptive* information about e. However, a statement of the kind '$P(e)$' (an "aspect statement" for e) does give descriptive

information about e, viz. it says that e has P. But then it also describes e as a P, i.e. as a specified structured event.

In general, then, we have the connection between a singular unstructured event, say e, and a corresponding structured event that the unstructured e *is*, as a structured event, e *as a* P (if P is the aspect property of e); or, alternatively, e is e *under a certain description* D. Here D could be just $P(e)$.

In our language we thus have, among other things, the following semantic devices for "speaking about" events. First, we have singular terms (proper names and definite descriptions) referring to singular events. We also have singular variables running over singular events. Next, we have aspect statements such as '$P(e)$'. Such a statement represents a *fact* (e.g., the occurrence of an event) just in case '$(Ex)(x = e)$' is true, viz. if and only if 'e' indeed refers to an existing entity e. (That the existential statement captures the occurrence of e may not be beyond dispute, but here we cannot pause to discuss it.) In addition, there are of course an endless amount of other descriptive statements, both true and false, about e.

Redescribing a structured singular event does not present any special semantical problem for us. It just amounts to using different new aspect statements and other statements containing new predicates such that these statements are true of the event.

Can we say anything about *generic* events within our framework? I think we can, and even in several different ways. First, *open sentences* like '$P(x)$', where x is an individual variable, represent *generic events* or *event types*, e.g., "arm-raising". When x is replaced by a name of a singular event we have a statement describing a singular event. (I have here, as well as in the case of singular events, omitted the explicit linguistic representation of locus and time. In any case, they may be represented by variables in a rather obvious way, if needed.) Secondly, *existential statements* of the form '$(Ex)P(x)$' represent generic events in another sense, viz. they say that a singular event exemplifying the generic event $P(x)$ exists. Thus, the existential statement '(Ex) Buttered (Jones, toast, x)' represents the non-pure generic event of Jones' buttering the toast (see Section 3). Similarly, the existential statement '(Ex) Buttered (A, p, x)' represents the pure generic event of buttering, provided A and p are variables. Furthermore, we may also quantify over predicates (if our language is a second-order one) and in that way discuss, e.g., the existence of generic events and relationships between them.

To summarize our view on event identities, the meaning of identity statements is given by Leibniz's criterion (L); (D′) and (PT) can be taken

3. Actions and action statements

In order to get a better grip on action statements (and on action-explanations), we shall now introduce a formal (or semiformal) framework. Our main purpose here is not to give a theory of the logical form of psychological statements or anything like that but rather to introduce a framework for the logical characterization of psychological statements within which philosophical problems can better be discussed.

We start by considering the following action statement about Jones' buttering a piece of toast which has often been discussed in the literature (cf. Davidson (1966)):

(*1*) Jones buttered the toast slowly, deliberately, in the bathroom, with a knife, at midnight.

One may say that the task of formalizing action statements is to exhibit their logical structure. Furthermore, one may require that this be done so that it is shown how the meanings of action sentences depend on their structure and indeed so that Tarski's semantical theory of truth be respected (cf. Davidson (1966) and (1975)).

If one accepts a strict "bringing about"-analysis of actions such as Chisholm (1966) or von Wright (1968) and (1971) one might think of the formalization of action statements as follows. An action consists of an agent's bringing about an event (or state). To analyze this we, for simplicity, first consider the statement

(*2*) Jones buttered the toast.

Here, the agent Jones is then said to have brought about the state of affairs that the toast is buttered. Speaking in logico-linguistic terms, on analogy with the case when we represent intentional attitudes (e.g. belief) by sentential operators, we may now think of the agent functioning as a sentential operator, which operates on the sentence describing the result of the action. More exactly, following Pörn (1974), we may represent (*2*) by a formal statement '$E_A p$', which reads 'A brought it about that p'. The bringing about-operator E_A can still be analyzed further:

$$E_A p = D_A p \ \& \ C'_A p$$

where '$D_A p$' reads 'it is necessary for something that A does that p' and '$C'_A p$' reads 'but for A's action it might not be that p'. '$D_A p$' thus stands for the necessity aspect of action according to which an action is necessary

for its result (see von Wright (1968)). But $C'_A p$ does not guarantee the satisfaction of p, which it should. (This may be called the sufficiency requirement.) But we are not interested here in patching such flaws but rather in the general feasibility of this type of approach. For there are many difficulties attached to this approach to formalizing action statements. I shall now briefly take up some of the most troublesome of them (cf. Davidson (1966)). The problems to be mentioned below seem to me to apply to the entire approach and not only to some particular version of it.

In order to get the agency aspect correct, (*2*) must be rendered not, e.g., as 'Jones brought it about that the toast was buttered' but rather as

(*3*) Jones brought it about that Jones buttered the toast.

But now we can see that no philosophical illumination and hardly any logical gain is to be obtained from such an analysis. First, what does (*3*) tell us about agency (and about bringing about) that the contained sentence 'Jones buttered the toast' does not? Does the difference between these two statements represent something? If it does, it has to be clarified so that the "standard" objections to agent causation do not apply (cf. Sellars (1969a)). If it does not, it is hard to see how (*3*) clarifies the meaning of 'Jones buttered the toast'. Naturally, a logical analysis of action sentences should somehow clarify their meaning.

Secondly, the above kind of "bringing about"-analysis at least does not show how the meanings of action statements depend on the meanings of their contained parts. This is seen for instance from the fact that this analysis gives no account of the so called *variable polyadicity* problem (cf. Kenny (1965)).

Consider again statement (*1*) and compare it with (*2*). The approach we are criticizing gives no solution to the problem of how to handle polyadicity of action verbs. (*2*) is represented in the bringing about-theory by means of the two-place predicate 'Buttered (Jones, the toast)'. But in order to account for the fact that Jones buttered his toast in the bathroom a new argument apparently has to be introduced into the predicate, and so on. So it would seem that there is no end to adding new arguments and hence to increasing the polyadicity of the action predicate. Thus, the logical form of action statements seems to be strongly affected by rather arbitrary additions. One problem that follows from this is that it does not seem possible to account for such obvious logical inferences as that (*1*) entails (*2*).

There are also difficulties connected to problems of action identity and the modification of action verbs, which affect the above kind of approach, although we shall not discuss them now.

In view of our criticisms it seems that we are well motivated to look for a completely different type of approach. We thus turn to an interesting approach advocated by Davidson (see, e.g., Davidson (1966)).

Let us consider statement (*2*). Does it describe a singular action? It does not in the sense of picking out one particular action. What Jones did in the bathroom that night (and what is more fully described by (*1*)) is something that makes (*2*) true, to be sure. But still (*2*) does not describe only that particular action, for it is satisfied by all singular butterings of the toast by Jones. Thus (*2*) seems to describe something like a generic (or partially generic) event, but not a singular event. It should be noticed that if, as a matter of fact, Jones buttered the toast only once in his life and if hence (*2*) is, as a matter of fact, satisfied by only one singular event, this does not change the semantic (or logical) situation. As to its logical form, (*2*) is *existential*, and satisfiable *in principle* by several singular events.

Now Davidson, having made this observation, makes an interesting suggestion for exhibiting the logical form, or "deep structure", of (*2*). Its philosophically interesting content is that the ontology of statements should be made explicit in the formalization. In this case the logical form of (*2*) becomes:

(*4*) (*Ex*) Buttered (Jones, the toast, *x*).

Thus the idea is to consider the action verb 'butter' as a three place predicate rather than as a dyadic one.[2] The symbol '*x*' is a variable which runs over singular events (actions). This variable can be replaced by singular terms naming singular events. In fact, the standard semantical device we have for referring to singular events is just singular terms. Whether or not a statement not containing such singular terms succeeds in picking out a singular event will always be a contingent matter.

Singular actions are thus to be named by singular terms such as proper names or definite descriptions (see Chomsky (1970) for argument that nominalizations occur in linguistic deep structures). For instance, nominalizations such as 'Jones' buttering the toast' can be construed as singular terms by the stipulation that 'Jones' buttering the toast' names just that singular action we have been speaking about.

It should be emphasized that Davidson's idea of exhibiting singular ontological entities in the formalization does not apply only to events. It can be applied to any kind of entities that one's language treats as particulars. Thus, it applies to singular states (also those not included here in the class of events), situations and processes. Thus, whenever singular entities are somehow considered philosophically and methodologically important Davidson's approach can be applied. As we remarked in the previous section, this technique has proved to be interesting and even necessary, for instance, in the treatment of causation and determination (cf., e.g., Davidson (1967), Berofsky (1971), Tuomela (1974b)). Furthermore, it seems to show new interesting applications for the notion of quantificational *depth* when the quantifiers range over, possibly complex, singular events. (Cf. the philosophical applications in Hintikka (1973) and the methodological importance given to the notion of depth in Tuomela (1973a) and Chapter 4 below.)

Let us now return to see how Davidson's formalization works in the case of the more complex action statement (*1*). Davidson omits from his treatment such locutions as 'slowly' and 'deliberately'. The phrase 'slowly' is omitted in his analysis mainly as it does not introduce a new entity in the formalization and as it is not only a problem for action sentences but for many other kinds of discourse as well (cf. 'Jones is a good actor'). Phrases like 'deliberately' (or 'intentionally', etc.) are omitted by Davidson because they are tied to the agent rather than to the action.[3]

Let us also forget about the phrase 'with a knife' in (*1*) for a moment. With all these omissions we are left with

(*5*) Jones buttered the toast in the bathroom, at midnight.

I take it that Davidson's formalization (formal translation) of (*5*) would be:

(*6*) (Ex) Buttered (Jones, the toast, x) & In (the bathroom, x) & At (midnight, x).

Verbally (*6*) reads: There is a singular event such that this event is a buttering of the toast by Jones and such that this event occurs in the bathroom and at midnight.

The Davidsonian way of formalizing action statements may more generally be represented by

(*7*) (Ex) Verbed (Agent, Object, x) & T(Object, x).

This represents a simplified case, however, for in general the predicate 'Verbed' may involve several agents and objects, but that does not affect the logical and philosophical situation, it seems. The predicate T is assumed to take care of the place (in the bathroom) and time (at midnight) of the event as well as all the other relevant attributes of the event (e.g., the instrument). I have simplified the account here in that there should in fact be several conjuncts with predicates like T and with different objects (cf. statement (*6*)). But our simplification does not affect the central theoretical issues involved.

It is easy to see that, given the formalizations (*6*) of (*1*) and (*4*) of (*2*), statement (*1*) entails statement (*2*). Notice also that the formalizations have been performed by means of standard first-order logic only. Davidson's approach is in fact a continuation of the programme for formalizing natural language advocated (in a sense) by Frege, Tarski, and Reichenbach.

Let us here still mention two more illustrations of the Davidsonian type of formalization. We first consider

(*8*) Shem kicked Shaun before Shaun kicked Shem.

We can understand the problematic temporal 'before' as a predicate between singular events and get this as the logical form:

(*9*) (Ex)(Ey) (Kicked (Shem, Shaun, x) &
Kicked (Shaun, Shem, y) & Before (x, y)).

For another illustration of the usefulness and flexibility of this approach, consider the following "numerical" sentence:

(*10*) Shem kicked Shaun at least twice.

The logical form of (*10*) can be taken as given by:

(*11*) (Ex)(Ey) (Kicked (Shem, Shaun, x) &
Kicked (Shem, Shaun, y) & $x \neq y$).

It is again easy to see how (11) explicitly entails that Shem kicked Shaun at least once. This fact and such related facts as that (*1*) trivially entails (*2*) within Davidson's approach can be taken to show something important. Namely, this indicates that the approach indeed is able to recursively account for the meanings of sentences on the basis of their parts. In other words, in the manner of Tarski's well known Convention T ('T' for 'true' here) inferential relationships between sentences are accounted for

It should be possible to construe a general theory of predicate modification in the above manner by means of just a finite number of "core" predicates and only a finite number of recursive modifying operators. Then and only then can we have a recursively learnable language (cf. Davidson (1965)). Such a theory should of course account for inferences in the earlier discussed manner. Thus, a statement containing a context . . . slowly-buttered . . . should entail another, otherwise the same, statement, except that the latter contains the replaced context . . . buttered. . . .

For our present purposes we shall still make another assumption or rather a restriction. We assume that a theory-to-be of predicate modification has generated a finite amount of structured predicates which suffice for the purposes of psychological theory formation. Thus below, when we use the placeholder 'Mod-Verbed ( )' for action predicates, we assume that the verb in question is either a core predicate or a modified predicate obtained by means of admissible transformations from core predicates. Thus, we can leave most of the problems involved in verb-modification to linguists and logicians.

Let us now return to the Sellars–Clark formalization (*12*). The first important thing to notice about it is that it is incorrect. Even if some adverbs in the original statement (*1*) might be taken to modify the verb 'Buttered', not all of them do. I think that place and time do not modify the verb but are properties of the singular event itself. Thus 'Slowly' and 'With-the-knife' modify 'Buttered' and do not introduce a new entity in the formalization, whereas 'In-the-bathroom' (place) and 'At-midnight' (time) in a sense do. So we would get (leaving the nature of the verb-modification implicit):

(*14*) (Ex) Mod-Buttered (Jones, the toast, x) &
At (midnight, x) & In (bathroom, x).

But have we really proved that predicate modification is necessary? I think not. That is very hard to do. The matter is more a question of convenience until we arrive at a broad linguistic theory of logical form and meaning telling us what is right.

If someone has strong nominalistic and anti-essentialist intuitions it is very hard to convince him of the necessity of predicate modification. It should be kept in mind that the event variable x can in principle be taken to refer to the whole behavior situation, i.e., a complex event including more than a bodily movement.

One of the best pair of sentences for showing the need for (at least)

predicate modification would be of the following type:

(*15*) Clumsily, John kicked the ball.

(*16*) John clumsily kicked the ball.

One might claim that on a certain occasion (*15*) could be truly uttered, whereas (*16*) would be false (the kick itself was perfect). But, if this is granted, perhaps we must include both 'clumsily$_1$' and 'clumsily$_2$' in our deep structures. However, that seems to make deep structures (logical forms) strongly dependent on linguistic (and perhaps extralinguistic) context.[5]

Another example apparently showing the need for predicate modification is given by:

(*17*) John moved quickly but raced slowly.

John might be a runner left last in the final heat of 100 m dash at the Olympic games. A mechanical application of Davidsonian formalization might seem to yield (after rather obvious steps):

(*18*) (*Ex*) (Moves (John, *x*) & Quick (*x*) & Races (John, *x*) & Not-Quick (*x*)).

Of course, (*18*) is a logically inconsistent sentence. However, we can easily see that 'quickly' and 'slowly' in (*17*) implicitly contain *different* standards of reference. Thus, 'quickly' should presumably be taken relative to (something like) the class of human beings and 'slowly' relative to the class of competitors in the final heat. Making these relativizations explicit obviously blocks the inference to (*18*). Still we may ask, as above, should we make deep structures dependent on background information such as comparison standards for the adverbs?

The list of examples of the above kind can be continued and similar remarks and counterremarks can be made. In this book we shall not discuss this problem further, but allow for the possibility of predicate modification as sketched earlier. Notice, that our way of treating the situation in no way destroys the general and most basic features of the Frege–Davidson programme; it essentially involves only a new selection of basic predicates.

For the time being it is too early to evaluate the possibilities and (final) adequacy of the Frege–Davidson programme. Theoretically viewed, linguistic theory is still, in spite of the wealth of data produced, almost

in its infancy. (For a good, recent discussion of the status of the Frege–Davidson programme see Harman (1972)).[6]

We shall end this chapter by noticing that the modern approaches to intensional logic and the Davidsonian approach are not formally very far from each other (even if they may differ interpretationally as to their ontologies). For one thing, Cresswell (1974) has shown that Davidson's approach can be embedded in Montague's approach under one condition. The condition is that two distinct events never occur in exactly the same set of possible worlds. However, for this condition to be maintained for singular events such as our unstructured events, the set of possible worlds would become extremely large, since it would have to correspond to all the possible ways of conceptualizing substratum events. The situation thus depends very much on what one accepts as one's set of possible worlds.

One can also approach the same matter from a slightly different angle. (Here I am indebted to Prof. Zoltan Domotor.) In Kripke–Montague-type logic one usually interprets an n-place predicate P by an interpretation function (I) as follows, using ordinary set-theoretic notation:

(a) $\quad \mathrm{I}(P) : D^n \rightarrow \mathbf{2}^S$

Here D is a set representing the domain of interpretation, $\mathbf{2} = \{0, 1\}$ represents the truth values, and S represents one's set of possible worlds, situations, events, or indices (or whatever one calls them). Finally, $\mathbf{2}^S$ stands for the set of functions from S to $\mathbf{2}$. Now, in Artificial Intelligence S is usually taken to be the set of situations and the interpretation of the predicate P is given by:

(b) $\quad \mathrm{I}(P) : S \rightarrow (D^n \rightarrow \mathbf{2})$.

Within Davidson's approach the interpretation clearly becomes this:

(c) $\quad \mathrm{I}(P) : S \times D^n \rightarrow \mathbf{2}$.

But it is easily seen that the interpretations (a), (b), and (c) are set-theoretically equivalent as long as just ordinary functions are allowed.[7] Furthermore, they remain equivalent under homomorphic mappings if instead of S we use a structure $\langle S, R \rangle$, where R is an "alternativeness" relation (partial ordering) familiar from modal logic.

Thus, purely formally, Davidson's approach is embeddable into the "possible worlds" approach. For a more general argument that corresponding to any intensional language there exists an extensional (perhaps somewhat artificial looking) language see Parsons (1970) and especially

Lewis (1974). This extensional language is "Ramsey-equivalent" (i.e., after quantifying existentially over metalinguistic semantic talk) with the intensional language. A price to be paid for extensionality is a somewhat complicated ontology. Formally speaking, one may, however, always "go extensional" it seems; the extensional language in question being in the general case a suitable predicate logic, which *most comfortably* (although perhaps not logically necessarily) is allowed to quantify over at least two types of entities and hence to be at least second-order (cf. Tuomela (1973a) on the scope of first-order formalization).

NOTES

[1] One of the best analyses of the "bringing-about" aspect of action is in von Wright (1968). In this work the following four elementary action types are distinguished: 1) producing a change (or result), 2) preventing a change, 3) forbearing to produce a change, 4) forbearing to prevent a change. However, as we shall accept and operate with a somewhat different notion of action than that of von Wright, we shall not here go into the details of his theory.

Let us remark here that we shall in this book concentrate on "positive" actions and thus leave out such "negative" actions as forbearances and negligences. In principle, however, our treatment covers them as well. Basically, such negative actions are to be analyzed in terms of suitable negated action descriptions. In our ontology, however, we do not have any *properly* negative actions any more than other negative events.

[2] One can handle the tenses of verbs either as in (4) or one can, alternatively, restrict oneself to present tenses and add in contexts like (*4*) that x is past.

[3] Such adverbs as 'intentional' can be formalized by using the locution 'It is intentional of A that p', where 'A' names the agent and p says what the agent did (cf. Davidson (1966)). It is essential that in p we use a name of A under which he recognizes himself (e.g. we may use 'A' or a pronoun). Thus we may in our logic use a *sentential* intentionality operator '$I^*_A(p)$' which reads 'It is intentional of A that p' to account for the intentionality of the action. Thus, if we in our example, formalized by (6) below, want to add that Jones performed his action intentionally, we just add the conjunct '$I^*_{\text{Jones}}(p)$'. Of course this logical method is not meant to be a *philosophical* solution of intentionality at all (cf. Chapters 3, 7, and 10).

Another, and perhaps more satisfactory way of formalizing that-clauses in contexts like 'It is intentional of A that p', 'A intends that p', 'A believes that p' is provided by Harman's operator '$\doteq$' (see Harman (1972)). This is an operator that converts any logical part of a logical structure (or deep structure or "proposition") into the corresponding part of the *structural* name of the proposition in question. In other words, $\doteq$ will grind out the syntactical deep structure when operated with on a surface sentence. E.g. '$\doteq$' applied to 'p & q' gives the logical form '$\doteq p \doteq \& \doteq q$'. As this operator clearly changes the logical properties of whatever it applies to it can be used to clarify issues such as the opacity of belief-contexts. We shall later occasionally employ "dot-quotes", originally introduced by Sellars, basically in the same role as Harman uses his $\doteq$.

[4] We shall understand Convention T in a strong sense here. Thus we assume that it, in particular, entails all the appropriate instances of 'x is true if and only if p' where 'x' is replaced by a name *descriptive of the logical form* of a sentence s of the object language

and 'p' is replaced by s itself. (This version of Convention T escapes the counterexample proposed in Hintikka (1975).)

[5] Parsons (1970) gives a somewhat similar example for the necessity of predicate modification. Consider

(a) John painstakingly wrote illegibly.

(b) John wrote painstakingly and illegibly.

Here it may seem clear that 'painstakingly' modifies the illegible writing in (a), whereas in (b) it only modifies the writing. But somebody with strong nominalistic intuitions may object.

[6] Harman (1972) discusses various approaches to formalizing natural language statements and finding their logical forms. He defends a somewhat strict, nominalistic, version of what we have called the Frege–Davidson programme against, e.g., various operator-approaches.

Harman proposes that the following five criteria of adequacy should be imposed on any theory of logical form: 1) A theory of logical form must assign forms to sentences in a way that permits a finite theory of truth for the language (in our above sense). A theory of logical form should minimize 2) the number of new rules of logic (i.e., standard quantification theory) and 3) the number of new logical and nonlogical axioms. 4) A theory of logical form should avoid unnecessary ontological commitments and 5) it should be compatible with syntax.

Even if we subscribe to these principles, we have above argued for the need of predicate modification, contrary to Harman. At the present state of art there is relatively much room for disagreement within the scope of the above five conditions of adequacy.

[7] To give an example, consider again Jones' buttering the toast. To keep consistent with the above treatment we now, instead of (*4*), consider the translation '(Ex) Butters $(x$, Jones, the toast)', and we consider the interpretation given by functions of kind

$$(*) \qquad \mathrm{I}\,(\text{Butters}): S \times D_1 \times D_2 \to \mathbf{2},$$

where S is the event domain, D_1 the domain of agents (including, e.g., Jones) and D_2 the domain of objects (including, e.g., the piece of toast). Our (*) is an instance of (c), which is in one-one correspondence with (b). (b) again is merely a notational variant of (a). This reduces the present example of Davidsonian formalization to Kripke–Montague semantics.


image
not
available


image
not
available

Index of Names

Index of Subjects