Executions, Motivations and Accomplishments

David Israel, John Perry and Syun Tutiya

[From *The Philosophical Review* (October, 1993): 515–40.]

1 Introduction

Brutus wanted to kill Caesar. He believed that Caesar was an ordinary mortal, and that, given this, stabbing him (by which we mean plunging a knife into his heart) was a way of killing him. He thought that he could stab Caesar, for he remembered that he had a knife and saw that Caesar was standing next to him on his left, in the Forum. So Brutus was motivated to stab the man to his left. He did so, thereby killing Caesar.

We have explained Brutus's act by citing a complex of beliefs, desires and perceptions that motivated it. Our explanation provides a causal account of Brutus's act. The beliefs, desires and perceptions in such a *motivating complex* are particular cognitions. The act was also a particular, an event that occurred at a certain place and time. The cognitions caused the act.[1]

Our explanation also provides a rationale for Brutus's act. The beliefs, desires and perceptions of Brutus's that we cite had *contents*. The desire we cited had the content that Brutus kill Caesar. The first belief we cited had the content that Caesar was an ordinary mortal. The act was of a certain type. The explanation provides a rationale because the contents of the cognitions mesh in a certain way with one another and with the type of the act. It was the type of act that would satisfy Brutus's desire to kill

Caesar, if the beliefs we cited were true. If the person next to him is Caesar, and Caesar is mortal, and stabbing is a way of killing the mortal next to one, then an act of that type will satisfy Brutus's desire. The beliefs in the motivating complex "close the gap" between the type of act motivated and the motivating desire.[2]

Our goal in this essay is limited. We assume that some complexes of cognitions motivate acts. This requires that the complexes cause the acts in the appropriate way, and also that they *rationalize* the acts. We try to provide a sufficient condition for rationalization. The basic idea is very simple. One of the cognitions must be a desire that a certain result be accomplished. The content of the beliefs, perceptions and other doxastic cognitions must be *i)* that the type of act performed is a way of accomplishing that outcome, under certain circumstances, and *ii)* that those circumstances obtain. We do not attempt to say what under what conditions a complex of cognitions causes an act in the right way to motivate.

Working out this simple idea is not straightforward, however. The properties of cognitions that determine how the body of the agent changes are local, as are the features of bodily changes that are directly caused by these cognitions. The external circumstances of the agent have no bearing on the causal transaction. But the familiar ways with which we classify acts and the contents of cognitions are *circumstantial*. The contents we use to characterize cognitions are determined in part by circumstances outside of the agent, and it is in this circumstantially determined content that the crucial links between cognitions and acts occur. The motivation we ascribe to Brutus depends on it being Caesar that he desires to kill, Caesar that he believes to be to his left, and Caesar whom he will stab if his belief is true. Our normal ways of classifying acts (stabbing the man to one's left; killing Caesar) depend in part on their results, and not

merely on the type of movements involved. Useful as these circumstantial methods of classification are for ordinary purposes, they can obscure the relationships among contents of cognitions and types of acts. We shall work through a series of versions of our sufficient condition, explaining the problems that arise for each from the circumstantial nature of classification.

Our account emphasizes the importance of relations between actions, such as one action being a way of doing another.[3] We will begin by formulating, in the next section, a preliminary version of our conditions in terms of this relation, together with a common sense conception of beliefs and desires. In §3, we improve on this account by developing a more explicit account of the structure of beliefs that allows us to discriminate between the contribution of the mental state and the contribution of circumstances. We note two defects in this improved formulation, both deriving from the circumstantial classification of acts, which leads to lack of attention to the movements an agent performs in acting. We attend to movements in §4 and use the theory we develop there in formulating a new version in §5. We end, in §6 by constructing a version that combines the virtues of its predecessors and, in terms of this last account, provide a suitably tedious reconstruction of Brutus's tyrannicide.

2 SCR-0

We distinguish between actions and acts. Actions characterize agents at times; we shall take them a species of properties.[4] Acts are concrete particulars. Our preliminary version of a sufficient condition for rationalization (**SCR-0**) is developed in terms of a species of actions we call *accomplishments* and the *way of* relation between accomplishments.

Accomplishments are properties of individuals at times, identified by the results of their acts, such as bringing it about that Caesar is killed. "Brings it about" corresponds to an operator that forms properties from propositions. So conceived, accomplishments have much in common with propositional attitude properties, such as believing that Caesar is a tyrant, and this makes it possible to state conditions on the relationship between the types of action and the contents of cognitions. We symbolize the accomplishment of bringing it about that **P** with $\mathcal{B}[\mathbf{P}]$.[5]

As normally understood, "killing Caesar" implies "bringing it about that Caesar is dead", but it cannot be identified with it. "Killing Caesar" imposes further conditions on manner and circumstances. In general, accomplishments don't exactly correspond to actions as characterized by ordinary locutions. For this reason, use of the somewhat stilted "bringing it about" locution usually suggests that the act in question doesn't fit the standard paradigm; in particular, it is often used to suggest a certain indirection, in particular the involvement of another agent and his acts. It is not our aim in this paper to analyze ordinary locutions, nor to honor them. We use "bringing it about" with no such suggestions of deviation from the paradigm and we use ordinary action locutions as if they matched accomplishment locutions more closely than they do.

Accomplishments need not be intended. Later we will consider a case in which Brutus makes a mistake and kills Marc Antony and one in which he ends up stabbing himself rather than Caesar. His accomplishments of killing Marc Antony (bringing it about that Marc Antony is dead) and stabbing himself were not intended, but on our usage they are still accomplishments.[6]

Brutus's killing of Caesar is accomplished by way of his stabbing the man to his left. This can also be thought of as an accomplishment, bringing it about that the

man to his left is stabbed. Accomplishing the latter is a way of accomplishing the former only in certain circumstances. Thus we can think of the *way of* relation as a relation holding among accomplishments and circumstances: bringing it about that **P** is a way of bringing it about that **Q**, given **C**. Later, we shall have more to say about this relation—and about its relata. For now we take these concepts as primitive. We symbolize this with $WO(\mathcal{B}[\mathbf{P}], \mathcal{B}[\mathbf{Q}], \mathbf{C})$.[7]

Given these notions, we formulate our first version as follows:

SCR-0 An act A involves an agent α at a time t bringing about a result that **P**. A motivating complex K of cognitions of α at t rationalizes A if it contains the following:

1. a desire that **R**;
2. a belief that α is in circumstances **C**;
3. a belief that $WO(\mathcal{B}[\mathbf{P}], \mathcal{B}[\mathbf{R}], \mathbf{C})$.

A will be *successful* with respect to **R** if the beliefs (2)-(3) are true.

In our first example, Brutus brought it about that the person to his left was stabbed. In line with **SCR-0** this act was motivated as follows:

1. Brutus had a desire that Caesar be dead;
2. Brutus had a belief that Caesar was mortal and that Caesar was to his left;
3. Brutus had a belief that bringing it about that the person to his left is stabbed was a way of bringing it about that Caesar was dead, given that Caesar was mortal and Caesar was to his left.

The last clause of **SCR-0** is an attempt to capture the requirement that the beliefs in a motivating complex for an act close the gap between the act caused by the complex and the motivating desire. It is important to note that it provides only sufficient conditions for success, not necessary conditions. Lucky agents often achieve their goals, even though the beliefs that (partially) motivate their acts are false.

This requirement of closing the gap imposes a heavy burden on an agent's beliefs. We can see that much more than we have listed has to be included in Brutus's beliefs, if every way he can fail is to be represented by a belief that could turn out to be false.

In a complete theory of motivating complexes for types of agents, this burden would be lessened in (at least) two ways. On our conception, an account of motivation is a part of a theory of action for a type of agent. Agents are attuned to (or, in the case of artificial agents, designed for) a certain limited range of environments, and their cognitive apparatus is used to pick up and store information about factors that vary within those limits.[8] All that should be required of a motivating complex for an act is that within that range the truth of the beliefs (and other "positive doxastic attitudes") guarantees success. The uniformities that are definitive of the environment to which the agents are attuned are things they act in accord with; these facts are *reflected* in the way these agents are put together physically and cognitively. But they are not (or need not be) explicitly represented.

The remaining burden needs to be shared among a variety of positive doxastic attitudes, only the more explicit and nontransitory of which would usually be called "beliefs". For example, when Brutus stabs Caesar, he will rely on subtle perceptual and kinaesthetic cues to guide his hand; this reflects both awareness of circumstances as is required by (2) and knowledge of the sort required by (3), about just what Brutus

must do to accomplish his goal. But we would not ordinarily call the transitory doxastic states involved in such perceptual/motor coordination "beliefs".[9]

We make some distinctions among types of attitudes later (§5). First, however, we are going to look at a problem that arises at the level of explicit belief.

3 SCR-1

Suppose that the situation was this. As before, Brutus wanted to kill Caesar. He stood next to Caesar in the Forum one Ide of March. Brutus remembered to bring his knife but he forgot his spectacles. He looks to his left, sees a particular individual he can't identify; he believes that individual to be to his left. Given that the individual was Caesar, does Brutus satisfy the second part of (2)?

According to a familiar and plausible account of the semantics of belief attributions, he does. According to that account, (2) requires him to believe a singular proposition involving Caesar and he does believe it. But intuitively, Brutus didn't have the right belief to rationalize his act. He didn't satisfy the conditions that (2) and (3) were aiming at. He didn't realize that stabbing *that man*, the one standing to his left, was a way of killing Caesar, for he didn't realize that *that* man, the one to his left, was Caesar.

This is a problem of *unreflected identity*, a species of problem that derives from the circumstantial nature of reference. A person may clearly have two quite different beliefs about the same object without knowing it, because the same object is presented to him in different ways in different circumstances. This is what has happened to Brutus in our second example. If the contents of beliefs are individuated in terms of the objects they are about, one cannot get at what is different about these beliefs.

To get at the conditions that (2) and (3) are aiming at, we need an account of belief

that finds more structure than is provided by agents, times and singular propositions.[10]

Beliefs and desires are concrete cognitive structures: they are particulars that belong to an individual agent, come into existence at a particular time, endure, and go out of existence. They stand in complex causal relations to an agent's acts, perceptions, and other cognitive structures and abilities. These causal properties of beliefs are what account for the fact that beliefs can be classified in terms of their propositional content. A belief constrains an agent's reasoning and action in a way that is conducive, if the belief's content is true, to the agent's achieving what she desires.

Beliefs are structured entities that contain *ideas* as components. Ideas, like beliefs, are concrete cognitive particulars. These ideas represent objects in the world: properties and relations, individuals, etc. Among the ideas are *notions*, which stand for individuals.[11] We take Brutus's belief that Caesar was a tyrant to involve two ideas of his, associated in the way that yields a belief.[12] One is a notion of Caesar, the other an idea of being a tyrant. His belief that Caesar was smart involves the same notion, and a different idea. The fact that the two beliefs involve the same notion means that the beliefs cognitively interact as is appropriate for two beliefs about the same person. Brutus will not act as if he has two acquaintances, one a tyrant and one smart, but as if he has a single acquaintance who is both a tyrant and smart.

Any system of storing information about particular individuals will have some analogue of notions. It is hard to imagine such a system that does not allow for the possibility of two notions or notion-analogues for the same object. Even in so careful an operation as the Social Security System, one person occasionally ends up with two numbers. When this happens, information about the person is entered in two files, and is not pooled as it should be. This sort of thing is what we see as having happened

in the various puzzle cases discussed in the philosophical literature, such as Kripke's Pierre.[13] The fact that a system has two files on the same person is not always a fault, however. Multiple files that can be linked provide a way of dealing with partial information. Suppose you lose your card and go to a Social Security office to get a new one. The clerk takes your name and address and other pertinent information and enters it in an address file. At this moment the Social Security System has two files on you, one keyed to your number, one to your name and address. When the clerk uses her computer to find your number, the two files will become linked; the fact that they are of the same person will be reflected (and maybe explicitly represented) in the system, and the information will be pooled.

In our examples, Brutus has two notions of Caesar. One he has had for a long time. It is associated with ideas of such properties as being a tyrant, having once conquered Gaul, being called "Caesar" and the like. Most Romans had a similar notion of Caesar, one associated with ideas of these same properties. Brutus acquired a second notion of Caesar, when he saw the man to his left. This notion of Caesar is associated with ideas of such properties as being next to Brutus, being a man, wearing a toga, and the like. In the original example these two notions were quickly linked: Brutus recognized Caesar.

In our second example, if Brutus found his spectacles and put them on, he would recognize the man to his left as Caesar. His two notions would then be linked—as they are in the first example—and ideas associated with the one would tend to become associated with the other.

Brutus's two notions play importantly different roles in the way he deals with information. We have methods of acquiring information about individuals and acting upon them, that depend on their standing in more or less specific relations to us. For

example, one can find out about the person to one's left by turning to one's left and looking; one can can kill the person to one's left by turning to one's left and stabbing. We call such relations *epistemic/practical* relations.[14] Brutus's first notion is not associated with any of these relations. It provides Brutus with a motive for killing, but not a method. Brutus's second notion is associated with these relations; it provides him with a method of killing (turning to the left and stabbing) but not a motive. We call a notion that is associated with an epistemic/practical relation **R** an *R-notion*; hence we speak of Brutus's *person-to-the-left* notion. Brutus already had a notion of Caesar when he came to the Forum; he acquired a person-to-the-left-notion after getting there. In our first case, these notions were quickly linked, because he was able to associate ideas with the second notion sufficient to identify Caesar. At this point motive was joined with method, and Brutus stabbed Caesar. In the second case, in which Brutus forgot his spectacles, the notions remained unlinked, method and motive were kept separate, and the opportunity was lost.

When an agent's notions of a given object are all linked, we'll say that agent is *well-oriented* with respect to that object.[15] Our ordinary apparatus of belief reporting, which classifies beliefs by agent, time, and singular proposition believed, works well enough for agents that are well-oriented towards the relevant objects, as Brutus was in our first example: "He believed Caesar was to his left, desired to kill him, and so turned to his left and stabbed him." This apparatus doesn't work so well for agents with unlinked notions of a given object, as in the second example. The belief cited in the first part of the of the account just given could be attributed in virtue of Brutus's person-to-the-left notion of Caesar. The desire cited in the second part could be attributed in virtue of his longstanding notion. But because these notions are not linked, the action

will not ensue. For Brutus to be motivated to stab the person to his left, the identity should show up in the internal structure of his beliefs and desires, not merely in their circumstantially determined content. Our ordinary concepts are not rendered helpless by such cases. We all understand about the possibility of unlinked notions, and can describe such cases in terms of failure to recognize, misrecognition and the like, as we did when we first described the case. But these descriptions require going beyond talking about agents and singular propositions.

Thus we characterize the beliefs and desires of our agents by their internal structure—the ideas and notions of which they are composed. Such internal structures have their content in virtue of *anchors*, functions from ideas and notions to relations and individuals. Anchors represent the circumstances that determine the reference of these ideas. The propositions believed and desired are those obtained by anchoring these internal structures with anchors that represent the connections of the ideas to their referents[16].

In our examples we would represent Brutus's desire as

Q: $\langle \text{Dead}, f \rangle$

where f is Brutus's longstanding notion of Caesar and Dead is his idea of the property of being dead. Notice that "$\langle \text{Dead}, f \rangle$" identifies a complex cognitive particular by specifying its components and suggesting its structure. The content of this desire—*what* Brutus desires—is a proposition determined by his desire and the way his notions and ideas are connected to the world. Where a is an anchor that represents Brutus's connections, we symbolize this proposition as $Q\,[a]$.[17]

For the most part, we treat ideas of relations and notions—ideas of individuals—differently in this paper. One could imagine cases in which Brutus had two unlinked ideas of the property of being dead, and in which this had to be taken account of to

properly appreciate why he had or lacked a motivation in that case. Nevertheless we simply represent ideas of relations in beliefs with the relation itself. We do discuss the multiple ideas one might have of a given property in §5, in connection with our discussion of know-how.

Brutus's belief as to his circumstances we represent as:

C: $\langle\langle \text{Left-of}, i, n\rangle, \langle \text{Mortal}, n\rangle\rangle$

where n is a perceptual notion (of Caesar) acquired by looking to his left at the Forum.[18] It is associated with the epistemic/practical relation of being-to-the-left-of. This notion is a component of beliefs he managed to form about the fellow he saw in spite of his poor vision. The notion i is Brutus's self-notion. Not every belief about one's circumstances needs to involve one's self notion; for example Brutus's belief that Caesar is mortal, does not. But some must. Brutus has to believe that Caesar is to *his* left.[19]

What Brutus believes in virtue of this belief and his connections is $C[a]$. This is a singular proposition with Caesar and Brutus as constituents—a proposition that, as we noted, is too coarse-grained for our purposes.

Brutus realized that stabbing the mere mortal to his left was a way for him to bring it about that this person gets killed. It was, in fact, a way for him to bring it about that Caesar gets killed. But this depends on an identity not reflected in Brutus's beliefs. It is reflected only in the common value that the anchors provide for the two different notions. Thus, Brutus's cognitions did not rationalize his stabbing the person to his left.

Suppose now, as in our original example, that Brutus manages to recognize Caesar. In this case, he has the additional belief:

C': $\langle$Identical$, f, n\rangle$.

This is part of the motivating complex Brutus had in the first case. Note that this belief requires that Brutus recognize the person to his left. In the original case, this identity was true and Brutus recognized it to be true. In our second case, it was was true but he failed to recognize it. A third possibility is that he believed it, but it was not true—perhaps he mistakenly stabbed and killed Marc Antony.

Given this new belief, we would expect him to have a belief of type (3) (from **SCR-0**). We need to characterize that belief.

In order to have belief (3) explicitly, Brutus needs more ideas and notions than we have given him thus far. We assume that he has the ability to form the following ideas and notions:

Notions of propositions. Where P is a belief of α's and a is the anchor for P, $P[a]$ is a proposition. α will be able to form a complex notion of this proposition that involves the same ideas and notions as P. This notion we symbolize as P^*. The anchor a will yield the same proposition as the reference of the notion of P as the proposition it determines to be the content of P: $P[a] = P^*[a]$. a reflects the circumstances that determine the reference of the notions in P, which in turn determine the proposition, $P[a]$, that α believes in virtue of having the belief P. In these same circumstances, P^* should refer to this same proposition. (The relation between the belief P and the notion P^* is analogous to that between the sentence "Elwood is sleeping" and the noun phrase "that Elwood is sleeping". The former expresses a proposition to which the latter refers.)

Ideas of accomplishments. For a proposition **P** of which α has a notion P^*, α will

have the ability to form a certain idea of the relation $\mathcal{B}[\mathbf{P}]$ which we symbolize $\mathcal{B}[P^*]$.

An idea of *way of***.** This is an idea of a relation between pairs of accomplishments and circumstances.

Now we can say more about what it is Brutus believes, in the case in which he recognizes Caesar.

We represent Brutus as having the following belief.[20] Where $P = \langle \text{Stabs}, i, n \rangle$,

$$\langle WO, \mathcal{B}[P^*], \mathcal{B}[Q^*], \langle C^*, C'^* \rangle \rangle$$

We can now reformulate our account as follows:

SCR-1 An act A involves an agent α at time t bringing about a result that **P**. A motivating complex K of cognitions rationalizes A if it contains the following cognitions:

1. a desire: $\langle Q^k, n_1, \ldots, n_k \rangle (= Q)$;
2. a belief: $\langle C^n, i, \ldots n_{n-1} \rangle (= C)$;
3. a belief: $\langle WO, \mathcal{B}[P^*], \mathcal{B}[Q^*], C^* \rangle$.

Let a be an anchor that represents the way α's ideas and notions are connected to the world. A act will be *successful* with respect to the content of α's desire, $Q[a]$, if the contents of α's beliefs (2)-(3) as anchored by a are true.

SCR-1 is an improvement over **SCR-0** with respect to the problem of unreflected identities. It is itself susceptible, however, to another problem, which we will call

the *problem of the wrong movement*. This problem takes two forms. In the first, we imagine that Brutus's beliefs, as far as we list them in **SCR-1**, are true, but that when he goes to stab the person to his left, he makes the wrong movement, and stabs himself in the left arm. He simply does not make the movement he was trying to make. In the second, he makes the movement he was trying to make, but misses Caesar completely. The movement he tries to make is not the right one; he is wrong in thinking that *that* movement would get Caesar stabbed. In each case, he doesn't succeed in bringing about Caesar's death.

None of the beliefs we have listed are false in either of these cases, so there is something missing from **SCR-1**. Brutus's motivating complex needs to reflect which movement he is trying to make, and what he thinks its effect will be. So far, these factors have been omitted from our account.

This missing factor is connected with another problem with our account. The connection between the act A and the motivating complex for A is made at the wrong level. We have made it in terms of the accomplishment, bringing it about that P, that the agent takes to be a way of bringing about the motivating desire Q. But the connection should instead be made at the level of movement.

To see why this is so, let us reflect on the problems that would beset us if we made the connection at the "highest" level, that of the motivating desire. To do this we would replace **P** with **Q** in the beginning of **SCR-1**:

> An act A involves an agent α at time t bringing about a result that **Q**.

But this isn't right, because the act to be rationalized may be one that does not conform to the desires of the agent at all. Consider our third Brutus example, where he kills Marc Antony by mistake. In this case the act to be rationalized is the killing of

Marc Antony; this is what Brutus did, even though it was not what he intended to do. The connection between the act motivated and the motivating complex can be made at the level of the motivating desire only when the act is successful.

But the same problem comes up at the level of the accomplishments which are taken to be a way of doing what one intends. This point becomes clearer, the more bizarrely unsuccessful we imagine Brutus's act to be. Suppose that no one was next to Brutus in the Forum; he was fooled by a play of light and shadows. Here the act was neither a killing nor a stabbing of anyone; the only possible link between act and motivating cognitions was the movement.

These two problems are closely related: they both point to the need for considering movements. It was the lack of movements that prevented us from accounting for what went wrong in the case where Brutus stabbed himself. And it is at the level of movement that the connection between act and motivating cognitions has to be made.

It should not be surprising that an account that is to accommodate content and causation cannot ignore movements. When we think of beliefs and desires causing things to happen, we are thinking of them as mental particulars causing other particulars. Those mental particulars have effects on bodies and tyrants only by having an effect on the movements of the body in which they are lodged. Ultimately, it is the movement that must be made intelligible by the cognitions that motivate it.

There is a temptation to think that the cognitions need not concern movement. But this would amount to mental causation at a distance. There would be a gap between the motivating cognitions and the act they cause. Brutus's cognitions can cause him to stab his neighbor only by causing him to move in one way or another. Which movement will work to constitute a stabbing depends on circumstances. If his cognitions do not

reflect considerations that would favor one sort of movement over another, they do not render the fact that he made one movement rather than another intelligible.

4 From Movements to Accomplishments

4.1 Executions

To deal with movements, we need to recognize a second kind of action: *executions*. Executions do the work of basic acts in Goldman's theory; but our distinction is ontological rather than epistemological.[21] Executing M, like bringing it about that **P**, is a property of an individual at a time, which we'll symbolize with $\mathcal{E}[M]$. Accomplishments are typically circumstantial properties, individuated by the results brought about, which in turn depend on intervening circumstances. In contrast, executions are individuated by the type of movement executed, whatever the result, and so are not circumstantial. Executions are properties that agents have at times in virtue of the types of movements they perform. Accomplishments are properties that agents have at times in virtue of the types of movements they execute and the circumstances in which the movements occur, for it is on these factors that results depend.[22]

An act involves an agent at a time *effecting* a movement of its body. When an agent effects a movement m of type M, we say the agent *executes* M. Before we can develop an account of the the relations between movements and accomplishments, we need to develop some ideas about movements themselves. We understand movement to include both lack of movement and changes that do not involve observable motion of the external parts of the body, such as tensing one's muscles, generating torque, etc. We take *effecting* as primitive. We assume that all intentional acts involve an agent effecting movements; we do not take a position on whether the converse is true.[23]

These effectings are particular, unrepeatable events of various types. The movement may itself be complex; it may involve submovements of various bodily parts and may itself consist in a series of movements of the whole body executed over an interval of time. On Davidson's view, this movement is identical with the act. On Goldman's, it is the basic act that generates the others. This issue is peripheral to our concerns.

Executions are properties of agents formed from types of movements. A theory of action for a given type of agents must include an analysis of the types of movements open to agents of that type, which will depend on its anatomy or architecture. Agents typically have systems of somewhat independent effectors or actuators, each with its own range of movements (trajectories). The movements of the whole system is determined by the movements of its parts. A theory of movement for an agent may be motivated by the circumstances in which the agent typically operates. But movement types abstract from those circumstances. Think of repairing a dot-matrix printer. The printer is removed from the system, emptied of paper and ribbon and sent to a shop. The pins can still go through exactly the same movements individually and in combination and succession as they can when they are printing. The range of movements is the same whether the printer is connected in the office or disconnected in the shop, but the effects are quite different.

M needs to be a total movement type, one that specifies everything relevant about what the agent is doing. At first this seems a bit counterintuitive. When we think of Brutus stabbing, we focus on the movements of his arm and hand. We don't focus on his feet. But if he were to take a big step back while he swung his arm up, he wouldn't stab anyone. To avoid such cases of interference by movements of parts of the agent's body, M needs to be total; it needs to specify what certain parts of the body do, and to

limit what the others do.

4.2 *Modes* and *Ways*

We use the term *mode of* for the relation between executions and accomplishments. We say that $\mathcal{E}[M]$ is a *mode of* $\mathcal{B}[\mathbf{R}]$ in **C** iff executing a movement of type M in **C** causes it to be the case that **R**. We use the notation $MO(\mathcal{E}[M], \mathcal{B}[\mathbf{R}], \mathbf{C})$.

Consider Brutus's accomplishment. He killed Caesar at a certain time. This accomplishment does not, by itself, determine which movement type needs to be executed. Different movements will get the job done in different circumstances. Even once we fix the circumstances, a variety of movements might result in Caesar's death. These we call the *modes of* killing Caesar, in those circumstances.

We are now in a position to say more about what it is for one action to be a *way of* performing another action. As we have mentioned, this relation may hold in some circumstances, but not in others, so we need to define it relative to circumstance. We say that bringing it about that **P** is a way of bringing it about that **Q** in **C** if any M whose execution is a mode of bringing it about that **P** in **C**, is also a mode of bringing it about that **Q** in **C**. In our notation:

$WO(\mathcal{B}[\mathbf{P}], \mathcal{B}[\mathbf{Q}], \mathbf{C})$ iff :

$(\forall M)$ if $MO(\mathcal{E}[M], \mathcal{B}[\mathbf{P}], \mathbf{C})$ then $MO(\mathcal{E}[M], \mathcal{B}[\mathbf{Q}], \mathbf{C})$.

We will discuss some examples to show how this definition works.

Consider Brutus again. Stabbing the man to his left was a way for him to kill Caesar in the circumstances he was in. (We remind the reader that by stabbing we mean plunging a knife into the heart). There were a variety of types of movements

whose execution would have resulted in the man to his left being stabbed. He could have done it more or less quickly than he did, with his fingers more or less tightly wrapped around the knife, and so forth. All of the ways of moving that were modes of stabbing the man to his left were also modes of killing Caesar.

Let's turn to a case posed by Fred Dretske, which appears to raise a problem for our definition. Suppose that Elwood and Ethel live apart, in homes that get cable TV from the same company. The wiring to their houses is independent. The fuse blows at the transmitting station. Later, the technician flips a switch to turn it back on. The movement he makes is a mode of bringing the signal back to Elwood, and also a mode of bringing the signal back to Ethel. So far, so good. But note that any other movement he might make that would get the signal restored to Elwood would also get it restored to Ethel. That is, every mode of restoring the signal to Elwood is a mode of restoring the signal to Ethel and vice versa. But then, by our definition, bringing the signal back to Elwood is a way of bringing it back to Ethel and vice versa. This seems counterintuitive.

The problem here is that we have lost track of the circumstances. We need to distinguish three relevant sets of circumstances. One comprises the facts connecting the transmitting station to Ethel's television, another comprises the facts connecting the transmitting station to Elwood's, and the third includes both of these. Relative to the first and second sets of circumstances, we don't get the unwanted result. Relative to the first set, the technician's movement is a mode of restoring the signal to Ethel's TV but not a mode of restoring it to Elwood's. Relative to the second set, the technician's movement is a mode of restoring the signal to Elwood's TV, but not a mode of restoring it to Ethel's. These facts are the basis for the intuition that doing the one is not a way

of doing the other.

When we focus on the third set, our intuitions shift. If he had clearly in mind that if he takes care of Ethel's problem, he takes care of Elwood's too, he might well say that doing the one was a way of doing the other.

One needs to keep in mind that the circumstances in our various relationships are not total circumstances. Thus, different circumstances need not mean different possible worlds. In this case, all three sets of circumstances obtain. But they are relevant to different results, and will be taken into account in thinking about different accomplishments.

5 SCR-2

We return to the structure of motivating complexes. The study of movements and their relationship to accomplishments of the last section was motivated by the shortcomings of **SCR-1**. There was the problem of the wrong movement, and the problem of capturing the connection between an act and the cognitions that motivate it.

Before formulating a version that resolves these problems, we need to add to our inventory of mental equipment. We introduce the concept of *volition* to execute M.[24] When beliefs and desires motivate acts, they do so by causing a volition to execute a movement of a certain type, which are the proximate causes of the movement.

Suppose that, in fact, executing M is a mode of bringing it about that **P** in circumstances **C**. And suppose that, when he wants to bring it about that **P**, and believes that he is in circumstance **C**, α forms a volition to M. We might say that he *knows how* to bring it about that **P**, especially if he succeeds. We conceive of this "know-how" as a doxastic cognitive particular, involving ideas and notions:[25]

α has the know-how: $\langle MO, \mathcal{E}[M], \mathcal{B}[P^*], C^* \rangle$.

To have know-how agents need ideas of various executions. Not just any old idea of an execution will do. The idea must be *executable*: the sort of idea that, perhaps aided by perceptual and kinaesthetic information, could guide the formation of a volition.

For example, Brutus might think of a certain type of movement as "the type of movement that will be required to kill Caesar when I have the opportunity". This may be an idea of the very type of movement he needs to execute. But it is not the sort of idea of a movement that can guide the formation of a volition.

Now suppose that α is just as he was, two paragraphs above, when he had know-how, but in fact M is not a way of bringing it about that P. α has, as it were, a bit of false know-how. But "false know-how" won't do, for "knowledge" implies truth. We seem to need, and hereby introduce, *belief-how*, an attitude an agent can have to a proposition about modes, whether false or true.

With this additional equipment, we can now state **SCR-2**.

SCR-2 An act A is determined by an agent α, a time t and movement m. A complex K of cognitions *motivates* A if it causes a volition v to M (in the right way) that causes m (in the right way), and K *rationalizes* A. K rationalizes A if K contains:

1. a desire: Q;
2. a belief: C;
3. a belief-how: $\langle MO, \mathcal{E}[M], \mathcal{B}[Q^*], C^* \rangle$.

Let a be an anchor that represents the way α's ideas and notions are connected to the world. A will be *successful* with respect to the content of his desire, $Q[a]$,

if the contents of α's beliefs (2)-(3) as anchored by a are true and m is of type M.

SCR-2 handles the problem of the wrong movement. In the case in which he stabs himself in the arm, Brutus's volition to execute a certain type of movement does not cause a movement of that type. In the case in which he misses Caesar completely, Brutus's volition to execute a certain type of movement leads to a movement of that type, but his belief (3), pertaining to the results of such a movement, is false.

As promised, **SCR-2** connects the motivating complex with the act it motivates at the level of movement, not at the level of desired result. The desire that **Q** may be part of the motivating complex of an act with the wholly unintended result that **Z**. In the Marc Antony case, the desire that Caesar be dead enters into the motivating complex for an act that brings it about that Marc Antony is dead. Our analysis allows for cases in which there is no relevant accomplishment at all, as in the case in which Brutus is fooled by a play of light and shadows. His act does have various irrelevant results, such as making some molecules move, but it kills no one.[26]

According to **SCR-2**, in the case in which he failed to execute the movement he was trying to execute, Brutus's stabbing himself was rationalized. Stabbing himself was something he accomplished because of a movement caused by his motivating complex. On our account the connection between a motivating complex and the act it rationalizes is made at the level of the particular movement. None of the accomplishments achieved need be among the ones aimed at. And none of the executions achieved need be among the ones tried.

This feature of our account is consistent with the strategy that we have followed all along. Motivations do not fail to be motivations for acts, simply because the attitudes

involved in the motivating complex are not satisfied. An act does not cease to be intentional because it is unsuccessful. If Brutus's self-destructive movement had been an involuntary twitch, Caesar might have felt bad about Brutus's injured arm. But if Caesar had known the facts of the case as imagined, he would feel betrayed.

6 SCR-3

We believe that **SCR-2** states sufficient conditions for a motivating complexes to rationalize an act. However, it is misleading in an important way. In our examples, at the time Brutus acted, he believed that a movement of a certain type was a mode of bringing about Caesar's death, as required by **SCR-2**. But this belief depended on a number of other beliefs, which connected intermediate accomplishments with the movement and the desired end, for example, the accomplishment of stabbing the person to his left. The sorts of circumstances that link the executed movement with this intermediate goal are of a quite different kind than those that link the intermediate accomplishment with the desired end. **SCR-2** obscures the role of beliefs about these intermediate accomplishments and the circumstances necessary for each to be a way of doing the next. **SCR-3** explicitly highlights the role of these beliefs. Note that it reduces to **SCR-2** in the case in which there are no beliefs about intermediate accomplishments.

SCR-3 An act A is determined by an agent α, a time t and movement m. A motivating complex K of cognitions *motivates* A if it causes a volition v to M (in the right way) that causes m (in the right way), and K *rationalizes* A. K rationalizes A if K contains

1. desires: $P_1, \ldots, P_n$;

2. beliefs: $C_1, \ldots, C_n$;

3. a belief-how: $\langle MO, \mathcal{E}[M], \mathcal{B}[P_1^*], C_1^* \rangle$, and

4. for each $1 < i \leq n$, a belief: $\langle WO, \mathcal{B}[P_{i-1}^*], \mathcal{B}[P_i^*], C_i^* \rangle$

Let a be an anchor that represents the way α's ideas and notions are connected to the world. A will be *successful* with respect to the content of α's desire, $P_n[a]$, if the contents of α's beliefs (2)-(4), as anchored by a, are true and m is of type M.

Now we will return to our original example using **SCR-3** to construct a fairly tedious account of what we described in a short paragraph using unadorned English. We will go through the clauses of **SCR-3** in order.

Brutus desires (1) that Caesar be dead; this desire we have represented as $P_n : \langle \text{Dead}, f \rangle$. Recall that f is Brutus's longstanding notion of Caesar. We may assume that he has had this desire for some time, and that he has recently formed the intention to act upon it. He has done some planning, and come up with a strategy that will work if he can get near Caesar while holding a knife. Given his strategy and ulimate goal, when he gets near Caesar, he forms the subgoal of stabbing the person to his left.

Brutus's beliefs (2) cover the circumstances required for his movement to be mode of stabbing the person to his left, C_1, and those required for stabbing the person to his left to be a way of killing Caesar, C_2.

The former will include, for example, that Brutus has a knife in his hand. This belief will involve Brutus's self-notion i, and (at least normally) his tactual notion k of the object in his right hand:

$\langle \langle \text{Holds-in-right-hand}, i, k \rangle, \langle \text{Knife}, k \rangle \rangle$

The latter include a bit of common-sense that only Caesar's most ardent admirers could have doubted, that he was mortal, as well as the fact that Caesar was to Brutus's left.

$\langle\langle$Left-of$, i, n\rangle, \langle$Mortal$, n\rangle, \langle$Identical$, f, n\rangle\rangle$.

When Brutus acts the movement he tries to execute is M: a coordinated movement in which he turns his body to the left, raises his right arm and plunges it down in a powerful arc, while moving his own left arm out of the way.

Brutus has a belief-how (3) that executing M is a mode of stabbing the person to his left, in certain circumstances. We represent this belief as

$$\langle MO, \mathcal{E}[M], \mathcal{B}[P_1^*], C_1^*\rangle.$$

Brutus has no doubt known for a long time that executing M with a knife in his right hand is a way of stabbing the person to his left. Brutus may have learned how to stab in this way at martial arts classes as a youth, or by watching spectacles at the Coliseum. Brutus may have no words to describe M. If he were describing his plan to Cassius, he would have to show the movement in question: "I'll get beside him and do *this*." If he heard the movement described (say by a Roman anatomist), he might not recognize it. But he has an executable idea of it.

Brutus believes (4) that stabbing the man to his left is a way of killing Caesar, given that Caesar is the man to his left, and is mortal. Where P_1, P_n and C_2 are as above, we represent this belief as

$$\langle WO, \mathcal{B}[P_1^*], \mathcal{B}[P_n^*], C_2^*\rangle$$

This belief is derived from a belief he has also held for a long time, that stabbing Caesar is a way of killing him, given that he is mortal. This belief in turn is an instance of the more general belief that stabbing a mere mortal is a way of killing him. The instantiation to Caesar would naturally be triggered by Brutus's resolving to kill him, based on the longstanding desire that he be dead. As we imagine things, Brutus pondered the chances of his getting up close to Caesar in the normal course of events that day, and then perhaps planned to make sure it happened if it weren't otherwise going to happen.

We imagine Brutus's belief C_1 coming mainly from perception just before he acts. He sees someone to his left. He knows that he has something in his hand because he can feel himself holding it; he remembers that he picked up a knife and hasn't let go of it, so he knows that the thing he feel is a knife. His belief C_2 is also based partly on perception, and partly on memory. He needs to recognize the man to his left as Caesar. And he has believed for a long time that Caesar, like all men, was mortal.

7 Discussion

As we noted in the preceding section, the beliefs in Brutus's motivating complex are of various types, but at least as we imagine them, they all are triggered by Brutus's attempt to carry out a certain plan.

In our tedious reconstruction, we imagined Brutus having formed and adopted a plan to kill Caesar by stabbing him, and then acting on that plan. We think of plans as themselves complex mental structures, consisting of decisions or intentions (typically forward-looking intentions), structured by relations between the actions that are the objects of those intentions, e.g. by *way of* and *mode of* relations between them. Beliefs

as to these relations support this structure. Some of these beliefs are brought into play by memory, others by inference—these processes, in turn, triggered and guided by the means-ends problems posed by planning. Plans are typically partial; in terms of our account, this means that they are not typically grounded in beliefs as to *modes*. (Of course they may also be partial with respect to *ways*.)[27]

The partiality of the plan, and the conditional nature of various of its constituent intentions, pose a series of cognitive problems that the agent must solve in carrying out the plan. These will involve both recognition of appropriate circumstances and appropriate instantiation of general beliefs as to modes and ways. Thus the adoption of the plan will shape the agent's cognitions: leading him to be on the outlook for opportunities (and obstacles) in his changing circumstances and to figure out specific ways and executable modes appropriate to those circumstances.[28]

In our simple example, we imagine that a certain general belief plays a role in Brutus's decision to kill Caesar by stabbing him. This decision then poses the problems of getting hold of a knife and getting close enough to Caesar to stab him with it. We need not imagine, though, that Brutus must figure out in advance how to stab Caesar when he gets close enough. Nor do we need to imagine that Brutus considers in advance the problem of recognizing that the circumstances at the time of the act fall within the range of circumstances in which stabbing is a way of killing. The plan can remain partial. The most that seems to be required is that Brutus believe that, if and when he has gotten in the right position relative to Caesar, he will recognize himself as being in those circumstances and he will know how to stab him in those circumstances.[29]

Footnotes

[1]We accept the contention of Bratman and Bratman, Israel and Pollack that intentions play a role in motivating complexes; see Michael Bratman *Intention, Plans and Practical Reason* (Cambridge: Harvard University Press, 1987) and Michael Bratman, David Israel and Martha Pollack, "Plans and Resource-Bounded Practical Reasoning," in *Computational Intelligence*, Vol. 4, No. 4, 1988, ps. 349-355. Reprinted in *Philosophy and AI: Essays at the Interface*, R. Cummins and J. Pollock, eds., (Cambridge: MIT Press, 1991) ps. 7-22. For the purposes of this paper, however, we will for the most part ignore intentions. But see the last section.

[2]If one thinks that the beliefs and desires in such a motivating complex constitute the reason for the act, then the present conception is in agreement with Davidson's thesis that reasons are causes. If one thinks of the reason as the contents of the beliefs and desires in the motivating complex, then reasons are not causes. If one thinks that "reason" shares the ambiguity of "belief" and "desire," sometimes meaning cognition, sometimes the content thereof, then reasons are causes in one sense, and not in another. These quibbles aside, our conception of a motivating complex is clearly a spiritual descendant of Davidson's primary reasons in "Actions, Reasons and Causes," reprinted in *Essays on Actions and Events* (New York: Oxford University Press, 1980).

[3]In this emphasis our account resembles Alvin Goldman's in *A Theory of Human Action* (Englewood Cliffs N.J.: Prentice Hall, 1970), to which it is much indebted.

[4]We follow Goldman in not restricting properties to static properties. Our actions are what Goldman calls *act-types* and *act-properties*. See *A Theory of Human Action*, p. 10. See also Jaegwon Kim, "On the Psycho-Physical Identity Theory," *American*

Philosophical Quarterly, III, 1966, p. 231.

[5]We emphasize that we recognize there are also quite important differences between cognitive attitudes, like belief and desire, and bringing it about.

[6]We shall argue (see §5) that, though in each case the accomplishment is unintended, the act in virtue of which Brutus brings about the unintended result is motivated.

[7]For more on these concepts, see David Israel, John Perry and Syun Tutiya. "Actions and Movements," in *Proceedings of IJCAI-'91*, (Mountain View, CA: Morgan Kaufmann), August, 1991.

[8]See David Israel and John Perry. "Fodor and Psychological Explanations," in Barry Lower and Georges Rey, eds., *Meaning in Mind*, (Oxford: Basil Blackwell, 1991) ps. 165-180, John Perry, "Circumstantial Attitudes and Benevolent Cognition," Jeremy Butterfield, ed., *Language, Mind and Logic*, (Cambridge: Cambridge University Press, 1986), and David Israel, "The Role of Propositional Objects of Belief in Action, " (Stanford: CSLI Reports, 1987).

[9]Perhaps, he must also have a positive doxastic attitude that he has the requisite sensory-motor abilities. But see §5.

[10]This account is in the spirit of Jon Barwise and John Perry, *Situations and Attitudes*, (Cambridge: Bradford Books, MIT Press, 1983), chapter 10, as developed and modified in Mark Crimmins and John Perry, "The Prince and the Phone Booth," *Journal of Philosophy*, LXXXVI, no. 12 (1989), 685-711. See also "Fodor and Psychological Explanation" for a discussion of the problem of unreflected identity as it manifests itself in psychological explanations.

[11]Among the individuals are propositions; among the notions are notions of propo-

sitions. Having a belief with **P** as its propositional content has to be distinguished from having a notion of **P**. See below.

[12]There are many different ways of accounting for the structure of beliefs in general and of logically complex beliefs in particular. Thus, one could postulate that one begins with *supposings*, which are concrete structures composed of ideas and notions and that beliefs, desires, etc. are simply supposings that play specific functional roles in the cognitive architecture of agents. Alternatively, one could bypass supposings and posit beliefs, desires, etc. as directly constructed out of ideas and notions, where the functional difference is realized by different modes of construction. Thus one might associate the idea of being dead and the notion of Caesar in a doxastic mode, resulting in a belief or in an appetitive mode, resulting in a desire. The talk of ideas "associated in the way that yields a belief" suggests this conception, but is meant to be neutral. With respect to logical structure, again, there are alternatives. It may be that the ability to form conjunctive beliefs does not require any idea or notion of conjunction, but rather an ability or propensity to make certain kinds of inferences to and from noncompound beliefs. Indeed, one may not want to posit a uniform procedure; thus, one may treat disjunctive beliefs differently from conjunctive beliefs. We cannot be completely silent on such issues, but we shall be as silent as our present purposes allow.

[13]See "The Prince and the Phone Booth" for a discussion of these cases.

[14]See John Perry, "Self-Notions," *Logos*, 1990.

[15]This condition, akin to Hintikka's condition on knowing who in *Knowledge and Belief* (Ithaca: Cornell University Press, 1962), is too strong and certainly stronger than we need. The appropriate condition must be stated in terms of relevance to an agent's desires and beliefs as to means and ends. Thus, for instance, Brutus might have a

notion of Caesar derived from his childhood acquaintance with him on the playgrounds of Rome, but this notion might be unlinked to his notion of Caesar, the conqueror of Gaul and tyrant. He might have forgotten this earlier acquaintance or might never have made the connection with the Caesar he met in young adulthood. This lack of a link, however, does not affect Brutus's being well-oriented with respect to Caesar, relative to his current plan to kill Caesar.

[16] See *Situations and Attitudes* and "The Prince and the Phone Booth".

[17] *Italic* capital letters stand for particular beliefs and desires. Propositions are symbolized in two different ways. If *Q* is a particular belief and a is an anchor for it, $Q[a]$ is a proposition. Where propositions are not identified in terms of beliefs and anchors, we use **boldface**: **R** is a proposition.

[18] See fn. 10 for our studied silence on the representation of logically complex beliefs.

[19] See "Self-Notions" and "Circumstantial Attitudes and Benevolent Cognition". We could have conceived of Brutus as desiring that he kill Caesar; that is, as having the desire: $\langle \text{Kill}, i, f \rangle$. Such agent-centered cognitions do play a central role, as we have noted, but we have chosen to segregate this aspect of cognitions to beliefs.

[20] This represents part of Brutus' motivating complex; his reasoning may involve other things. For example it is at least likely that he arrives at this belief by instantiation from a general belief relating stabbing and killing. But here, we are interested only in what might be taken to be the primary form of the idea of the *way of* relation: that in which an agent α relates ideas of two accomplishments by α, as conceived under α's self-notion and an idea of α's own circumstances.

[21] See *A Theory of Human Action*.

[22]One might take the action $\mathcal{E}[M]$ to be identical with some accomplishment such as bringing it about that one executes M. We do not know whether this is a good idea or not, so we leave the matter open.

[23]There are events that share much of the structure of acts but that are not acts, for instance a tree falling. We assume that trees do not effect movements; they just move. Thus they are not agents of acts. Nothing in our theory explains the difference between acts and other movements or between agents and nonagents.

[24]We do not take a position on whether there are also volitions directed at accomplishments, that is, volitions to bring it about that P.

[25]We will be silent on how like such cognitions are to beliefs.

[26]If one takes the option described in fn. 22, there would be the accomplishment of bringing it about that Brutus has executed M.

[27]There is at least one more important relation between actions that we have ignored—that of one action *enabling* another. Being close to Caesar is a necessary condition for stabbing him via the modes available to ordinary humans. That is, executing M is a mode of stabbing the person on the left, given (inter alia) that the agent is close enough to that person—indeed, only if he is close enough. Walking over is not a mode of killing, nor of stabbing. It's a mode of getting close, given that one is where one is relative to the target. Similarly with getting a knife and stabbing. We say that one action A (execution or accomplishment) enables another A', relative to a circumstance **C**, just in case there is some action A'' (typically an accomplishment) and propositions **P** and **P′** such that $WO/MO(A, \mathcal{B}[\mathbf{P}], \mathbf{C}) \,\&\, WO/MO(A', A'', \mathbf{P}')$, where **P′** entails **P**. In the case of an enablement relation holding between two actions, there will not be a single chain of WO/MO beliefs, with compatible circumstances. Rather the enabling action

will result in a change of circumstance. Intuitively this corresponds to there being at least two acts—and two motivating complexes, linked together in a plan.

[28]For more on this, see "Plans and Resource-Bounded Practical Reasoning."

[29]We are grateful to Michael Bratman, to Fred Dretske, to many other colleagues at Stanford and SRI and to the referees for the *Philosophical Review* for criticisms and suggestions. The work on this paper was supported in part by a grant from the System Development Foundation to the Center for the Study of Language and Information at Stanford.