Logique & Analyse **143-144** (1993), 359-391

ADAPTIVE LOGIC IN SCIENTIFIC DISCOVERY: THE CASE OF CLAUSIUS*

Joke MEHEUS**

1. *Aim and survey*

It is the aim of this paper to review some logical systems with respect to their adequacy in a particular type of creative processes, namely those that involve inconsistent constraints. I shall show that these creative processes have some very interesting properties: they require *reasoning* (from inconsistent premises) in order to detect *specific* contradictions and to resolve them. I shall argue in detail that neither classical logic nor *mixed* non-monotonic logics[1], nor Rescher-like mechanisms, nor monotonic paraconsistent logics are suitable to understand this kind of processes. We need a logic that enables us to reason sensibly in the presence of (explicit and implicit) inconsistencies and that nevertheless is almost as powerful as classical logic, in other words an inconsistency-adaptive logic.

The logical systems will be reviewed with respect to a particular episode in the history of nineteenth century thermodynamics. As is generally agreed upon, modern thermodynamics was founded by Rudolf Clausius, who 'reconciled' two *incompatible* approaches to the phenomena of heat and work: the theory of Sadi Carnot on the one hand, and some ideas (and experimental results) of James Prescott Joule on the other. Several authors describe Clausius's contribution to thermodynamics as the result of a simple standard procedure: eliminate some parts of Carnot's theory in such a way

* I am indebted to Yorgos Goudaroulis for interesting discussions on nineteenth-century thermodynamics, to Jean Paul Van Bendegem for his enthusiastic help with some mathematical passages in Clausius's work, to Erik Weber for some useful comments on an earlier version of this paper, and especially to Diderik Batens for the many enlightening discussions on the application of inconsistency-adaptive logics and for his (as always) sharp and challenging comments.

** Aspirant navorser van het Belgisch Nationaal Fonds voor Wetenschappelijk Onderzoek.

1 This term refers to non-monotonic logics as commonly understood. They are called *mixed* because, as Batens shows in his 199+a (this volume), a deductive and a preferential component are *blended* in them. (The first component may lead to inconsistencies whereas the second component selects the correct half from each inconsistency.)

that the result is compatible with Joule's ideas. As I show in my 199+ (the companion to the present paper), this description is mistaken. Clausius arrived at his theory through a complex creative process, involving several inconsistencies the elimination of which was far from obvious.

In the companion paper, I present an analysis of the problem as it manifested itself to Clausius as well as a detailed reconstruction (based on his *Ueber die bewegende Kraft der Wärme und die Gesetze, welche sich daraus für die Wärmelehre selbst ableiten lassen* of 1850) of the process by which he arrived at his theory. This reconstruction is *not* a 'rational reconstruction' in the sense of Lakatos (a view on the way in which Clausius *should* have made his discovery), but an attempt to reconstruct, on the basis of an analysis of the original texts, as accurately as possible Clausius's reasoning process.

In the present paper, I proceed as follows. After some preliminary remarks (§2), I introduce those elements from early thermodynamics that are needed to understand the arguments in later sections (§3). In §4, I give a brief summary of the most important findings of my 199+. Relying upon these, I discuss the central mechanisms behind Clausius's reasoning process (§5). Next, I approach the question by which logic it is possible to make sense of Clausius's reasoning process (§6) and illustrate the answer to this question with a detailed discussion of two stages of Clausius's reasoning process (§7).

2. *Some preliminary remarks*

As I already mentioned, the results of the present paper are based on a detailed reconstruction of Clausius's discovery of his 1850 theory. In this reconstruction Clausius's discovery is viewed as a *problem solving process*. According to present-day methodologies of discovery, a problem consists of two components: a goal and a set of constraints — items of information that are relevant for the solution of the problem (Nickles 1980, 1981). In my terminology, the constraints may be of two kinds. *Limiting constraints* function as *conditions* on the solution (for instance, the requirement that the solution be consistent). *Constructive constraints* function as *premises* from which the solution can be *derived*. In accordance with this terminology, a problem is considered to be *well-defined* if, without further search, a solution satisfying the limiting constraints can be derived from the set of con structive constraints.[2]

Given the complexity of the problem, it is highly implausible that Clausius reasoned exactly in the way that it is represented in my 199+. However, the reconstruction reveals several interesting mechanisms (see section 5). Attempts to make the reconstruction more realistic will not undermine those mechanisms. Such attempts will only result in the discovery of more instances were Clausius detected new inconsistencies, more instances were he analyzed and reorganized the relevant constraints, and more instances were he fixed preferences in the light of a logical analysis of the set of constraints.

For the sake of clarity and brevity, I introduce some terms which do not occur in the original texts. I have taken care, however, that Carnot's, Joule's and Clausius's ideas are reproduced as accurately as possible.

3. *A sketch of early thermodynamics*

In section 3.1, I present a brief introduction to the central assumptions of Carnot's theory. Joule's ideas are briefly discussed in section 3.2. The central elements of both Carnot's theory and Joule's ideas are listed for further reference — where necessary the interpretation of terms and statements is rendered between square brackets. Section 3.3 contains some comments on the incompatibility between the two approaches.[3]

3.1. *Carnot's theory and the fall of caloric*

Central to Carnot's theory is the assumption that heat is a material substance [caloric] that can neither be created nor destroyed. In line with this assumption, the working of a heat engine was conceived by Carnot on analogy with the working of a water wheel: a heat engine receives heat from a hot body which it delivers to a colder body in the same way as a water wheel receives water from a high level which it emits at a lower

[2] This does not entail that a problem defined by an *inconsistent* set of constructive constraints is necessarily well-defined. The limiting constraints of most (if not all) problems include the (implicit) requirement that the solution should be *unique*. As *any* solution follows from an *inconsistent* set of constructive constraints (at least according to classical logic), one of the limiting constraints is not satisfied.

[3] For a detailed and accurate discussion of early thermodynamics, see Truesdell 1980.

level; in both cases, the production of work results from the 'fall' of the working substance. Hence, Carnot accepted (pp. 9-11, pp. 6-7)[4]

C1 The production of work by the agency of heat is always accompanied by the transfer of heat from a hotter to a colder body [fall of caloric]; the production of work by a heat engine is *not* due to the consumption of heat but results from the *mere* transfer of heat from a hot to a cold reservoir.

According to Carnot's theory, every difference in temperature can be exploited to produce work. However, when bodies of different temperature are in *immediate* contact with each other, there is a flow of heat which is *not* accompanied by the production of work. Consequently, Carnot accepted (p. 23, p. 13)

C2 Some processes result in a 'loss' of (potential) work.

In order to analyze the relation between heat and work, Carnot devised a *reversible* cycle in which an engine, after a series of operations between a hot and a cold reservoir, returns to its original state. The action of this engine (henceforth, a "Carnot-engine") is given by the following assumptions (pp. 36-37, p. 19):

C3 Operating a Carnot-engine in the normal direction results in the production of work and the transfer of heat [caloric] from the hot to the cold reservoir; operating a Carnot-engine in the reversed direction results in the consumption of work and the transfer of heat [caloric] from the cold to the hot reservoir [during these transfers heat is neither consumed nor produced].

C4 The amount of heat [caloric] absorbed by a Carnot-engine at one reservoir equals the amount of heat [caloric] delivered at the other.

For a good understanding of Carnot's interpretation of the Carnot-engine, also the following should be noted. (i) Both directions of a Carnot-engine annul each other: if the same amount of heat is transferred, the amount of work produced in the normal direction equals the amount of work consumed in the reversed direction. (ii) The amount of work a Carnot-engine

[4] The references to Carnot's paper are given by two numbers referring to the relevant pages in the original French text (Carnot 1824) and in the standard English translation (Mendoza 1960).

produces with a given amount of heat is a function of the temperature of the reservoirs between which the heat is transferred. (iii) The operations of a Carnot-engine are arranged in such a way that bodies of different temperature are never in contact with each other. Hence, a Carnot-engine is an *ideal* engine: the amount of work produced by it is a theoretical *maximum* (no work is lost during the operations).

Relying upon these ideas, Carnot derived the following theorem:

C5 It is impossible to design an engine that produces *more* work than a Carnot-engine, while operating between the same temperatures and absorbing the same amount of heat (from a hot reservoir).

Henceforth, such an engine is called a "more efficient engine". The argument that led to C5 will play an important role in section 7.1. Let me therefore give a brief reconstruction of it.

Suppose, Carnot argued, that C5 is false, and hence that it is possible to design a more efficient engine. Clearly, a *portion* of the work produced by this engine (in transferring an amount of heat from the hot to the cold reservoir) would suffice to operate a Carnot-engine in the inverse sense (thereby returning the *same* amount of heat from the cold to the hot reservoir). Hence, the combined operation of both engines would result in a perpetual motion machine. As Carnot considered the latter to be unacceptable, he concluded, by *reductio ad absurdum*, to C5 (pp. 20-21, pp. 11-12).

3.2. *Joule's ideas and the mutual conversion of heat and work*

Joule's first research concerned the design of electric motors. Early in his investigation he noticed the heating effect of a current (generated by some kind of dynamo). This discovery led to the conviction that work may be *converted* into heat, and hence, that the production of heat results from the *consumption* of work. Joule also accepted the converse, namely that heat may be converted into work, and hence, that the production of work (by the agency of heat) results from the consumption of heat.

It is important to note that Joule viewed these conversions as mechanical processes which are perfectly *reversible*. From Joule's point of view, the work consumed in producing heat (as an example one might think of heating water by stirring it with a paddle wheel) can be entirely recovered — it suffices to convert the heat produced back into work. Similarly, for the heat consumed in producing work. These ideas are summarised in the following assumption:

J1 Work and heat are *convertible* into each other: not only is the production of work by a heat engine due to the consumption of an equivalent amount of heat (and the production of heat due to the consumption of an equivalent amount of work), the heat (work) consumed during these conversions is entirely recoverable.

In view of J1, the following idea of Joule's seems self-evident:

J2 The total amount of work in the universe is *conserved* [no transformation can result in the loss of (potential) work].[5]

These ideas led to a series of remarkable experiments for the measurement of the 'mechanical equivalent of heat' (the amount of work that *corresponds* to one unit of heat). According to Joule's own account, the results were well in line with his ideas.

3.3. *The incompatibility between Carnot's theory and Joule's ideas*

At the beginning of the 1840's, Carnot's theory was generally accepted (notwithstanding the fact that heat was no longer conceived as a material substance but as a wave).[6] By that time, Joule's experimental results and ideas found entrance; the latter were soon accepted as fundamental principles. However, from the joint adoption of Carnot's theory and Joule's ideas *at least* two contradictions result. As the reader may verify, a first contradiction results from C1&J1 and a second one results from C2&J2.

It was generally agreed upon that these contradictions caused serious difficulties. This should not surprise us: Carnot's theory was commonly considered as the *only* theory available for thermodynamic phenomena, and Joule's ideas were generally accepted as fundamental principles. Moreover, resolving the inconsistencies that result from the joint adoption of Carnot's theory and Joule's ideas (in a satisfactory way) was *far from trivial*. Kelvin was even convinced that, if one accepted Joule's ideas, thermodynamics had to be restarted *from scratch*. This view was not due to lack of competence — Kelvin was one of the most competent researchers in the domain of thermodynamics of those days. It was related to the fact that Carnot's

[5] This principle should not be confused with the present-day principle of the conservation of *energy*. The (modern) concept of energy originated only in the late nineteenth century. (For a discussion of the development of the concept of energy, see Smith 1990.)

[6] For a discussion of the wave theory of heat, see Brush 1986.

theory (when combined with Joule's ideas) leads to contradictions, the elimination of which is not obvious.

The reactions to the situation were diverse. Put roughly: Joule rejected Carnot's theory, Kelvin (eventually) rejected Joule's ideas (in order to avoid 'innumerable other difficulties'), and Clausius, who favoured a mechanical model of heat, attempted to reconcile Carnot's theory with Joule's ideas. (According to his model, heat is a kind of motion and thus subject to the laws of mechanics. The mutual conversion of heat and work is well in line with this.) This attempt resulted in the first theory of modern thermodynamics.

It should be noted that also Holtzmann attempted to reconcile Carnot's theory with Joule's ideas. In a memoir published in 1845, he explicitly rejected C1 in favour of J1. But, as Clausius noticed (pp. 5-6, pp. 110-111)[7], the resulting theory is still inconsistent, for C1 is tacitly assumed in the equations from which Holtzmann derived his conclusions.[8]

4. *A reconstruction of the discovery of the first theory of modern thermodynamics*

In section 4.1, I briefly discuss the (components of the) problem that gave rise to Clausius's 1850 theory. A sketchy survey of the central stages in Clausius's reasoning process can be found in 4.2. Due to lack of space, the survey is unavoidably schematic and simplified —for a more detailed and more convincing account, I refer to my 199+.

4.1. *Clausius's problem: its goal and its constraints*

There can be no doubt about the *goal* Clausius was aiming at. Confronted with the incompatibility between Carnot's theory and Joule's ideas, Clausius wanted to develop a theory that would satisfy the following *limiting constraints*: (i) it would be at least as powerful as Carnot's theory for explaining thermodynamic phenomena, (ii) it would be compatible with his mechanical model of heat, and (iii) it would be consistent. The *constructive*

7 Where the references to Clausius's 1850 paper are given by two numbers, the first refers to the relevant pages in the original German text (Clausius 1850), the second to the relevant pages in the standard English translation (Mendoza 1960).

8 For a discussion of Holtzmann's contribution to thermodynamics, see Truesdell 1980, pp. 158-61.

constraints of Clausius's problem included the following: (i) his mechanical model of heat, (ii) some relevant experimental results, (iii) Joule's ideas and (iv) Carnot's theory.[9] As Clausius, like most of his contemporaries, found it inconceivable that there could be a loss of work in Nature, it is most likely that he originally aimed at a theory from which C2 would *not* follow. (But, as we shall see, the final solution turned out to be quite different.)

Someone might object to including Carnot's theory among the constructive constraints, for it is clearly inconsistent with Clausius's mechanical model of heat as well as with Joule's ideas. However, there is no alternative. The reason for this is not difficult to understand. Carnot's theory offers a detailed understanding of thermodynamic phenomena, but it is incompatible with Clausius's mechanical model of heat. Joule's ideas seem well in line with this model, but they are highly incomplete. Hardly anything interesting follows from them for the study of specific phenomena. The upshot is that Clausius, in order to arrive at a *consistent* and *complete* theory, had to rely not only on Joule's ideas but also on Carnot's theory.

Still, one might object as follows. Even if both Carnot's theory and Joule's ideas are needed, it must have been clear to Clausius that Carnot's theory is in need of modification and not, for instance, Joule's ideas. Hence, it makes sense to assume that Clausius assigned a higher *preference* to the latter than to the former. This would then result in the strategy to resolve inconsistencies by simply dropping the relevant parts of Carnot's theory.

For two reasons, this position is difficult to maintain. First, as we shall see, some of the relevant inconsistencies can *only* be detected while making inferences from a set of constraints which includes *both* Carnot's theory and Joule's ideas (see section 5.1). Next, as far as I know, there is no reason at all to assume that Clausius initially endorsed Carnot's theory, in general, less than Joule's ideas. (There seems to be only one exception: we have every reason to believe that Clausius originally assigned a lower preference to C2 than to J2). Moreover, as we shall see, Clausius's analysis led not only to the abandonment of some parts of Carnot's theory but also to the rejection of some parts of Joule's ideas — among which J2! So, whatever preferences Clausius originally assigned to Carnot's theory and Joule's ideas, particular parts of the former were eventually judged more reliable than particular parts of the latter.

In view of this, there are only two options. (i) We assume that Clausius started from a situation in which a (more or less) clear preference-ranking

[9] Clausius's problem includes some 'personal' constraints — constraints that are typical for Clausius. The mechanical model of heat, for instance, was by no means generally agreed upon by the relevant scientific community. In my 199+, I show that these and other personal constraints played a crucial role in redefining and solving the problem.

was assigned to the constraints of the problem. But then we also have to assume that this preference-ranking was adjusted while he was working on the problem (it is extremely implausible that Clausius originally preferred C2 above J2). (ii) We assume that Clausius assigned a high preference to those assumptions he judged valid right from the start (the relevant experimental results and his mechanical model of heat), but that he did not have any (internal or external) grounds to assign a clear preference-ranking to the other constraints of the problem. The upshot is that it really makes no sense to assume that the decisive preference-ranking was fixed *beforehand*: it was gradually developed while Clausius was working on the problem.

Given that there was no alternative than to work with an inconsistent set of constructive constraints, the problem is *ill-defined.* (From the point of view of classical logic, everything follows from it, for instance classical arithmetic. But clearly, this 'solution' does not satisfy the limiting constraints.)

The reasoning process that led to the solution of this problem can be made comprehensible by assuming that Clausius carefully investigated which parts of Carnot's theory and of Joule's ideas could be retained and which parts had to be modified or even rejected, but only after he had reorganized them. This reorganization resulted in several changes to the set of constructive constraints here defined: some were abandoned, others were modified, and new ones were added. Throughout these changes the original problem became gradually better defined, so that, finally, a solution was found.

4.2. *A brief survey of Clausius's reasoning process*

One of the earliest stages in Clausius's reasoning process concerned the simultaneous analysis of both Carnot's theory and Joule's ideas in order to arrive at a new interpretation of the Carnot-engine. From C3 and J1, Clausius drew the following conclusion:

R1 Operating in the normal direction, a Carnot-engine results in the production of work, in the transfer of heat from the hot to the cold reservoir, *and* in the conversion of heat into work; operating in the reversed direction, a Carnot-engine results in the consumption of work, in the transfer of heat from the cold to the hot reservoir, *and* in the conversion of work into heat.

Having derived this result, Clausius interpreted a Carnot-engine as a device which (operating in the normal direction) absorbs an amount of heat from a hot reservoir, transfers *part* of it to a cold reservoir and converts *part* of it into work (so, the transfer is 'partial' rather than 'total'). The transfer of heat he interpreted as the *conversion* of heat at high temperature into heat at low temperature. (Note that this interpretation, unlike Carnot's interpretation in terms of a 'fall of caloric', is in line with Clausius's mechanical model of heat.)

From R1 together with some common-sense assumptions, it is possible to derive

R2 The amount of heat absorbed by a Carnot-engine at one reservoir is *not* equal to the amount of heat delivered at the other.

Thus, the reinterpretation of the Carnot-engine results in the discovery of a 'new' contradiction: R2 contradicts C4. However, as Clausius must have realized, C4 is related to one of Carnot's *suppressed assumptions* (which is clearly inconsistent with Joule's ideas), namely

R3 Heat, being a material substance, can neither be consumed nor produced.

If R3 is abandoned, C4 is no longer derivable (and hence, the contradiction resulting from C4&R2 will be solved). After a series of derivations, in which Clausius established that the new interpretation of the Carnot-engine is at least as successful in deriving empirical results as Carnot's, he decided to eliminate R3. Although fundamental to Carnot's theory, it is incompatible with his mechanical model of heat. Moreover, he had by that time sufficient evidence that there were no independent grounds for sticking to this particular principle: abandoning it did not make the reorganized theory empirically less successful than Carnot's theory (quite to the contrary).

In view of this decision, Clausius also abandoned C1 but only after he had 'divided' it into several 'parts'; of these he retained

R4 The production of work by the agency of heat is always accompanied by the transfer of heat from a hotter to a colder body.

According to Clausius's own account, R4 forms the essential part of C1: it is significantly verified by experience, and some of Carnot's most important results are based on it. Moreover, having reinterpreted the Carnot-engine, Clausius realized that a transfer of heat does not necessarily conflict with the idea of an actual *consumption* of heat. As the example of the reinterpreted Carnot-engine shows, "it may very well be the case that at the same time a certain quantity of heat is consumed and another quantity transferred from a hotter to a colder body." (p. 7, p. 112)

There were several other changes to the problem. As R3 was abandoned, C4 did not follow any longer, and the interpretation of the other assumptions of Carnot's theory changed accordingly — C3 was no longer interpreted in terms of (unchangeable) caloric, and the interpretation of C5 was determined by the new interpretation of the Carnot-engine.

These changes led to the resolution of several inconsistencies (among which those arising from C1&J1 and C4&R2). However, Clausius's total set of constructive constraints was still inconsistent. More specifically, the contradiction resulting from C2&J2 was not yet solved. There was something more, however.

In the same way as Clausius reinterpreted (parts of) Carnot's theory in the light of Joule's ideas, he analyzed the latter in view of the former. At some point then, Clausius must have realized the following. Suppose that a body is heated by means of friction. According to J1 the work thus expended can be entirely recovered — it suffices to convert the heat produced back into work. But, if C5 holds true, it is impossible to design an engine that produces more work than a Carnot-engine. Given the new interpretation of the Carnot-engine, this entails that, if all the heat produced by friction is absorbed by even the most efficient engine, only *part* of it will be converted into work (the remainder being *transferred* to a second reservoir). This entails, in contradiction to J1,

R5 There are cases in which the work consumed in producing an amount of heat is not entirely recoverable.

It also entails, in line with C2 but in contradiction to J2, that some processes result in the loss of work.

These new difficulties must have been considered by Clausius as extremely serious. When he started working on the problem, he considered J1 as well as J2 as fundamental and indisputable principles. On the other hand, C5 formed one of the central elements of Carnot's theory from which several important results were derived.

How did Clausius react? First, he divided J1 into parts and identified the part that contradicts R5:

R6 The work consumed in producing heat is entirely recoverable.

(If C5 is adopted only this particular part of J1 has to be abandoned — the remainder can be retained.)

Next, he tried to find out whether or not C5 still follows from the reorganized theory — for a detailed discussion of this stage in his reasoning process, see section 7.1. As we shall see, Clausius was able to derive C5 by *reductio ad absurdum*: from the hypothesis that C5 is false, it follows that heat can be transferred from a cold to a hot body without the expenditure of work. This contradicts a principle Clausius considered absolutely fundamental (p. 32, p. 134):

R7 It is impossible, without the expenditure of work, to transfer heat from a cold to a hot body.

(Most probably, this particular principle was added to the set of constraints while working on the problem.) As we shall see in more detail in section 7.2, the derivation of C5 had an unexpected side-effect: it showed that the adoption of R6 would make it impossible to arrive at a consistent theory. This provided good (internal) grounds for assigning a low preference to R6 (and hence to J1), and also to J2.

Having obtained these results, Clausius started a new series of derivations, in which C5 was successfully put to the test. Eventually, Clausius eliminated R6 (and hence J1) as well as J2. This decision led to the resolution of all the remaining inconsistencies (but would have been inconceivable when Clausius started working on the problem).

5. *Some characteristics of Clausius's reasoning process*

5.1. *Detecting the relevant inconsistencies required reasoning*

It is typical for Clausius's search process that, although he was working with an inconsistent set of constraints, he had no alternative than to make inferences from it. Simply abandoning halves of the contradictions that were explicit when he started working on the problem did not work. The reason for this is threefold — in this section I discuss a first reason, two others will be discussed in sections 5.2 and 5.3.

When Clausius started working on the problem, he was well aware of two contradictions that result from the joint adoption of Carnot's theory and Joule's ideas — C1 contradicts J1 and C2 contradicts J2. However, these are not the only contradictions he had to worry about. As we have seen, contradictions also result, for instance, from C4&R2 and from R5&R6. It is typical for these contradictions that they are implicit or *hidden*: they cannot be recognized *at sight* but can only be detected through a thorough analysis of the set of constructive constraints.

In order to arrive at a consistent theory, not only the manifest contradictions but also the implicit contradictions had to be resolved. Besides, resolving the manifest contradictions presupposed the detection and resolution of inconsistencies that were more deeply entrenched — for instance, resolving the contradiction that results from C1&J1 presupposes the detection and resolution of the contradiction concerning R3.

In view of all this, Clausius had no alternative than to *analyze* both Carnot's theory and Joule's ideas and to make *explicit* the *specific* contradictions which result from their joint adoption. In order to achieve this end, it was necessary to make inferences from *inconsistent* constraints: some contradictions which result from the joint adoption of Carnot's theory and Joule's ideas can only be detected if consequences are derived from their *union* — it is seldom the case that one half of a contradiction follows from Carnot's theory and the other half from Joule's ideas. Take, for instance, the contradiction resulting from R5&R6. An analysis of Carnot's theory and Joule's ideas *independently* from each other would never have revealed that this particular contradiction results from their *joint* adoption, for R5 can *only* be derived from the *union* of both Carnot's theory and Joule's ideas. From Joule's ideas, the negation of R5 is derivable but not R5 itself; within Carnot's theory, R5 does not even make sense (for it refers to the *conversion* of work into heat). A similar point can be made for the contradiction resulting from C4&R2.

Note especially that, although Clausius was working with an inconsistent set of constraints, detecting the hidden contradictions was *far from trivial*. This holds especially true for the contradiction resulting from R5&R6. But also the contradiction that results from C4&R2 was not self-evident. As a matter of fact, this particular contradiction was not detected by Holtzmann: although he accepted the idea that work results from the *conversion* of heat, he retained Carnot's interpretation of the Carnot-engine, and hence (implicitly) assumed that *all* the heat absorbed from one reservoir is delivered to the other. But, from the combination of Carnot's interpretation of the Carnot-engine and the idea that work results from the conversion of heat, it also follows that this is impossible.

5.2. *Resolving the inconsistencies required a reorganization of the relevant constraints*

There is a second reason why Clausius had to continue reasoning from an inconsistent set of constraints. Resolving the (manifest) contradictions by simply abandoning parts of Carnot's theory and Joule's ideas would have resulted in a ‘theory’ that was much too poor (if there would be anything left at all). An example may help to clarify this.

Suppose that Clausius, contrary to the facts, would have decided to resolve the contradictions resulting from C1&J1 and from C2&J2 by simply abandoning C1 and C2. As he thus would have rejected the idea that the production of work (by a heat engine) results from the *transfer* of heat, it seems inevitable to eliminate C3 as well. But also C4 and C5 are in danger, for the idea of a transfer of heat is essential for Carnot's interpretation of the Carnot-engine. So, where this strategy comes to is that Carnot's theory is eliminated *in its entirety*.

This is precisely what confused Kelvin (and most of his contemporaries): abandoning parts of Carnot's theory in favour of Joule's ideas, seems to lead unavoidably to the *total* abandonment of the former. But also the converse seems to hold true: if Carnot's theory is adopted, it seems inevitable to abandon both J1 and J2. Both outcomes were considered as highly undesirable — remember that Joule's ideas were conceived as fundamental principles, and that Carnot's theory formed the only theory available for explaining thermodynamic phenomena.

Clausius realized that there is a way out. It is possible to retain the non-problematic parts of both Carnot's theory and Joule's ideas, *provided* that the former as well as the latter are *analyzed* and *reorganized*. We have seen, for instance, that Clausius divided C1 as well as J1 into parts, and that he rejected *only some of these* — all the non-problematic parts were retained.

So, an analysis of the set of constraints was required not only in order to detect the relevant inconsistencies but also in order to 'preserve' as much as possible of the non-problematic parts of both Carnot's theory and Joule's ideas. Note that here too it is necessary to reason from the inconsistencies: some of the parts which Clausius eventually retained, can only be derived from the *union* of Carnot's theory and Joule's ideas. (One might think here of R1 and also of R5.)

5.3. *Resolving the inconsistencies required the development of an adequate preference-ranking*

The third reason why Clausius had to reason from an inconsistent set of constraints, has to do with the assignment of preferences. It is typical that Clausius, when he started working on the problem, did not have an *adequate* preference-ranking for the halves of the relevant contradictions. Given the complexity of the problem, such a preference-ranking had to be based on *internal* grounds, and hence, required the logical analysis of the set of constraints. Let me try to clarify this.

As I mentioned in section 4.1, Clausius was aiming at a theory that would not only be consistent but also empirically successful. In view of this goal, it was of utmost importance that he retained the 'good' parts of both Carnot's theory and Joule's ideas. However, this could not be decided *at sight*. In order to identify those parts of Carnot's theory and Joule's ideas that were essential for the desired theory, it was necessary to make inferences: the relevant constraints had to be analyzed, problematic parts of the constraints had to be identified, and seemingly non-problematic parts had to be put to the test. Any preference-ranking suited to obtain the desired goal had to be based on this analysis: a lower preference had to be assigned to halves of contradictions that are problematic in the light of the desired theory than to those that are essential for it.

This is exactly how Clausius proceeded. The way in which he resolved the contradiction resulting from C1&J1 forms an excellent example. He analyzed C1, identified its problematic parts (those that were incompatible with his mechanical model of heat), and put the remainder (R4) to the test. The decision to assign a high preference to R4 was based on this analysis — he had discovered that the rejection of R4 would make the resulting theory empirically less successful, while adopting it did not cause any conflict with his mechanical model of heat. Similarly, for the contradiction resulting from C2&J2. Although it is most likely that Clausius originally assigned a higher preference to J2 than to C2, the former was eventually abandoned in favour of the latter. As we shall see in more detail in section 7.2, this decision too was based on a logical analysis of the relevant constraints.

5.4. *Reasoning from the inconsistencies resulted in a multiplicity of conceptual shifts*

It is remarkable that Clausius, while making inferences from the union of Carnot's theory and Joule's ideas, did not worry much about Carnot's and Joule's *interpretation* of the central terms and statements. The importance of this strategy can hardly be overestimated. It resulted in several interesting *conceptual changes* which played an important part in the resolution of the inconsistencies.

The reinterpretation of the Carnot-engine forms a nice example (other examples can be found in my 199+). If one takes into account Carnot's interpretation of "transfer of heat" and Joule's interpretation of "conversion", the conclusion that a Carnot-engine (in the normal direction) results in the transfer of heat from one reservoir to another *and* in the conversion of heat into work is a contradiction in terms. Nevertheless, Clausius derived this conclusion. Seemingly, he manipulated terms and statements, without paying much attention to their original interpretation, and looked afterwards for an interpretation that would fit the outcome. Thus, he was able to arrive at a new interpretation of a Carnot-engine: having derived the aforementioned conclusion, he assumed that only *part* of the heat is transferred and that the 'transfer' is actually a conversion of heat at high temperature into heat at low temperature. This interpretation makes the conclusion not only intelligible but also compatible with his mechanical model of heat.

Note especially that this and similar conceptual shifts (which formed key elements in Clausius's discovery) can *only* be comprehended as the result of reasoning from the *union* of Carnot's theory and Joule's ideas (see also my 199+). Analyzing Carnot's theory and Joule's ideas *independently* from each other would never have yielded this particular reinterpretation of the Carnot-engine: on the basis of Carnot's theory alone, one cannot possibly infer that a Carnot-engine *converts* heat into work; on the basis of Joule's ideas, it is impossible to derive that a Carnot-engine *transfers* heat.

This situation changes radically if inferences are made from the *union* of Carnot's theory and Joule's ideas. In that case, *but only in that case*, one is able to derive that a Carnot-engine (in the normal direction) transfers some heat from a hot to a cold reservoir *and* converts some heat into work.

Someone might object that the idea of a (partial) transfer is entirely compatible with Joule's ideas. This certainly holds true. Joule's ideas by no means exclude that some of the heat absorbed by a specific engine is transferred to a second reservoir (the remainder being converted into work). However, this does not undermine my claim that this particular interpretation of the Carnot-engine cannot be obtained on the basis of Joule's ideas alone: it is impossible to decide, on the basis of these ideas, whether or not the conversion of heat is, in some cases, accompanied by the *transfer* of heat.

5.5. *The inconsistencies were gradually eliminated while working on the problem*

It is typical for Clausius's search process that the contradictions resulting from the union of Carnot's theory and Joule's ideas were eliminated *bit-by-bit*. Clausius did not follow a two-stage-strategy in which, first, *all* the relevant contradictions were derived, and next, choices were made between halves of these. Quite to the contrary. He concentrated on *specific* contradictions and analyzed the set of constraints in view of these. As soon as the logical analysis provided sufficient grounds for the abandonment of one half of a specific contradiction, the latter was eliminated, even if this move did not make the total set of constraints consistent. (For instance, when Clausius eliminated the contradiction resulting from C1&J1 no attempt was made to resolve at the same time the contradiction resulting from C2&J2.) Having thus obtained a less inconsistent set of constraints, research was continued in order to detect other contradictions and to fix further preferences.

The stepwise elimination of inconsistencies enabled Clausius to define consistent subsets of constraints — the reinterpretation of the Carnot-engine forms a good example. These subsets were heuristically important, because they functioned as reference points: as soon as a consistent subset was defined, research was directed at extending this subset without making it inconsistent.

With the advantage of hindsight, someone might try to defend the idea that the solution to Clausius's problem is actually quite simple: take Carnot's theory, abandon one of its central assumptions (R3) together with its consequences, and combine the remainder with the relevant parts of J1. But, as should be clear from this and the previous sections, this really makes no sense, neither conceptually nor historically. (i) Clausius's theory is *not* a reduction of Carnot's theory combined with some ideas of Joule's. It consists of parts of a *reorganized* version of both Carnot's theory and Joule's ideas; in this reorganized theory all the central terms and statements obtained a new meaning. (ii) There was no procedure to eliminate the relevant inconsistencies in a single move — they had to be resolved step-by-step while working on the problem. (iii) When first confronted with the problem, neither Clausius nor one of his contemporaries would have thought of this particular solution. More specifically, nobody would have seriously considered the idea to reject J2. It was only through a logical analysis of the problem that this idea *gradually* forced itself upon Clausius.

6. *The logic underlying Clausius's reasoning process*

Starting from the characteristics discussed in the previous sections, I shall now show that classical logic as well as mixed non-monotonic logics (including Rescher-like mechanisms) as well as (monotonic) paraconsistent logics are equally unsuitable to understand the reasoning in this and similar creative processes. The only logic known to me by which it is possible to make sense of Clausius's reasoning process is an inconsistency-adaptive one.

6.1. *Why Classical Logic is inadequate*

It is well known that, from the point of view of classical logic, it is impossible to reason (sensibly) in the presence of inconsistencies. This leaves one with a single procedure to deal with inconsistencies: given a set of premises from which inconsistencies arise, one decides on *external* grounds, *without making any inference from the premises at issue*, which (parts of the) premises will be eliminated (in such a way that the result is consistent); having thus obtained a consistent set of statements, and only under this condition, it makes sense to apply the rules of classical logic.

The problem with this procedure is that it only works under the following conditions: (i) the premises involved are divided (beforehand) into workable 'parts', and (ii) there are clear criteria for the *elimination* of (parts of the) premises such that the outcome is consistent. The rationale for (i) is easy to understand: being unable to make sensible inferences from an inconsistent set of statements, one has no means for analyzing the premises.

It is clear that these conditions were not satisfied in the case of Clausius: (i) the constraints had to be *analyzed* and even *reinterpreted* in order to obtain workable parts (see section 5.2), and (ii) adequate criteria for the elimination of parts of the constraints had to be developed while working on the problem (see section 5.3). Given that the desired reorganization of the constraints and even the assignment of preferences involved reasoning from the *union* of Carnot's theory and Joule's ideas, and hence from inconsistencies, it is clear that classical logic is inadequate to understand Clausius's search process. There is something more, however.

As I have argued, it is central for our understanding of Clausius's reasoning process to recognize (i) that several of the relevant inconsistencies were *hidden* and (ii) that they could only be detected by making *inferences* from the *union* of Carnot's theory and Joule's ideas in a way that was *far from trivial* (see section 5.1). However, for the classical logician, these phrases do not even make sense. The classical logician is convinced that all sensible reasoning stops in the presence of inconsistencies: from an inconsistent set of premises *any* contradiction can be derived in a *completely trivial way*. So, from his point of view, the difference between explicit contradictions and hidden contradictions is meaningless: as soon as one has detected only one contradiction, *all* the others can be *trivially* derived. Still this is not all.

I have argued that the various contradictions which result from the joint adoption of Carnot's theory and Joule's ideas were resolved *step-by-step*; throughout this process the set of constraints became gradually less inconsistent (see section 5.5). All this is incomprehensible from the point of view of classical logic: from an inconsistent set of premises *all* possible contradictions follow. Hence, for a classical logician, the distinction between more and less inconsistent sets of premises makes no sense.

Someone might object that classical logic can be sensibly applied to consistent subsets of the set of constraints. This indeed holds true. Once Clausius arrived at a consistent reinterpretation of the Carnot-engine, classical logic could be applied to it. Note however that this and similar consistent subsets could only be defined by making inferences from inconsistent constraints. Hence, classical logic is inadequate to understand the construction of consistent subsets.

In view of all this, it should not surprise us that the standard reconstruction of Clausius's discovery, which *is* based on classical logic, is wanting (for the standard account and its shortcomings, see my 199+). The authors I consulted pay attention *neither* to the fact that a lot of reasoning was required in order to detect and resolve the various inconsistencies, *nor* to the fact that Clausius developed a *novel* theory (in which all the central terms of Carnot's theory obtained a new meaning). These shortcomings are comprehensible if one takes into account that these authors presupposed that Clausius (implicitly) followed the classical logician's procedure for dealing with inconsistencies. Starting from this presupposition, one is completely blind for the many interesting arguments (at the moment it makes sense to apply classical logic to the total set of constraints the problem is *solved*) as well as for the novelty of Clausius's theory (how could it be novel if Clausius merely dropped parts of Carnot's theory and combined the remainder with Joule's ideas).

Such reconstructions have disastrous consequences for the understanding of discovery, and so has the underlying classical logician's view in general. As soon as one realizes that the elimination of inconsistency is not self-evident (as in the case of Clausius), the classical logician's view does not enable one to study the reasoning process that led to the solution. One can just pretend, in accordance with the romantic view on creativity, that the problem solver, for no reason at all, suddenly had a flash of insight.

6.2. *Why a reconstruction from the point of view of mixed non-monotonic logics does not do either*

In one respect, mixed non-monotonic logics show a strong resemblance with classical logic: from both points of view it is impossible to make inferences (in a sensible way) from an inconsistent set of premises. The difference is that, from the point of view of mixed non-monotonic logics, one is rather reluctant to drop (parts of the) premises (in order to eliminate inconsistencies). Instead one tries to 'prepare' the set of premises in such a way that no inconsistencies arise from it. The idea is actually quite simple: the occurrence of inconsistencies is precluded by stipulating that some universally quantified formulas of the form $(\forall x)(Px \supset Qx)$ are turned into rules-with-exceptions of some form. Roughly speaking, a rule-with-exceptions applies only under circumstances in which it does not lead to an inconsistency. Specific mixed non-monotonic logics differ from each other in the way rules-with-exceptions are formalized. (For a survey of mixed non-monotonic logics, see Łukasziwicz 1990.)

This technique results in the following procedure for dealing with inconsistent theories: given a set of premises from which inconsistencies may arise, one decides which universally quantified formulas will be turned into rules-with-exceptions. Note that also from the point of view of mixed non-monotonic logics, the decision which is crucial for the elimination of inconsistencies, *cannot be taken but on external grounds*. It is easy to see why. Just like from the point of view of classical logic, one is not able to reason sensibly in the presence of inconsistencies. So, mixed non-monotonic logics do not offer a means to *analyze* inconsistent sets of premises.

As far as I know, a problem of the complexity of Clausius's has never been studied from the point of view of mixed non-monotonic logics. However, this does not imply that a reconstruction of Clausius's reasoning on the basis of a mixed non-monotonic logic should be impossible. Such a reconstruction might have great advantages in comparison to reconstructions based on classical logic. On the latter one can only decide to abandon part of the premises, whereas the former seems to allow for keeping all premises involved, provided some universally quantified statements are turned into rules-with-exceptions.

Matters are not that simple, however. Any reconstruction of Clausius's reasoning process has to take into account at least the following. (i) At the moment Clausius started working on the problem, the set of constructive constraints was *not* prepared in a non-monotonic way (inconsistencies *did* arise from it). (ii) Before analyzing the *union* of Carnot's theory and Joule's ideas, it would have been impossible to decide which universally quantified statements had to be turned into rules-with-exceptions. The reason for this is twofold. First, most problematic statements could only be recognized as such after a thorough analysis (see sections 5.1 and 5.2). Next, adequate criteria to choose between problematic statements were not fixed *beforehand* but had to be designed while working on the problem (see section 5.3). The upshot is that Clausius's creativity cannot possibly lie in the *application* of a mixed non-monotonic logic, but in the *reorganization* of the set of constraints that eventually led to a consistent result. As we have seen, this reorganization was not due to some flash of insight, but resulted from the *simultaneous analysis* of Carnot's theory and Joule's ideas.

Starting from these facts, one easily sees that a reconstruction on the basis of a mixed non-monotonic logic meets several problems. (i) It makes no sense to assume that Clausius, in order to arrive at his theory, implicitly used a mixed non-monotonic logic, for this presupposes that *all* problematic statements were identified and that preferences were adequately fixed for all of them. These conditions were not satisfied when Clausius started working on the problem. As a matter of fact, they were only satisfied at the moment the problem was *solved*. (ii) It might be possible to devise a reconstruction according to which Clausius's theory results from non-monotonic versions of Carnot's theory and Joule's ideas (and the available experimental results). But such a reconstruction, which is only possible *after* and in view of Clausius's analysis and reorganization, would reveal *nothing* about the creative process itself because the latter consists precisely in that analysis and reorganization. (iii) One might look for a reconstruction according to which the non-monotonic reorganization of both Carnot's theory and Joule's ideas was designed by Clausius while he was working on the problem. But again, the logic underlying *this* enterprise should have allowed Clausius to reason sensibly in the presence of inconsistencies, and hence cannot possibly have been a mixed non-monotonic logic. (iv) Even if the previous problems were surmountable, the fact remains that Clausius's reorganization of both Carnot's theory and Joule's ideas was *conceptual*: the meaning of all central terms and statements changed. So, it seems that it even does not make sense to try to obtain Clausius's theory by simply reducing some of Carnot's and Joule's general assumptions to rules-with-exceptions.

Still this is not all. From the point of view of mixed non-monotonic logics, like from the point of view of classical logic, all sensible reasoning stops in the presence of inconsistencies: from an inconsistent set of premises anything follows. Hence, mixed non-monotonic logics are inadequate to understand some of the most important characteristics of Clausius's search process, namely (i) that he had to *reason*, in a non-trivial way, in order to detect *hidden* contradictions, and (ii) that, as a result of the stepwise elimination of specific contradictions, the set of constraints became gradually *less* inconsistent.

Summarizing: whatever the advantages, for the present purpose, of mixed non-monotonic logics as compared with classical logic, they are equally inadequate for the understanding of creative processes which involve inconsistent constraints. From neither point of view it is possible to study the *reasoning* which eventually led to the resolution of inconsistency.

6.3. *A note on Rescher's 1964 mechanism*

Readers familiar with Rescher-like mechanisms (Rescher 1964) may have noticed that these are not suitable to reconstruct Clausius's reasoning process either. The construction of maximally consistent subsets presupposes that the premises involved are divided into *appropriate* parts — otherwise there are only two maximally consistent subsets: Carnot's theory on the one hand and Joule's ideas on the other. This condition was not satisfied when Clausius started working on the problem. In defense of a reconstruction in terms of Rescher-like mechanisms, one might suggest that Clausius analyzed both sets of premises independently of each other, thus dividing each of them into several parts, *before* he looked for maximally consistent subsets of those parts.

There are, however, two difficulties with such an analysis. (i) As we have seen, Clausius did *not* analyze Carnot's theory and Joule's ideas *indepen-*

dently of each other. Quite to the contrary. From the very outset he reinterpreted Carnot's theory and Joule's ideas in the light of each other, and hence reasoned from an inconsistent set of premises. (As we have seen in section 5.4, it is only by virtue of the *combination* of Carnot and Joule that one is able to understand the multiplicity of conceptual changes: the latter can only be comprehended as the result of reasoning from the *inconsistencies*, and not as the result of selecting some consistent subset of the union of both theories.) (ii) Even if Clausius had analyzed Carnot's theory and Joule's ideas independently of each other, it is by no means clear in what way one should define maximally consistent subsets of the obtained 'pieces' of these sets. It is well-known that Rescher's mechanism is highly sensitive for the formulation of the premises. But there was no obvious single way for Clausius to cut up Carnot's theory and Joule's ideas.[10] In other words, applying Rescher's mechanism to such cases would introduce a high amount of arbitrariness in creative processes (and hence would obfuscate the programme to approach creativity in terms of a problem solving methodology).

6.4. *Why monotonic paraconsistent logics do not offer a way out*

Monotonic paraconsistent logics offer a procedure for dealing with inconsistent sets of premises that is quite different from the procedures previously discussed. Rather than trying to find a consistent reformulation of the premises at issue (for which the price may be quite high), one reconciles to the inconsistencies involved and looks for a logic that allows one to reason sensibly in their presence. The idea is that deriving consequences from an inconsistent set of premises does *not* lead to triviality, provided that one uses a properly designed logic. In practice, this comes to using a logic that is *less powerful* than classical logic (one in which *disjunctive syllogism* and many other rules do not hold).

At first sight, this strategy looks quite promising. Unlike the logics previously discussed, monotonic paraconsistent logics seem adequate to understand central aspects of Clausius's search process. If the logic underlying Carnot's theory and Joule's ideas is replaced by a more restricted one, only *some* contradictions follow from their union and deriving a particular contradiction may be far from trivial. Hence, it seems almost self-evident that some contradictions which follow from the joint adoption of Carnot's theory and Joule's ideas were *hidden* and that a good deal of *reasoning* was

[10] That there are ways to do so is a preconception that is typical for manuals, but is unrealistic with respect to creative science.

required in order to detect them. Moreover, as monotonic paraconsistent logics allow one to reason sensibly in the presence of inconsistencies, the many conceptual shifts (which resulted from reasoning from the inconsistencies) seem easy to account for. So, are (monotonic) paraconsistent logics suitable for understanding the process by which Clausius arrived at his theory? Obviously not.

First, Clausius did *not* reconcile to an inconsistent set of premises. The inferences he made from the union of Carnot's theory and Joule's ideas were meant to *detect* and *eliminate* the inconsistencies involved. Next, Clausius eventually succeeded in formulating a consistent theory. The logic underlying this theory is clearly classical. So, whatever logic Clausius may have used while *searching* for a new theory, monotonic paraconsistent logics are inadequate to understand the result. But, there is a further and even more convincing argument. It is a remarkable fact that Clausius, although he was reasoning from an inconsistent set of constraints, used a logic that was, at particular points, *as powerful as classical logic*. (In section 7.1, I discuss an argument in which *reductio ad absurdum* is applied; other examples of 'rich reasoning' are found in my 199+). Monotonic paraconsistent logics are clearly too poor to make sense of this. Hence, monotonic paraconsistent logics are not only inadequate to understand Clausius's final result, they are, more importantly, inadequate to understand his reasoning process.

6.5. *What we need*

Inconsistency-adaptive logics were developed by Diderik Batens (see his 1985, 1986, 1989, and 199+b) for the logical reconstruction of the following paradigm case. A theory T, based on classical logic, and hence intended to be consistent, turns out to be inconsistent (one might think of Russell discovering his paradox in Frege's set theory). If T is considered to be interesting, one will want to replace T by a (consistent) theory T'. Batens correctly observes that T itself is heuristically important in this connection: we want a large number of theorems of T, actually all 'good' ones, to be theorems of T'. However, we can only make sense of this requirement if we know which sentences are theorems of T. Here comes the rub. If we stick to classical logic, *all* sentences are theorems of T. If we substitute some (monotonic) paraconsistent logic for classical logic in T, the resulting set of theorems is *too poor*.

So, how should we proceed? As Batens argues, the problem with monotonic paraconsistent logics is that they restrict the rules of inference *globally*; for instance, they take *disjunctive syllogism* to be incorrect *in general*. In order to specify the 'good' theorems of T, we need a logic that *localizes* inconsistencies and *modifies* the rules of inference in view of these. This is precisely what inconsistency-adaptive logics do: 'in the neighbourhood of' inconsistencies, they go paraconsistent; everywhere else, classical logic is applied in its full strength. Inconsistency-adaptive logics 'oscillate' between a (poor) paraconsistent logic and classical logic, and they do so in view of the specific inconsistencies of the theory.

Inconsistency-adaptive logics based on the minimal paraconsistent logic PI and its predicative version PIL are, at this moment, best understood (PI contains the full positive fragment of classical logic as well as the negation-completeness half, $A \vee \sim A$, of the meaning of negation). As Batens has shown in his 1989 (pp. 211-213), inconsistency-adaptive logics based on PI, localize inconsistencies 'as much as possible'. This feature makes them preferable to alternative inconsistency-adaptive logics for the reconstruction of the paradigm case described above.

The axioms for PIL are listed in Batens 199+a (this volume). A detailed discussion of the proof theory of inconsistency-adaptive logics is found in Batens 1989 and 199+b. I merely summarize the idea behind the proof procedure for APIL1 (an inconsistency-adaptive logic based on PIL).[11]

Let $\mathrm{DEK}(A_1, \ldots, A_n)$ be $\exists(A_1 \& \sim A_1) \vee \ldots \vee \exists(A_n \& \sim A_n)$, where $\exists(A_i \& \sim A_i)$ is $(A_i \& \sim A_i)$ proceeded by an existential quantifier over each free variable that occurs in A_i. In accordance with Batens 199+b, we say that A behaves consistently at a stage of a proof iff $\exists(A \& \sim A)$ does not occur in the proof at depth zero[12] at that stage; we say that the consistent behaviour of A_1 is connected to the consistent behaviour of $A_2, \ldots, A_n$ at a stage of a proof if $\mathrm{DEK}(A_1, \ldots, A_n)$ occurs in the proof at depth zero whereas $\mathrm{DEK}(A_2, \ldots, A_n)$ does not occur in the proof at depth zero. Given these stipulations, the rules of inference of APIL1 are of two kinds.

Unconditional rules include all those rules of classical logic that hold in PIL — for instance, *conditional proof* and *modus ponens*. These rules may be applied in an-APIL1-proof in all circumstances (even if applying them leads to inconsistency).

For each of the *conditional rules* a requirement is specified (the requirement always concerns the consistent behaviour of specific formulas). A conditional rule may be applied in an-APIL1-proof provided that the requirement is fulfilled; if a wff A is added to the proof by application of a conditional rule and the requirement is not any more fulfilled at a later

11 In Batens 199+b, a slightly different inconsistency-adaptive logic APIL2 is also studied. The differences between both systems are not essential for the purposes of the present paper.

12 In other words: not depending on any hypothesis.

stage of the proof, then A as well as all wffs derived from A are deleted (at the later stage).[13] Some examples of conditional rules:

$A \vee B, \sim A \vdash B$	*provided* that A behaves consistently and the consistent behaviour of A is not connected to the consistent behaviour of any other wffs.
$A \supset B, \sim B \vdash \sim A$	*provided* that B behaves consistently and the consistent behaviour of B is not connected to the consistent behaviour of any other wffs.

In view of section 7.1, I also specify the requirement for *reductio ad absurdum*:

> given a derivation of $B \& \sim B$ from the hypothesis A, one may conclude to $\sim A$ *provided* that (i) B behaves consistently *with respect to the premises* and (ii) the consistent behaviour of B is not connected to the consistent behaviour of any other wffs.

As the reader may expect, inconsistency-adaptive logics are non-monotonic. However, unlike *mixed* non-monotonic logics, they do not involve non-logical preferences (Batens, 199+a). As such, they can cope with sets of inconsistent premises to which no clear preference-ranking is assigned (yet).

Unlike the logical systems previously discussed, inconsistency-adaptive logics offer an excellent tool for understanding creative processes that involve inconsistent constraints. This type of logic enables one to *analyze* an inconsistent set of premises 'as classically as possible' (without the risk of falling unwarranted into triviality), and thus to *reorganize* the set of constraints and to *detect* the *specific* inconsistencies that arise from it. As an inconsistency-adaptive logic *localizes* the specific inconsistencies and *adapts* itself to these, instances of rich reasoning can be accounted for. As it does not involve non-logical preferences, the logical analysis does not presuppose that a preference-ranking is fixed beforehand. This makes it possible to analyze the set of constraints, to fix preferences on the basis of this analysis, and to eliminate inconsistencies in view of these. If the set of

[13] Inconsistency-adaptive logics have a *dynamic* proof procedure. Batens (see his 1989, 199+b) has shown, however, that for any set of inconsistent premises there is a unique set of wffs *finally derivable* from it. (A wff A is finally derivable from a set of premises α if and only if it is possible to construct a proof of A from members of α in such a way that A will remain derivable at any stage in any continuation of this proof from α.)

constraints is still inconsistent, research can be continued in order to detect further inconsistencies and to fix further preferences.

A logic with these features is precisely what is needed to make sense of Clausius's reasoning process. If one assumes that Clausius implicitly used a logic of this type, all the characteristics of his search process can easily be accounted for. Note again, however, that an inconsistency-adaptive logic is *not* a device for *resolving* inconsistencies. The question which half of an inconsistency should be *abandoned* is a matter of *preferences*, and these are not embodied in the logic itself. Still, in interesting cases preferences may be assigned on the basis of a *logical analysis* of the premises (see also section 7.2).

In his 199+a, Batens proposes an interesting procedure for the *resolution* of inconsistencies in which inconsistency-adaptive logics play an important part (the procedure is intended for the reconstruction of mixed non-monotonic logics). The procedure consists of three components. A *deductive* component leads from the premises to a possibly inconsistent consequence set — inconsistency-adaptive logics prove most suitable for this component. The two other components 'weed out the inconsistencies'. A purely logical component connects a set of consistent models to the 'set' of (possibly inconsistent) models of the premises. A preferential component selects a subset of the consistent models. In the same paper, it is also shown that the stepwise elimination of inconsistencies leads to the same result as eliminating them all at once. However, as Batens notices himself, this result has little importance for the reconstruction of creative processes. As consistent models may be ruled out prematurely (choices are unavoidably made on the basis of present insights), and as the set of constraints may be extended while working on the problem (new constraints may be added), there is no guarantee that, in a creative process, the deterministic character of the procedure is safeguarded (see also section 7.2).

7. *Two applications of* APIL1

In the following sections, I discuss two stages of Clausius's reasoning process in somewhat more detail. Each of them is revealing for the underlying logic. The argument discussed in section 7.1 indicates that Clausius implicitly used a logic that *localizes* the specific contradictions and *adapts* itself to these. In section 7.2, I show that the underlying logic did not *presuppose* that an adequate preference-ranking was fixed beforehand, and hence, cannot have been a mixed non-monotonic logic.

7.1. *Adapting to inconsistencies — a new derivation of* C5

In section 4.2, I discussed the context in which Clausius tried to find a new derivation for C5. Clausius himself makes some comments on this particular stage in his reasoning process in his 1863 (p. 313). Apparently, he first tried to 'replicate' Carnot's argument (in which C5 is derived from the impossibility of a perpetual motion machine) and combined this 'replication' with Joule's ideas. This combination leads to a contradiction similar to the one Carnot used in his application of *reductio ad absurdum*. Nevertheless, Clausius did *not* consider the replication as a valid derivation of C5. From the point of view of classical logic, this is remarkable and even incomprehensible. Clausius's decision makes sense, however, if one assumes that he implicitly used an inconsistency-adaptive logic. Let me try to phrase this somewhat more concretely.

Like Carnot, Clausius started from the hypothesis that it is possible to design an engine that produces *more* work than a Carnot-engine, while receiving the *same* amount of heat (henceforth, a "more efficient engine"). Next, he derived that a more efficient engine allows one to construct a perpetual motion machine. However, if the combined operation of a more efficient engine and a Carnot-engine results in a surplus of work (as Carnot claims), then a certain amount of heat is *converted* into work (according to J1). Hence, it is *not* the case that a more efficient engine allows one to construct a perpetual motion machine (for this presupposes that, during the combined operation, work is produced 'out of nothing').

Given this contradiction, it seems easy to derive C5. Clausius concluded, however, that C5 cannot be derived in this particular way (see his 1863, p. 313). Why is that so? A contradiction *is* derivable from the hypothesis together with the set of premises used for this particular argument. Why does he not apply *reductio ad absurdum*? Actually, there are two possible explanations. The first is that Clausius decided to reject the argument because he realized that it rests upon ideas already eliminated — for instance, the idea that no heat is consumed when work is produced. The other is that Clausius did not apply *reductio ad absurdum* because he realized that the contradiction is derivable from the *premises* themselves, *without* the hypothesis being added to them. It is intuitively clear that such a contradiction is of no use in an argument based on *reductio ad absurdum*. The contradiction does *not* establish the impossibility of the hypothesis but merely reflects the inconsistency of the premises. Therefore, it makes no sense to use this particular contradiction to refute the hypothesis. (The reader will have noticed that this is exactly what one would expect from the point of view of APIL1 — as the requirement for *reductio ad absurdum* is not satisfied, the rule is not applied.)

As far as I know, the historical record is insufficient to choose between both explanations. In any case, having discovered that some of the steps in the argument are problematic, Clausius developed a new argument, based *again* on *reductio ad absurdum*. (For all I know, this can only be accounted for by an inconsistency-adaptive logic.)

According to Clausius's own account, the new argument is only 'slightly' different from Carnot's argument (see his 1863, p. 313). It is indeed remarkable that Clausius's argument is structurally the same as Carnot's. Let us have a closer look at it.[14]

Following Carnot's example, Clausius starts from the hypothesis that C5 is false, but immediately skips to a *reformulation* of it (obviously in the hope that the reformulation, unlike the original hypothesis, would lead to a 'useful' contradiction). Suppose, Clausius argues, that it is possible to design an engine that produces the *same* amount of work as a Carnot-engine, while receiving a *smaller* amount of heat. (The reader may verify that this hypothesis is equivalent to the hypothesis that C5 is false.) The work produced by this (more efficient) engine (in transferring an amount of heat from the hot to the cold reservoir) could be used *in its entirety* to operate a Carnot-engine in the inverse sense. All the work produced would be consumed, or, what comes to the same (according to Joule's ideas), all the heat converted into work would be recovered. However, since the more efficient engine transfers a *smaller* amount of heat, there would be a difference in the way the heat is distributed over the two reservoirs: more heat would have passed from the cold to the hot reservoir than from the hot to the cold one. "By repeating these two processes alternately it would be possible, without any expenditure of force or any other change, to transfer as much heat as we please from a *cold* to a *hot* body" (p. 32, p. 134). This result contradicts R7, a principle conceived by Clausius as absolutely fundamental.[15] Having derived this contradiction, Clausius concludes, by means of

[14] I follow the steps of Clausius's argument as it is presented in his 1850 (pp. 31-32, p. 134).

[15] As far as I know, it cannot be determined with certainty whether R7 originally belonged to the (explicit) constructive constraints of Clausius's problem, or, that he 'discovered' it while he was trying to find a new derivation for C5 . In any case, as soon as Clausius became 'aware' of this principle, he assigned the highest preference to it. In a note added in 1864, he even claims that it is as fundamental as the principle that neither heat nor work can be created 'out of nothing' (Clausius, 1864, p. 55). The concept of entropy, which Clausius developed in later work, goes back on R7.

reductio ad absurdum[16], to C5.

The reader will remember from section 4.2 that Clausius is still working with inconsistent constraints. Nevertheless, he applies *reductio ad absurdum*. Why does he do so? Why does he not, as in the failed argument discussed above, decide that this rule of inference leads to an invalid argument? There are two differences. First, in this new argument, Clausius only makes inferences from high-preference-constraints: he relies nowhere upon 'doubtful' results of Carnot's theory. Next, and this is the most important difference, the contradiction concerning R7 is *not* derivable from the premises alone: starting from the *premises*, *without* the hypothesis that C5 or its reformulation is false, one is able to conclude to R7 but *not* to its falsehood. This means that the contradiction is not related to the fact that the constructive constraints are inconsistent but truly reflects the impossibility of the hypothesis. Hence, from the point of view of APIL1, it makes sense to apply *reductio ad absurdum.*

7.2. *Assigning preferences — a solution for the remaining inconsistencies*

As I mentioned in section 4.2, the new derivation of C5 has to be seen against the background of several contradictions: contradictions result, for instance, from C2&J2 and from R5&R6. It is most likely that Clausius at first did not have any (internal or external) ground for preferring some halves of these contradictions above others. C2 as well as R5 follow from C5 (see section 4.2), the latter being conceived by Clausius as one of the central elements for a general theory of thermodynamic phenomena. On the other hand, he regarded C2 as well as R6 as fundamental principles.

This undecided situation changes with the argument discussed in the previous section. The argument can be generalized to account for any engine (whether or not transferring heat between a hot and a cold reservoir). In its generalized form the argument is directly relevant for the decision between R5 and R6. The elimination of this particular contradiction has immediate consequences for the remaining inconsistencies. Let me explain this.

Suppose that R6 holds true, and thus that the work consumed in producing heat is entirely recoverable. This entails that it is possible to design an engine that converts heat into work, *without* heat being transferred to a second reservoir (for if it were not, it would, in some cases, be impossible to recover all the work expended in producing an amount of heat — see sec-

[16] A slightly different reconstruction proceeds in terms of *conditional proof* and *modus tollens*, but has not the slightest effect on the correctness of my analysis.

tion 4.2). Clearly, the work produced by such an engine could be used to operate a Carnot-engine in the inverse sense. The result would be that an amount of heat is transferred, *without* any expenditure of work, from the cold to the hot reservoir. As we have seen, Clausius regards the latter as totally unacceptable.

What about the assignment of preferences to R5 and R6? In the light of the logical analysis of the relevant constraints just described, this does not cause any further difficulties. Indeed, at this moment, Clausius *does* have good grounds for preferring the one above the other. R6 inevitably leads to the conclusion that it is possible to transfer heat from a cold to a hot reservoir, without any work being expended. But, this conclusion contradicts a principle Clausius is *by no means* willing to abandon. Consequently, the adoption of R6 makes it impossible to arrive at a consistent theory. On the other hand, the adoption of R5 does not result in the same kind of difficulty. Therefore, it is certainly reasonable to assign a higher preference to R5 than to R6. This, in turn, has consequences for the contradiction resulting from C2&J2. The adoption of R5 leads to the adoption of C2; the adoption of J2 leads to the rejection of it. As a high preference is assigned to R5, it makes sense to assign a higher preference to C2 han to J2. (Note especially that thus a high preference is assigned to the idea that work may be lost, an idea which Clausius initially considered to be highly implausible, and a low preference to the idea that work is conserved.)

If constraints are eliminated in view of this particular preference-ranking, the resulting set of constraints is consistent. There remains, however, one important question to be answered. Did Clausius's problem have a *unique solution*? Put in other words: would an arbitrary researcher, starting from the same set of constraints and from the same initial preferences, necessarily have arrived at the same solution? These questions have to be answered in the negative. Central in the assignment of preferences is R7 — the principle that it is impossible, without the expenditure of work, to transfer heat from a cold to a hot reservoir. But this principle did, most probably, not belong to the original set of constraints; it was *added* to it while working on the problem. Adding a different principle could have resulted in a solution significantly different from the one Clausius arrived at. If, for instance, the highest preference were assigned to some kind of symmetry principle, then, most probably J1 and J2 would have been retained and C5 would have been rejected. This indicates at once that the order in which the contradictions are resolved, may influence the final solution. Had Clausius concentrated on the contradiction resulting from C2&J2 in an earlier stage of his search process, then, maybe, he would never have attempted to derive C5 (and hence, would never have 'discovered' R7).

All this may sound rather speculative. One should not forget, however, that different researchers, seemingly starting from the same set of constraints and from the same initial preferences, did arrive at very different solutions. Till the end of the nineteenth century, several researchers explicitly rejected the principle which Clausius considered to be fundamental. Seemingly, they assigned the highest preference to the idea that the mutual conversion of heat and work is *perfectly reversible*.

All this cannot possibly be accounted for from the point of view of mixed non-monotonic logics, for the latter presuppose that preferences are fixed beforehand. Moreover, from the point of view of mixed non-monotonic logics, the same result is obtained whenever one starts from the same set of constraints and the same initial preferences. The situation makes perfectly good sense, however, from the point of view of inconsistency-adaptive logics. As inconsistency-adaptive logics do not involve non-logical preferences, there is no objection to use a logic of this type to analyze the relevant constraints, to add further information in the light of this analysis and to make choices on the basis of the present best insights.

8. *Conclusions*

Inconsistency-adaptive logics are a fascinating enterprise. Not only do they have a lot of interesting properties which are worth studying for their own sake, they also offer an excellent tool (actually, the only tool we have at this moment), for a better understanding of creative processes that involve inconsistent constraints. In this paper, I discussed some characteristics of such processes. As it turned out, these characteristics can easily be accounted for by an inconsistency-adaptive logic. There are other applications, however. Inconsistency-adaptive logics have a *dynamic* proof procedure. This seems to capture the *dynamics* in problem solving processes — the discovery of new inconsistencies leads to the refinement or even abandonment of previously accepted findings.

This paper constitutes a first step. Hopefully, others will follow. For one thing is clear. Inconsistency-adaptive logics open new perspectives, not only in the domain of logic, but also, and perhaps more importantly, in the domain of the methodology of science.

Centrum voor Logica en Wetenschapsfilosofie
Universiteit Gent
Joke.Meheus@rug.ac.be

REFERENCES

Batens, Diderik

1985 "Dynamic Dialectical Logics as a tool to deal with and partly eliminate unexpected inconsistencies." in J. Hintikka & F. Vandamme (eds.), *The Logic of Discovery and the Logic of Discourse.* New York, Plenum Press, 1985, pp. 263-271.

1986 "Dialectical dynamics within formal logics." *Logique et Analyse* 114, pp. 161-173.

1989 "Dynamic dialectical logics." in G. Priest, R. Routley & J. Norman (eds.), *Paraconsistent Logic. Essays on the Inconsistent.* München, Philosophia Verlag, 1989, pp. 187-217.

1994 "Inconsistency-Adaptive Logics and the Foundation of Non-Monotonic Logics." *Logique et Analyse* 145, pp. 57-94.

199+b "Inconsistency-Adaptive Logics." in Ewa Orłowska (ed.), *Essays Dedicated to the Memory of Helena Rasiowa.* (forthcoming).

Brush, Stephen

1986 *The Kind of Motion we call Heat. A History of the Kinetic Theory of Gases in the 19th Century. Book 1, Physics and the Atomists.* Amsterdam, North-Holland Physics Publishing.

Carnot, Sadi

1824 *Réflexions sur la Puissance Motrice du Feu et sur les Machines propres à développer cette Puissance.* Paris, Bachelier. Reprinted, 1990, Sceaux, Éditions Jacques Gabay.

Clausius, Rudolf

1850 *Ueber die bewegende Kraft der Wärme und die Gesetze, welche sich daraus für die Wärmelehre selbst ableiten lassen.* Reprinted in M. Planck (ed.), *Ueber die bewegende Kraft der Wärme und die Gesetze, welche sich daraus für die Wärmelehre selbst ableiten lassen.* Leipzig, Verlag von Wilhelm Engelmann, 1898, pp. 1-52.

1863 *Sur un axiome de la théorie mécanique de la chaleur.* Reprinted and translated in F. Folie (ed.), *Théorie mécanique de la chaleur par R. Clausius.* Paris, Librairie scientifique, industrielle et agricole, 1868, pp. 311-335.

1864 *Sur la force motrice de la chaleur et les lois qui s'en déduisent pour la théorie même de la chaleur.* Revised edition, reprinted and translated in F. Folie (ed.), *Théorie mécanique de la chaleur par R. Clausius.* Paris, Librairie scientifique, industrielle et agricole, 1868, pp. 17-106.

Łukasziwicz, Witold

1990 *Non-Monotonic Reasoning. Formalization of Commonsense Reasoning.* New York, Ellis Horwood.

Meheus, Joke

199+ "Clausius's case reconsidered: a retrial of the process." (in preparation).

Mendoza, E. (ed.)

1960 *Reflections on the Motive Power of Fire by Sadi Carnot and other Papers on the Second Law of Thermodynamics by E. Clapeyron and R. Clausius.* New York, Dover Publications.

Nickles, Thomas

1980 "Scientific Discovery and the Future of Philosophy of Science." in Th. Nickles (ed.), *Scientific Discovery, Logic, and Rationality.* Dordrecht, Reidel, 1980, pp. 1-59.

1981 "What is a Problem that we may solve it?" *Synthese* 47, pp. 85-118.

Rescher, Nicholas

1964 *Hypothetical Reasoning.* Amsterdam, North-Holland.

Smith, Crosbie

1990 "Energy." in R.C. Olby, G.N. Cantor, J.R.R Christie, & M.J.S. Hodge (eds.), *Companion to the History of Modern Science.* London, Routledge, 1990, pp. 326-356.

Truesdell

1980 *The Tragicomical History of Thermodynamics 1822-1845.* New York-Heidelberg-Berlin, Springer Verlag.