



## Regret and Retroaction

Jeanne Peijnenburg

Faculty of Philosophy, University of Groningen, The Netherlands  
(eMail: Jeanne.Peijnenburg@rug.nl)

**Abstract** Often, regret implies the wish not to have performed certain actions. In this article I claim that this wish can to some extent be fulfilled: It is possible, in a sense, to influence the character of actions that have already been performed. This possibility arises from combining a first person perspective with an outlook on actions as expressions of tendencies, where tendencies are identified on the basis of a number of actions. The idea is specified within the framework of Carnapian reduction sentences, but this technique is in no sense mandatory: It can be formulated in other vocabularies as well.

*Keywords* regret, action, retroactivity

### 1. Introduction

Feelings of regret and remorse seem to be universally human. It is true that some people say that they never regret anything, but such claims are generally to be understood more as stylistic hyperbole than as literal fact. Criminals lacking all remorse are as rare as they are horrifying; at any rate they can count on much less understanding than can a contrite wrongdoer. A person for whom the phenomenon of regret is foreign, who has never experienced it, is fit for the psychiatrist.

Moreover it seems that feelings of regret and remorse are typically human. Admittedly, animal behaviour exists that we might be inclined to call remorseful, in particular if it is displayed by our beloved pets. Ethologists assure us however that such behaviour expresses not remorse but rather fear of punishment. When your favourite poodle creeps on his belly to you after having stolen a steak from the kitchen, he is actually attempting a reconciliation with the aim of diminishing his punishment – there is not a whimper of remorse in his canine mind (De Waal 1996, Vollmer 1977).

Assuming that regret is felt by all humans, and only by humans, it is

scarcely surprising that psychologists and philosophers have waxed eloquent on the matter. It is striking, however, that most psychological and philosophical studies on regret are of a decision-theoretic nature. In the 1980s in particular a plethora of articles appeared about how people try to avoid future regrets by anticipating their possible occurrence. Some of these studies, for example the regret theory of Loomes and Sugden (1982), developed Savage's minimax regret rule of 1951 further, and in this sense they may be regarded as being refinements of the classical theory of the maximization of utility. Where decision theory is flaunted, phenomenology seems often like a poor relative. Phenomenological studies of regret are indeed but sparse in psychology, Janet Landman's (1993) being a rare example. In philosophy, the meagre phenomenological analyses of regret have a predominantly ethical flavour; and, in the wake of Bernard Williams (1976), many philosophers see regret as a 'moral remainder', an unfortunate but unavoidable byproduct of moral dilemmas or other situations in which we are forced to choose between displeasing alternatives.

In this paper I propose to analyze regret from a somewhat different perspective. I shall not be concerned at all with preventing feelings of regret (as in decision theory), and neither with the question as to what regret actually is (as in phenomenology). Much less shall I ask whether remorse is rational (see Betzler 1999, Hurka 1996, Dancy 1993: Chapter 7), or reasonable (see Bittner 1992), or functional (see Zeelenberg 1999). Similarly, I shall ignore the much-discussed topic of the cognitive content of feelings of regret, and also the query as to whether regret is actually necessary or even possible (if we really have good reasons for a choice, there is no need or reason at all for regret). Rather, I am particularly interested in the problem how, once feelings of regret have arisen, they can be *reduced* or even *removed*. I shall first claim that regret is accompanied by the desire for a world different from the actual one; in particular I shall defend the thesis that regret implies the wish not to have performed certain actions. After this hardly surprising contention, even if it is not accepted by everyone, I hope to make acceptable something that is less obvious, namely that this wish can to some extent be fulfilled: It is possible, in a certain manner and to a certain extent, to 'undo' actions that have already been done.

First a terminological matter. Regret is sometimes distinguished from guilt, remorse, and repentance – all of them moral remainders. Without wishing to deny that the distinctions are valid in certain contexts, I shall not make them here: As will appear in the sequel, the form of regret that I discuss will lie very close to guilt, remorse and repentance.<sup>1</sup>

---

<sup>1</sup> However, I do assume that regret often differs relevantly from two other moral remainders, namely shame and disappointment, although this assumption plays hardly any

## 2. Regret about what you yourself do

Regret is a relation-concept: You can only have regret for or about something. Linguistic usage is tolerant concerning the question what this 'something' might be; in particular it does not stipulate whether this something happens to you or whether you caused it yourself. If you say 'I am sorry that the downpour flooded your cellar', you express your regret about an incident in which you had no part at all: You think it is a pity that the catastrophe occurred, but you could not do anything about it, you are free from blame. The matter is different when you yourself bring about an event. Then regret is transformed into remorse, guilt, and perhaps repentance. Here it is that the phenomenon acquires that typically tragic force that is so often described in certain psychological novels. For whoever has remorse about what he himself has done is, as it were, doubly the victim: by what happened and by the fact that he caused it to happen.

Lexical dictionaries take this difference clearly into account. In the Oxford English Dictionary, for example, regret is described as a sort of sorrow that can take on one of two forms. The first is 'sorrow or disappointment due to some external circumstance or event', the second is 'sorrow or pain due to reflection on something one has done or left undone'. In this article I shall limit myself to this second kind of regret. Regret about 'some external circumstance or event' will be discussed only in a derivative sense, and the same applies to a kind of regret that one could see as an intermediate form between the two already mentioned. This is the remorse or guilt you feel, not for a natural disaster, nor for your own behaviour, but for the behaviour of others. Your compatriots who voted massively for Jörg Haider, for Jean-Marie le Pen, or for Pim Fortuyn, your director who cooked the books, your grandfather who was a Nazi sympathizer during the war: All these matters can give you a feeling of guilt and shame, even if they happened long before you were born. There are even people who feel guilty because they are part of the fallen human race. In the Calvinistic faith, with its notion of original sin, such collective guilt might perhaps be deemed fruitful; nonetheless, as I said, I shall leave this matter aside.

As with other forms of sorrow, regret implies the wish that reality had been different from what it actually was: Those who live in the world of their dreams have nothing to wish for and lack all occasion to be sorrowful. What distinguishes regret from other types of sorrow is the idea that reality easily *could have been different*. As has often been observed, regret

---

consequential role in my argument. For the differences between regret and shame, see Benedict, 1934; Piers and Singer, 1953; Rawls, 1971, Chapter 8. For the distinction between regret and disappointment, see Zeelenberg, 1996, Chapters 2 and 7, and Gerritsen, 1998.

is characterized by 'what might have been'. 'The constitutive thought of regret in general is something like 'how much better if it had been otherwise', claims Williams, 'and the feeling can in principle apply to anything of which one can form a conception of how it might have been otherwise' (1976: 27). The more one feels that the world could easily have been different and better, the greater the regret. But also: The more one feels that it is one's own act that could have been different and better (and not just that of your grandfather, your compatriots, or Mother Nature), the more the regret hurts and gnaws. Frijda (1986: 201) speaks in this context of 'painful self-evaluation due to some action evaluated negatively and for which the person holds himself responsible'. The consequences of this self-evaluation are well-known: 'mere suffering, wringing one's hands, and beating one's head' (201).

How to evade such painful feelings? You can try to repress and not feel them, but such an attempt is merely palliative. There is only one way in which regret for your own action can be combated satisfactorily, and that is to undo the action.<sup>2</sup> But precisely that seems impossible. Apart from a few trivial cases (I steal 100 euros from your piggy bank and, overcome by guilt, replace them before you notice the theft) what is done is done.

Is it? A negative answer seems bizarre. How could you arrange for an action that has already been performed not to have occurred? To claim that what is done can be undone seems as absurd as to say that what has happened can 'unhappen'. The reason for the absurdity is clear: The cancellation of executed actions or definite occurrences implies that we can change the past. That would mean that we could reverse the temporal order of the cause-effect relation, with the consequence that we might well become ensnared in puzzles like the grandmother paradox, for if I travel back in time to kill my grandmother before she has even met my grandfather, there can be no 'I' who travels back in time.

However, recent studies in physics and philosophy of physics teach us that time-reversal and retrocausality are less absurd than they seem at first sight: There appear to be several ways to resolve the grandmother paradox. I propose to say a few words about these ways, as well as about the objections raised against them, in Section 2. In Section 3 I shall explain in which sense we can, and in which sense we cannot be said to determine our own past.

---

<sup>2</sup> Frijda mentions another way, namely suicide: 'Since the self has caused the pain, only removal of the self can end the pain'. This option is indeed effective, if a trifle severe.

### 3. Retrocausality and the reversed arrow of time

Recent research on retrocausality and the arrow of time claims that it may be *physically*, and not merely logically, possible to affect the past (Horwich 1987, Savitt 1995, Price 1996). The physical possibility stems from the fact that physical laws are fully time-symmetric when applied to a small number of objects (a few elementary particles, several atoms, molecules, billiard balls, planets, and so on).<sup>3</sup> It is in fact only when there are very many objects, as in thermodynamics, that processes appear to be irreversible; and then we must adopt a time-asymmetric description of the observed phenomena. But great numbers of objects are made up of many sets of smaller numbers of objects, so the question naturally poses itself as to whence the time asymmetry has come. How can the irreversibility of thermodynamic and other processes with many degrees of freedom come into being, given that processes with few degrees of freedom are reversible? Similarly, if the cause-effect relation is also time-symmetric, then *why* should we not be able to influence the past?

As early as in the nineteenth century, the physicist Ludwig Boltzmann tried to explain thermodynamic irreversibility by means of what has been called *PI*<sup>3</sup>: the *Principle of Independence of Incoming Influences*. The idea is that very many objects together (in Boltzmann's case, molecules) give rise to very many interactions (collisions), which lead to dependencies or correlations that were not there before (Boltzmann 1877). So the suggestion is that the existence of very many objects increases the chance of collisions, and this in turn increases the probability that dependencies or correlations come about.

Boltzmann's idea was however severely and rightly criticized on the grounds that it *stipulates* rather than explains: Boltzmann merely *asserts* the time asymmetry by claiming that the correlations exist after, but not before the interactions. Thus he introduces a distinction between the time before and the time after (namely before and after a collision), whereas time asymmetry was what had to be explained.

Physicists after Boltzmann, notably Wheeler and Feynman in the twentieth century, tried to explain time asymmetry in another way (Wheeler and Feynman 1945, Wheeler and Feynman 1949). But in their attempt,

---

<sup>3</sup> Strictly speaking, since *certain* so-called weak interactions between elementary particles *do* violate *T*-invariance (i.e. do violate time reversal), but not *PCT*-invariance (simultaneous reversal of parity, *P*, charge conjugation, *C*, and time, *T*), one should rather say that all physical laws are *PCT*-symmetric. This means that when one reverses the direction of time (*T*), one should interchange left and right (*P*), and replace particles by antiparticles (*C*). This technicality has however no consequence for the burden of my argument: whenever I write time-symmetric, one should understand *PCT*-symmetric, in those cases where it makes a difference.

too, the very time asymmetry seems to enter the story as a *deus ex machina*. Hence both Boltzmann as well as Wheeler and Feynmann seem to be guilty of arguing in a circle: the very time asymmetry that is the explanandum, they use as an explanans.

Boltzmann, as well as Wheeler and Feynman, all believe that claims about altering the past are ultimately incoherent. However, Michael Dummett asserts the opposite. In his article 'Bringing About the Past', Dummett argues such claims can be coherent, on condition that, at the moment the attempt is made, it is impossible to ascertain what exactly *has* happened in the past. One of Dummett's two illustrations concerns a tribe that has the following custom:

Every second year the young men of the tribe are sent, as part of their initiation ritual, on a lion hunt: They have to prove their manhood. They travel for two days, hunt lions for two days, and spend two days on the return journey; ... While the young men are away from the village, the chief performs ceremonies – dances let us say – intended to cause the young men to act bravely. We notice that he continues to perform these dances for the whole six days that the party is away, that is to say, for two days during which the events that the dancing is supposed to influence have already taken place (Dummett 1964: 348).

Dummett comments:

There is generally thought to be a *special* absurdity in the idea of affecting the past, much greater than the absurdity of believing that the performance of a dance can influence the behaviour of a man two days' journey away ... (ibid., 349, Dummett's italics).

In other words, for most people retrocausality is absurder than action at a distance. Mistakenly, in Dummett's view. On condition that one cannot possibly find out, before the occurrence of the later cause, whether the earlier effect has taken place, a belief in retrocausality is perfectly coherent. Of course this belief can be false, but that is not the point; the point is that it does not lead to contradictions.

Dummett's argument is not the only way to defend the coherence of retrocausality, but it is, as Huw Price has remarked, 'by far the strongest defense' (Price 1996: 174). However, Price has also mounted a telling attack on Dummett. He observes that retrocausality is often understood in analogy with causality: The reverse arrow of time is like the forward arrow, only the direction is different. In the normal causal situation, in which the past is fixed but the future is not, we change that future by intervening (to greater or lesser degree). In the retrocausal situation, the future should be considered fixed but the past not, and thus our interventions change the past. It is precisely this last feature, according to Price, that is not properly

accounted for by Dummett. Dummett's argument implies that something *did* happen in the past, and on condition that we do not know what it is, we are able to exert some influence. Obviously this argument is at odds with the analogy between causality and retrocausality, which entails that as yet *nothing* happened in the past.<sup>4</sup>

Like Dummett, I maintain that there is nothing incoherent in the idea that you can influence the past. In a certain sense I shall even go a step further: I shall claim that there is a sense in which you really can determine the character of an act that you have already performed. As we will see, this sense reveals itself when we combine a first person perspective with a focus on actions as symptoms of tendencies. But we will also see that such an approach differs from that of Dummett, and one of the advantages of this is that it is immune to Price's criticism.

#### 4. The first person perspective and actions as symptoms of tendencies

Imagine that I am a member of a football club which is continually in straitened financial circumstances, and meetings are scheduled every Saturday to discuss money matters. I enjoy football but I hate administration and very often I invent an excuse to shirk these tasks. During one of the weekly meetings at which I have once more absented myself, it is decided to revoke my membership, since it seems all too obvious that I seek only the pleasures and will not shoulder the responsibilities of a club member. Now there are at least three reasons why I can regret my absenteeism. I can regret it because it has led to my expulsion, or because it is inconsistent with my duty as a member of the club, or because it is symptomatic of one of my bad habits, namely the tendency to let other people do my chores. The first form of regret is that of the utilist, the second relates to duty ethics and the third to virtue ethics. Of course the one form does not exclude the others. Mixed reasons are possible and even common: An unmixed utilist reasoning is as rare as are pure duty or virtue motivations. What I want to argue here is that feelings of regret of the third type, namely those pertaining to considerations of virtue, can in some cases be reduced. That is to say, *insofar as* an agent regrets her own action as a potential symptom of a fault or other objectionable tendency, it can sometimes be 'undone' in a nontrivial sense. This can be further explained by taking a closer look at the semantics of tendencies.

---

<sup>4</sup> Both the classical scenario that the past is fixed and the future is malleable, as well as its inverse, according to which the future is fixed but the past is changeable, are but extreme standpoints. A more flexible and defensible position is that *some* future and *some* past events can be influenced by present intervention.

Often (albeit not always) a single act is insufficient to allow one to claim that a particular tendency is present. For a tendency is normally manifested in several actions. In this connection, Ryle (1949: 43–44) spoke of ‘higher-grade dispositions’, Hempel (1965: 472) of ‘broadly-based dispositional traits’, and Carnap (1936–1937:444 ff; 1956: 64) of ‘multiple dispositions’. Carnap particularly exerted himself to provide a good definition of these multiple dispositions.<sup>5</sup>

The typical characteristic of multiple dispositions is their multifarious manifestation, depending on the situation. To take a familiar example: The magnetization of an iron bar can be manifested by the fact that iron filings attach themselves to the bar’s ends, but also by the fact that one end attracts the north pole, the other the south pole of a compass needle, and equally by the fact that, if the bar breaks in two, both halves retain the same properties. In Carnap’s analysis, each of these manifestations is described in a so-called reduction sentence. A reduction sentence has one of the following two forms:

$$\forall x(Mx \rightarrow (Sx \rightarrow Rx)) \quad (i)$$

or

$$\forall x(Sx \rightarrow (Rx \rightarrow Mx)) \quad (ii)$$

In words, (i) reads, for all objects  $x$  (where  $x$  is for example a bar), if  $x$  has the disposition  $M$  (here: is magnetic), then it is the case that, if  $x$  is in situation  $S$  (e.g. iron filings are placed closed to  $x$ ’s ends), then  $x$  will display reaction  $R$  (e.g. attract the filings). Formula (ii) reads, for all objects  $x$ , if  $x$  is in situation  $S$ , then it is the case that, if  $x$  displays reaction  $R$ ,  $x$  has disposition  $M$ . There are well-known difficulties connected with this procedure, arising from the fact that Carnap interprets the symbol ‘ $\rightarrow$ ’ as a material implication. In principle, however, it is possible to model reduction sentences in another way, for example as indicative or as counterfactual implications, whereby the most intransigent problems are avoided.<sup>6</sup>

My concern here is not the precise form of a reduction sentence, but

---

<sup>5</sup> Attempts, such as those of Carnap, to *describe* dispositions should not be confused with attempts to answer the question whether dispositions must have bases or grounds. The latter concern the *realization* of dispositions in *properties*, categorical or not, the former the *manifestation* of dispositions in *observable behaviour*. The two things are clearly different: a disposition can be realized in several different ways and manifest itself in one way only, or it may instead manifest itself in different ways, while having but one realization.

<sup>6</sup> For the convenience of readers not familiar with the symbolism used here, or with the problems of material implication, I have added an Appendix in which these matters are explained.

rather the fact that each manifestation of a disposition must correspond to a reduction sentence. For this means that every multiple disposition is described by a *sequence* of reduction sentences, and also that different sequences can partially overlap. If the dispositions are behavioural ones, the latter property allows us, by acting in a particular way, to continue an arbitrary sequence in a specified manner, with the consequence that earlier reduction sentences become part of one, and not another sequence.

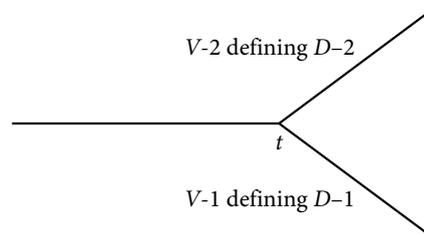
The idea can be illustrated clearly by means of behavioural dispositions that we experience on the one hand as being very different, but on the other hand as having considerable overlap as far as their manifestations are concerned. Bravery and recklessness, for example. If I am brave, I will react under certain circumstances in a particular way, but equally so if I am reckless. Often it is not possible to ascertain if I am brave or reckless on the basis of just one act, and even a number of acts may be insufficient, the problem being that the manifestations of these two dispositions resemble one another so greatly. Just as people in the Middle Ages had difficulty in deciding whether someone was possessed by the devil or had been touched by God (for the behaviour in the two cases often was thought to be largely the same), so we often cannot tell from a single action if someone is brave or merely reckless, miserly or simply prudent, self-conscious or only vain. The lesson is not that we ascribe dispositions to someone on grounds other than her verbal or nonverbal behaviour, but simply that we must observe *more* of her behaviour before we can conclude precisely which dispositions she has. Sometimes only by observation of someone's later behaviour can we determine the tendency that was expressed in her earlier deeds. But also – and here the alliance with the first person perspective is made – sometimes we ourselves have the opportunity, through our future deeds, to determine what precisely was the character of our earlier actions. The reason we may have this opportunity is that many actions do not express a tendency aside from the set of which they are an element, and moreover we can take advantage of this from a first person perspective. Of course, all this assumes that someone's life shows a certain continuity, but such an assumption is surely unproblematical; it is rather a life history lacking all continuity that causes problems.<sup>7</sup>

Let me try to make idea somewhat more precise, keeping within the framework of Carnap, and not troubling myself with the problems connected to the use of material implications. Take two multiple dispositions,

---

<sup>7</sup> I do not deny that now and then, in a person's life, a discontinuity can occur. As with Saul/Paul, some people experience radical turnings, where there seems to be little connection between what they did in the past and what they are doing now. In the same manner, I do not deny that there are situations in which one act suffices to call a person brave or reckless. These are just not the situations that I am talking about.

$D-1$  and  $D-2$ . Each will be described in a sequence of reduction sentences such as (i) and (ii) with  $D-1$  and  $D-2$  as substitutions for  $M$ . The two sequences, let us call them  $V-1$  and  $V-2$ , are typically open, this meaning for the early Carnap that they can be extended by the adjunction of new reduction sentences (Carnap 1936–37: 449), and for the later Carnap that neither sequence is long enough to give a complete description of  $M$  (Carnap 1956: 67). In the first case the meaning of  $D-1$  and  $D-2$  is determined entirely by the reduction sentences, in the second case it depends partly on the theory we have about  $D-1$  and  $D-2$ . For our purposes it does not matter which of the two interpretations we choose. It is however important that  $V-1$  and  $V-2$  have an overlap, and thus that they have a number of reduction sentences in common. This overlap does not need to be as large as it is in the case of bravery and recklessness, but it should be sufficiently great that, at a given time, it is plausible that either  $D-1$  or  $D-2$  might be applicable to a given person. If  $t$  is the latest time for which this plausibility holds, then  $t$  is said to be a bifurcation point between two different extrapolation possibilities. This could be the case if the agent himself tries to continue the sequence of acts in such a way that it resembles  $V-1$  more than  $V-2$ , so that finally  $D-1$ , and not  $D-2$ , is applicable to the agent. In this way the agent determines, through his actions after  $t$ , to which sequence the subsequence of actions before  $t$  belongs. In a diagram, the situation can be pictured thus:



It is important to understand this diagram correctly. It is not that an act can be described in different ways, so that an action performed before  $t$  acquires a new description after  $t$ . Neither is the point that often we only discover through later actions what the tendencies were that were expressed by the earlier acts. What is being contended has ontological implications, not merely semantic or epistemological ones. We are not just talking about a fixed set of actions ‘out there’ that can be described first such and then so, nor are we talking about a set out there with a determined character that is only revealed through later observations. What we are talking about concerns the adding, at an ontological level, of elements to a set, and hence the very creating of a set out there as we go along.

There are cases in which, by acting in a particular way after  $t$ , I can establish that my acts before  $t$  become a subset of another set of reduction sentences. In these cases I establish in retrospect which tendency is expressed by my earlier act, and I thus fix *post actu* the character of that act. Of course, it is always *possible* to take a semantical or epistemological view, just as it is always possible to be a hard nosed skeptic, or to maintain an animistic outlook, or to regard all events as consequences of divine intervention. Those positions are tenable, in the sense that they are not inconsistent. But the relevant question is whether they are fruitful, and there I have my doubts. I think that there are actions that are better accounted for by saying that their character is *determined* by later actions, than by saying that their character *acquires a new description later on* or by claiming that their character *is revealed or discovered through the observations of later actions*.<sup>8</sup>

This has consequences for our attitude towards regretting our own actions. The traditional, often religiously tinted reaction to such feelings of regret is to beg forgiveness, or alternatively to try to diminish your karma. In the first case you are dependent on somebody who imparts the desired forgiveness (whether it be God or another person), in the second case the idea is that your good deeds should outweigh your bad ones. Both cases are accompanied, if the agent is sincere, by the resolution 'never to do it again', and in both cases it is assumed that the character of the regretted deed remains more or less fixed. It is precisely this last assumption, which is also implicit in Williams (1976), that is undermined. In our approach, we can combat our regret in the most effective manner conceivable, namely by annihilating, in a certain sense, the deed that we regretted. The possibility is made feasible thanks to the combined outlook that we have

---

<sup>8</sup> Michael Smith, in private conversation, has objected to my view, stating that at any particular time a number of counterfactuals is true of the agent, and will forever be true. It will probably surprise Michael, as it did me, to learn that the controversy between us reflects a famous dispute between Jesuits and Dominicans, where he represents the former and I the latter. In the introduction to the second volume of her *Collected Papers*, G.E.M. Anscombe recalls: 'I read a work called *Natural Theology* by a nineteenth century Jesuit. I read it with great appetite and found it all convincing except for (...) the doctrine of *scientia media*, according to which God knew what anybody would have done if, e.g. he hadn't died when he did. This was part of theodicy, and was also the form in which the problem of counterfactuals was discussed. I found I could not believe this doctrine: it appeared to me that there was not, quite generally, any such thing as what would have happened if what did happen had not happened, and that in particular there was no such thing, generally speaking, as what someone would have done if ... and certainly that there was no such thing as how someone would have spent his life if he had not died as a child. I did not know at the time that the matter was one of fierce dispute between the Jesuits and the Dominicans, who took rather my own line about it' (Anscombe 1981, vii).

adopted, namely the first person perspective in conjunction with the view of actions as expressions of tendencies, where it is acknowledged that tendencies are often identified on the basis of a number of actions. Insofar as we are sorry for our actions as potential symptoms of undesirable tendencies, we are able to influence the character of these acts by means of our future behaviour.<sup>9</sup>

An example may be helpful. Imagine that I am a happily married mother of four small children. One day I go to a party where I have a great time, dancing exuberantly and accepting all the many glasses of wine offered to me. The following morning I find myself in a hotel room, lying next to a handsome man whom I do not recall having seen before. Then it seems all too clear what happened. Nevertheless, I still have some latitude in determining what I have already done. In particular I can make it the case, through my future actions, that this adventure becomes either a mere incident or the beginning of a long and secret affair. What is more, my knowledge of the fact that I have this possibility, and hence my understanding that I am situated at a bifurcation point, might motivate me to pursue the one rather than the other course. In this sense cognition can strengthen volition.

Note that the adoption of a first person perspective is essential here. A restless volcano might or might not erupt, but it cannot understand what is going on – let alone that it can be motivated through reflection on the situation to pursue the one course rather than the other. In addition, a certain indeterminacy is required with respect to the question which tendency exactly it is that is manifested.

An example based on musicological insights might make this even clearer. In Thomas Mann's novel *Doctor Faustus*, the main character is a German composer by the name of Adrian Leverkühn, whose tragic life story is told by his trusty friend Serenus Zeitblom. Zeitblom needs six chapters to relate Leverkühn's childhood and adolescence before Leverkühn himself, in Chapter 7, takes the stage. He explains to Zeitblom that the essence of music lies in relation (*Beziehung*) and in ambiguity (*Zweideutigkeit*). An arbitrary chord, for example the black keys F sharp, A sharp, C sharp, do not prescribe a definite key. It looks as though they are in F-sharp major, but if we add an E, the same chord belongs to B major, as its dominant. In this manner we imbue retroactively the notes that have already sounded with their key; thus we can shape, and in that sense change, the past. As Zeitblom narrates:

---

<sup>9</sup> In the Spring of 2007, when I was visiting the philosophers at the University of Miami, Michael Slote drew my attention to Chapter VII of his 1985 book, in which he discusses the relevance of future behaviour for past actions in the context of ethics.

To illustrate the meaning of the word [ambiguity – JP], he [Leverkühn] played me chord-progressions belonging to no definite key; demonstrated for me how such a progression fluctuates between C major and G major, if one leaves out the F, that in G major turns into F sharp; how it keeps the ear uncertain as to whether that progression is to be understood as belonging to C major or F major if one avoids the B, which in F major is flattened to B flat.

You know what I find?, he asked. That music turns the equivocal into a system. Take this or that note. You can understand it so or respectively so. You can think of it as sharpened or flattened, and you can, if you are clever, take advantage of the double sense as much as you like (Mann 1947: 49).

Leverkühn is a classical composer (Zeitblom compares his compositions to the works of Monteverdi, Mahler, Schönberg and Alban Berg), but the principle that he refers to above must sound familiar to contemporary improvising musicians. In an interview on Dutch television in January 2004, the jazz pianist Louis Van Dijk gave the following answer when asked why he preferred improvised music:

The nice thing about improvising is: You play a mistake, then you repeat that mistake eight times, and then not only have you legalised that mistake, but you even have elevated it to the point where it has become the starting point.

What Van Dijk describes here is that the nature of a particular note is often determined by the manner in which the musician proceeds. In this way, a note that was produced ‘by mistake’ or ‘accidentally’, can give rise to a new, interesting musical development (of course, the opposite can also happen: A sequence of challenging or promising notes might not be picked up by the fellow musicians so that it perishes without trace). In this sense, then, the later behaviour of a musician may determine the character of what he did earlier – just as, in our earlier example, my escapade last night becomes either a mere incident or the beginning of a long and secret affair, precisely because of my later actions. Rather than as a restriction, the indeterminacy involved in this approach functions as a liberating factor.

It will be clear that this manner of ‘bringing about the past’ is very different from that which Dummett describes in his article of the same name. Three differences are noteworthy. First, Dummett is interested solely in showing that the notion of retrocausality is not necessarily incoherent, but I myself have gone a step further, namely by showing that a particular form of retrocausality really takes place. Second, while Dummett means by the ‘past’ the actions of others, in this case those of the young men on the lion hunt, I am talking about a past consisting of someone’s own actions: As said before, the first person perspective is an indispensable part of what

I am trying to convey. The third difference pertains to the role of knowledge. In Dummett's account knowledge plays no role, in the sense that it must be fully absent. After all, Dummett requires that, before the occurrence of the later cause, it is impossible to know whether the earlier effect has taken place: The dances that are supposed to ensure the young men's bravery only make sense if the tribal chieftain cannot possibly know whether the men were indeed brave. In my approach, on the other hand, some knowledge of what has already occurred is necessary. I can only be motivated to continue the one rather than the other course if I have some idea about what I have done and about how that fits into possible networks of different dispositions. That is, I must know which terms 'S' and 'R' could be used as descriptions of my earlier actions, and I must also know which dispositions *M* could in principle correspond to *S* and *R*. In practice the number of possibilities will be limited by the normal scientific rules of simplicity, consistency, continuity, etc., but there will always be some leeway here. What we do when we review our past actions is to look for patterns that form networks of different dispositions. Since these networks are open and incomplete, we can, by acting in a particular way, continue and strengthen the one rather than the other pattern. In this sense we are, so to speak, weaving our own life by living it.

These differences all reflect the fact that we, unlike Dummett, have connected the phenomenon of 'bringing about the past' with occurrences of regret and remorse. This becomes especially clear in the third difference. Knowledge of what you have done is an inherent part of the sort of regret and remorse that we have been talking about. Without such knowledge, feelings of regret and remorse would not arise, and the concepts of regret and remorse would not make sense. A coherent notion of remorse includes knowledge of what one did, and in addition it opens up the possibility of bringing about one past rather than another, as Oscar Wilde apparently realised when he noted in *De profundis*<sup>10</sup>:

Of course the sinner must repent. But why? Simply because otherwise he would be unable to realize what he had done. The moment of repentance is the moment of initiation. More than that: It is the means by which one alters one's past. The Greeks thought that impossible. They often say in their Gnostic aphorisms, 'Even the Gods cannot alter the past'. Christ showed that the commonest sinner could do it, that it was the only thing he could do. Christ, had he been asked, would have said – I feel quite certain about it – that the moment the prodigal son fell on his knees and wept, he made his having wasted his substance with harlots, his swineherding and hungering for the husks they ate, beautiful and holy moments in his life. It is difficult for most people to grasp the

---

<sup>10</sup> Thanks to my colleague Allard Tamminga who brought this citation to my attention.

idea. I dare say one has to go to prison to understand it. If so, it may be worth while going to prison (Wilde 1949: 179).

Etymology teaches us that the words ‘regret’ and ‘remorse’ stem from the infinitive ‘to turn back’. The Hebrew root ‘shav’, the Arabic word ‘*tawbah*’, and the Greek ‘*metanoia*’ all denote not only regret, but also return. Theologians, irrespective of whether they are of Jewish, Moslem or Christian descent, have interpreted this as a return to God. But perhaps it is not too far-fetched to give it a secular twist and understand it as a way of going back to one’s past actions with the aim of determining them through future ones. In 1843 Søren Kierkegaard notes in his diary that philosophers rightly say ‘that life must be understood backwards’, but wrongly forget ‘that it must be lived forwards’ (SKJP 1 1030). He adds: ‘[L]ife in temporality never becomes properly understandable, simply because never at any time does one get perfect repose to take a stance: backwards’ (ibid.). However, the fact that we cannot find perfect repose to look back is only part of the reason why life in temporality is so difficult to understand. The other part is that what we aim to understand, our life, has not yet taken its full shape, and that we give it its full shape precisely by living it forwards.

## Appendix

The present appendix is offered for the convenience of readers who are not familiar with first order logic, or with the use that Carnap makes of it in defining dispositions in terms of reduction sentences.

In the most elementary part of formal logic, viz. propositional logic, the expression ‘if ... then ...’ is taken as a so-called material implication, represented by the symbol ‘ $\rightarrow$ ’. The meaning of the formula ‘ $p \rightarrow q$ ’ (where ‘ $p$ ’ and ‘ $q$ ’ are arbitrary propositions) is fixed by indicating under which circumstances exactly ‘ $p \rightarrow q$ ’ is true and under which circumstances it is false. According to propositional logic, ‘ $p \rightarrow q$ ’ is only false in case ‘ $p$ ’ is true and ‘ $q$ ’ is false; in the three remaining cases, ‘ $p \rightarrow q$ ’ is true.

This definition of ‘ $p \rightarrow q$ ’ corresponds reasonably well with our use of ‘if  $p$  then  $q$ ’. For example interpret ‘ $p$ ’ as ‘I am the director’ and ‘ $q$ ’ as ‘You receive a permanent position’. Then it is easy to see that there is only one situation in which ‘if  $p$  then  $q$ ’ is false, namely when you do not receive a permanent position whilst I am the director.

However, our use of ‘if  $p$  then  $q$ ’ is not entirely captured by the propositional definition of ‘ $p \rightarrow q$ ’. The reason for the mismatch is that ‘if  $p$  then  $q$ ’ typically presupposes that ‘ $p$ ’ and ‘ $q$ ’ are somehow semantically connected, while ‘ $p \rightarrow q$ ’ does not make this assumption. Let ‘ $p$ ’ for instance be ‘2 is an odd number’ and let ‘ $q$ ’ be ‘Amsterdam is the capital of the Netherlands’. Under this interpretation, ‘if  $p$  then  $q$ ’ is either false or senseless, since ‘ $p$ ’

(called the antecedent) and 'q' (the consequent) have nothing to do with one another. However, ' $p \rightarrow q$ ' is true under this interpretation, for 'p' is false and 'q' is true. The fact that a material implication is always true if it has a false antecedent or a true consequent, even if antecedent and consequent have no semantical connection whatsoever, has led one to speak of 'the paradoxes of material implication'. Strictly speaking, however, these facts are not real paradoxes, i.e. logical contradictions, but merely counterintuitive results of the definition of ' $p \rightarrow q$ '.

In the nineteen twenties Rudolf Carnap held the view that meaningful empirical concepts are always completely definable as observability concepts, i.e. concepts referring to properties, relations, states of affairs etc. that can be observed or otherwise detected by the senses. However, in the nineteen thirties he realised that dispositional concepts cannot be thus defined (Carnap 1936–1937). Consider the sentence 'Derek is allergic to adrenalin', in which a disposition ('being allergic to adrenalin') is meaningfully ascribed to Derek. Since this sentence can be true at a time when Derek looks a picture of health, it follows that its truth or falsity cannot simply be established by observing Derek at a particular moment. What we mean by saying that Derek suffers from adrenalin allergy is that under certain conditions, namely when Derek is injected with adrenalin, he will display certain reactions such as turning pale or fainting.

Let  $M$  be the disposition of being allergic to adrenalin, let  $S$  be the conditions under which  $M$  manifests itself and let  $R$  be the manifestation in question. Then it might seem reasonable to define  $M$  in the following way:

$$\forall x(Mx \leftrightarrow (Sx \rightarrow Rx)) \quad (\text{iii})$$

which means, for all objects  $x$  (where  $x$  might be a subject such as Derek),  $x$  has the disposition  $M$  (here: being allergic to adrenalin) if and only if it is the case that, if  $x$  is in situation  $S$  (e.g.  $x$  is being injected with adrenalin), then  $x$  will display reaction  $R$  (e.g. faint). The symbol ' $\leftrightarrow$ ' is a two-sided material implication, also known as material equivalence, corresponding to the expression 'if and only if'. The formula ' $p \leftrightarrow q$ ' is equivalent to the formula ' $(p \rightarrow q) \& (p \leftarrow q)$ ': Either both are true or both are false ( $p$  and  $q$  again are arbitrary propositions).

However, as Carnap noticed, the definition of  $M$  in terms of (iii) is problematic, for ' $Sx \rightarrow Rx$ ' is already true if ' $Sx$ ' is false (ibid.). In our example this would mean that never injecting a person with adrenalin is a sufficient reason for calling that person adrenalin allergic. Since that is clearly not what we mean by saying that someone is allergic to adrenalin, Carnap rejects (iii) as a definition of  $M$ .

Carnap's next step is to propose a definition of  $M$  in terms of *reduction*

sentences like:

$$\forall x(Mx \rightarrow (Sx \rightarrow Rx)) \quad (i)$$

or

$$\forall x(Sx \rightarrow (Rx \rightarrow Mx)) \quad (ii)$$

Applied to our adrenalin example, (i) means, for all  $x$ , if  $x$  is allergic to adrenalin, then if  $x$  is injected with adrenalin,  $x$  will faint. Formula (ii) reads, for all  $x$ , if  $x$  is injected with adrenalin, then if  $x$  faints,  $x$  is allergic to adrenalin. The objection raised against (iii) does not hold for (i) or (ii). For in neither (i) nor (ii) is the falsity of 'Sx' a sufficient reason for the truth of 'Mx'.

Formula (i) is said to express a *necessary* condition for the application of  $M$  in that the truth of 'Mx' and 'Sx' necessitates that of 'Rx'. Formula (ii) expresses a *sufficient* condition in that the truth of 'Sx' and 'Rx' is sufficient for the truth of 'Mx' (Carnap 1936-1937). Taken together, (i) and (ii) express a necessary and sufficient condition and hence can serve as a definition of  $M$ . The conjunction of (i) and (ii) is equivalent to  $\forall x(Sx \rightarrow (Rx \leftrightarrow Mx))$ .

### Acknowledgements

This article is an extended version of 'Is what is done done? On regret and remorse', *The Journal of Mind and Behavior* 26(4) 219–227 (2005). The central idea was first presented in March 2003 in Sydney, where Huw Price of the Centre for Time had convened a one-day conference around an early version of the paper. I thank the other contributors to the conference for their commentaries: Fred D'Agostino, Tim Bayne, Barbara Nunn, John Quilter, Natalie Stoljar. The paper was later presented in Groningen (Department of Theoretical Philosophy) and Canberra (Research School for the Social Sciences, ANU). I also used the idea in my opening lecture of the conference 'Knowledge and Imagination', June 2004, Amsterdam. Papers presented at this conference, including my lecture, were published in *Metaphilosophy* 37(4), 2006. I thank Manfred Holler for his invitation to write an extended version for *Homo Oeconomicus* and Matthew Braham for his instructive comments.

### References

- Anscombe, G.E.M. (1981), *Metaphysics and the Philosophy of Mind*, Volume Two of *The Collected Papers of G.E.M. Anscombe*, Blackwell.
- Benedict, R. (1934), *The Chrysanthemum and the Sword: Patterns of Japanese Cul-*

- ture, Houghton Mifflin.
- Betzler, M. (1999), Rational deliberations and reactive attitudes, in: J.L. Falguera, U. Rivas and J.M. Sagullo (eds.), *La Filosofía Analítica en el Cambio de Milenio (Analytic Philosophy at the Turn of the Millennium)*, Universidad de Santiago de Compostela Publicaciones.
- Bittner, R. (1992), Is it reasonable to regret things one did?, *Journal of Philosophy* 89: 262-273.
- Boltzmann, L. (1877), Ueber die Beziehung eines allgemeine mechanischen Satzes zum zweiten Hauptsatze der Wärmetheorie, *Sitzungsberichte, K. Akademie der Wissenschaften in Wien, Math.-Naturwiss.* 75: 67-73.
- Carnap, R. (1936-1937), Testability and meaning, *Philosophy of Science* 3: 420-472 and 4: 1-40.
- Carnap, R. (1956), The methodological character of theoretical concepts, in: H. Feigl and M. Scriven (eds.), *The Foundations of Science and the Concepts of Psychology and Psychoanalysis*, University of Minnesota Press.
- Dancy, J. (1993), *Moral Reasons*, Blackwell.
- Dummett, M.A.E. (1964), Bringing about the past, *Philosophical Review* 73: 338-59.
- Frijda, N.H. (1986), *The Emotions*, Cambridge University Press.
- Gerritsen, L. (1998), *If I Knew Then What I Know Now: Determinants of Regret and Disappointment*, Ph.D. thesis, Technical University of Eindhoven, The Netherlands.
- Hempel, C.G. (1965), *Aspects of Scientific Explanations and Other Essays in the Philosophy of Science*, The Free Press (Paperback edition 1970).
- Horwich, P. (1987), *Asymmetries in Time*, MIT Press.
- Hurka, T. (1996), Monism, pluralism and rational regret, *Ethics* 106: 555-575.
- Kierkegaard, S., Hong, H.V. and Hong, E.H. (1967-1975), *Søren Kierkegaard's Journals and Papers (SKJP)* (Vol. 1-6), Indiana University Press.
- Landman, J. (1993), *Regret: The Persistence of the Possible*, Oxford University Press.
- Loomes, G. and Sugden, R. (1982), Regret theory: An alternative theory of rational choice under uncertainty, *The Economic Journal* 92: 805-824.
- Mann, T. (1947), *Doctor Faustus: The Life of the German Composer Adrian Leverkühn as Told by a Friend (Doktor Faustus: Das Leben des deutschen Tonsetzers Adrian Leverkühn erzählt von einem Freunde)*, translated by H.T. Lowe-Porter, 1970, Penguin.
- Piers, G. and Singer, M.B. (1953), *Shame and Guilt: A Psychoanalytic and a Cultural Study*, Charles C. Thomas.
- Price, H. (1996), *Time's Arrow and Archimedes' Point*, Oxford University Press.
- Rawls, J. (1971), *A Theory of Justice*, Harvard University Press.
- Ryle, G. (1949), *The Concept of Mind*, Hutchinson.
- Savage, L. (1951), The theory of statistical decision, *Journal of the American Statistical Association* 46: 55-67.
- Savitt, S. (ed.) (1995), *Time's Arrow Today*, Cambridge University Press.
- Slote, M. (1985), *Common-Sense Morality and Consequentialism*, Routledge and Kegan Paul.
- Vollmer, P.J. (1977), Do Mischievous Dogs Reveal their 'Guilt'?, *Veterinary Medicine Small Animal Clinician* 72: 1002-1005.
- Waal, F. de (1996), *Good Natured*, Harvard University Press.

- Wheeler, J.A. and Feynman, R.P. (1945), Interaction with the absorber as the mechanism of radiation, *Reviews of Modern Physics* 17: 157-81.
- Wheeler, J.A. and Feynman, R.P. (1949), Classical electrodynamics in terms of direct interparticle action, *Reviews of Modern Physics* 21: 425-33.
- Wilde, O. (1949), *De Profundis* (reprinted in *De Profundis and other Writings* with an introduction by Hesketh Pearson, Penguin, 1976).
- Williams, B. (1976), *Moral Luck* (reprinted in *Moral Luck, Philosophical Papers* 1973-1980 (20-39), Cambridge University Press.
- Zeelenberg, M. (1996), *On the Importance of What Might Have Been: Psychological Perspectives on Regret and Decision Making*, Ph.D. thesis, University of Amsterdam.
- Zeelenberg, M. (1999), The use of crying over spilled milk, *Philosophical Psychology* 12: 325-340.