Studies in History and Philosophy of Modern Physics 66 (2019) 52–61

Contents lists available at ScienceDirect

Studies in History and Philosophy of Modern Physics

journal homepage: www.elsevier.com/locate/shpsb

# The reversibility objection against the Second Law of thermodynamics viewed, and avoided, from a logical point of view

Thomas Müller

*Department of Philosophy, University of Konstanz, Fach 17, 78457, Konstanz, Germany*

ARTICLE INFO

*Article history:*
Received 12 September 2017
Received in revised form
6 September 2018
Accepted 14 September 2018
Available online 27 December 2018

ABSTRACT

In this paper we offer a formal-logical analysis of the famous reversibility objection against the Second Law of thermodynamics. We reconstruct the objection as a deductive argument leading to a contradiction, employing resources of standard quantified modal logic and thereby highlighting explicit and implicit assumptions with respect to possibility, identity, and their interaction. We then describe an alternative framework, case-intensional first order logic, that has greater expressive resources than standard quantified modal logic. We show that in that framework we can account for the role of sortals in possibility judgments. This allows us to formalize the relevant truths involved in the reversibility objection in such a way that no contradiction ensues. We claim that this analysis helps to understand in which way the Second Law is, specifically, a law of thermodynamics, but not of systems of particles in general.



## 1. Introduction

Thermodynamics is a highly successful theory that describes the behavior of systems consisting of a vast number of microconstituents in terms of a few macroscopic variables such as volume, temperature, and entropy. Perhaps the most distinctive feature of thermodynamics as a physical theory is that it incorporates a time-asymmetric law, the so-called Second Law of Thermodynamics, which states that the entropy—a specific state function—of an isolated system does not decrease over time. (Milk poured in coffee mixes, but never unmixes, spontaneously.) This law sets quantitative limits, for example, to the amount of useful energy that can be extracted from a heat engine, and thus defines hard boundary conditions for a multi-billion-dollar industry. The Second Law is certainly technologically fundamental.

Still, there are worries about the physical fundamentality of the Second Law. In a nutshell, the main worry is this. Any system described by thermodynamics—e.g., a gas—is constituted by microparticles that are subject to time-symmetric dynamical laws, e.g., of Newtonian mechanics. Accordingly, these particles' motion can be reversed without infringing their laws of motion. The laws describing the macroscopic behavior of the system, which provide a coarse-grained description of the micro-dynamics, should therefore also be time-symmetric and allow for the reverse of any lawful dynamics as well. According to this so-called reversibility objection, the time-asymmetric Second Law could not be a law after all.

In this paper, we will consider the reversibility objection from a logical point of view, with the aim of holding on to the Second Law as a law for thermodynamic systems. After a brief introduction to the reversibility objection (§2), we will formalize it as a deductive argument in the logical framework of standard quantified modal logic, filling what we take to be an important lacuna in the sizable literature on the matter. We will show how seemingly innocuous assumptions about the dynamics of systems allow one to derive an explicit contradiction (§3). This motivates our move to a different logical framework, case-intensional first order logic (§4). We will show how that alternative logical framework allows one to circumvent the reversibility objection while doing justice to the scientific facts (§5). We end with some general conclusions in §6.

## 2. The reversibility objection

Thermodynamics (TD) is a universal theory: A vast number of actual physical systems, independent of their material constitution and almost independent of scale, are adequately described by the theory. But the class of *possible* physical systems clearly exceeds the class of systems to which TD applies. Consider any actual sample of a gas (e.g., isolate a volume of such a gas in the lab): TD will apply

*E-mail address:* Thomas.Mueller@uni-konstanz.de.

https://doi.org/10.1016/j.shpsb.2018.09.002
1355-2198/

beautifully. Most importantly for our purposes, the sample's entropy will be non-decreasing, i.e., the Second Law will be valid. But it is easy to describe a possibility for that system, an arrangement of the constituents of the gas isolated in the lab, in which TD will not apply. For a vivid example, take the gas in a set-up in which its entropy is increasing dramatically in accord with the Second Law (e.g., by undergoing free expansion). Then instantaneously reverse the direction of motion of all particles making up the gas. Assuming classical Newtonian micro-dynamics, the resulting system will show the temporal mirror image of the original behavior. Thus, its entropy will decrease over time, violating the Second Law.[1] This thought experiment is known as the *reversibility objection* to the universal validity of the Second Law, and hence, of TD.

Historically, the reversibility thought experiment was described by Loschmidt (1876) in a discussion of Boltzmann's attempts to derive the laws of thermodynamics from considerations of micro-dynamics. Loschmidt describes the envisaged status of the Second Law as "a theorem of analytical mechanics" like the First Law (which concerns the conservation of energy), and explicitly states that that theorem needs to be valid "not just under circumstances that are found in nature", but completely generally.[2] The objection would appear to show that this is asking too much. And indeed, today it seems universally agreed upon that the Second Law cannot be a theorem of analytical mechanics: It is not a law describing the possible behavior of systems of particles in general, because there are systems of particles whose possible behavior violates the would-be law. The Second Law is, however, a law of TD, forming an integral part of the successful practice of applying TD in a wide variety of circumstances. What are we to make of this?

A straightforward reaction to this difficulty would be to show that the concrete systems to which TD is actually successfully applied are *different in kind* from the imagined systems referred to in Loschmidt's thought experiment, and that the Second Law naturally applies to the first kind of systems, but not to the second. It appears to be difficult to say, however, what this difference in kind could consist of. Doesn't the reversibility thought experiment show that *the very same systems, consisting of the very same micro-constituents*, can show both thermodynamic and anti-thermodynamic behavior? What is different in these cases, apart from the contingent state that the systems are in? Materially, nothing is different: Actual systems showing thermodynamic behavior do not have any secret parts or special constituents telling them apart from the imagined systems after motion-reversal that show anti-thermodynamic behavior. So the difference between the two kinds of systems, it seems, can really only be a difference within one kind of entities, systems of particles, that points to two different classes of states, one associated with thermodynamic, the other, with anti-thermodynamic behavior. And the fact that naturally occurring systems behave thermodynamically can only mean that in nature, systems of particles are found (almost) exclusively in states belonging to the former class. This, in turn, may be spelled out as a difference in the probability of finding a system of particles in a state belonging to one of the two classes, given certain natural preconditions.

In this way, the task of dealing with the reversibility objection becomes one of explaining (A) why actual systems of particles are (almost) exclusively, or with exceedingly high probability, found, or produced, in states that do not lead to a decrease in entropy over time, and (B) how the micro-dynamics keeps the two classes of states well separated. We agree that this is a worthwhile enterprise. As discussions of issue (A) in the literature show, the enterprise turns out to be quite complicated.[3] The sought-for explanation will most likely have to refer to facts of natural history, such as facts about our universe and its past, and facts about our biological make-up that grounds our powers of discrimination, intervention, and manipulation.[4] Note that conceptually, what an explanation of issues (A) and (B) cannot establish is the status of the Second Law *as a law of a specific kind of entities*. The reason is that such an explanation has to assume that the difference between thermodynamic and anti-thermodynamic behavior is a difference among the possibilities for a single kind of entities, systems of particles. Thus, the explanation fails to support a difference in kinds of entities that could underwrite a difference in laws. After all, the laws for systems of particles have to be compatible with the possibilities for these entities' behavior, and these possibilities include both thermodynamic and anti-thermodynamic behavior. It follows that the Second Law cannot be a law for these entities.

Since this claim may be controversial, we offer replies to two relevant challenges. First, there are attempts to promote the facts on which the explanation of issue (A) is based (such as a low-entropy past of the universe, via the "past hypothesis") to the status of law-like necessities (e.g., Loewer, 2004). In this way, it would be a law that systems of particles can only be found in states giving rise to thermodynamic behavior. The issue is subtle and highly controversial (see, e.g., Winsberg, 2008). Over and above the challenge of Poincaré recurrence (see note 3 above), the problem is how to make sense of motion-reversal as a possibility for systems of particles, given that the resulting anti-thermodynamic behavior runs against the purported law-like necessity of thermodynamic behavior. If for a given system, some behavior is both possible and impossible, then we are either facing a contradiction, or two

[1] In this paper, we assume, in line with the discussion of the reversibility objection, that it makes sense to ascribe compliance to the Second Law, and the associated entropy non-decrease, to a system at an instant. This assumption is, to a certain extent, naive. For example, at small time scales, minute fluctuations of entropy in both directions will be observed. These fluctuations just point to a step of idealization, inherent in the simplest formulation of the Second Law, which is innocent from the point of view of the present discussion. The Newtonian micro-dynamics of a closed system is generally assumed to be deterministic, so that the whole temporal development of the system is given by its state at one instant. See also note 2.

[2] "[D]er zweite Hauptsatz der mechanischen Wärmetheorie [soll], gleich dem ersten ein Satz der analytischen Mechanik sein […], [gültig] nicht allein unter den in der Natur vorkommenden Verhältnissen, sondern ganz allgemein" (Loschmidt, 1876, p. 128f.).—Loschmidt brings forth his famous thought experiment towards the end of his article, in which he is explicitly not objecting to the Second Law, but rather to the claim that a system that is in equilibrium will remain in equilibrium forever. He holds that a criticism of the latter claim does not affect the former. We follow the mainstream discussion after Loschmidt in taking the thought experiment as a challenge to the Second Law itself.—Loschmidt (1876, 139) is explicit that he considers an instantaneous ("momentan") reversal of instantaneous velocities ("die momentanen Geschwindigkeiten […] plötzlich umzukehren").

[3] In light of Poincaré's famous recurrence theorem, explaining fact (B) in the sense of marking a mathematically strict, eternally stable distinction between the two classes of states is probably impossible. For now, we follow the common practice of ignoring issue (B) in the context of discussions of issue (A) that refer to natural history. This attitude, one might say, indicates a pragmatic approach to the persistence of thermodynamic systems over very long periods of time. We will comment on Poincaré recurrence further in note 16 below.

[4] We do not claim that it has been shown (and we will not attempt to show) that there can be no more fundamental explanation of the ubiquity of TD systems. Maybe the reversibility story can be dismissed because it ignores important details of actual micro-dynamics that already lead to irreversibility. Ismael (2016) provides a helpful critical evaluation of two pertinent attempts in this direction, one due to Albert (2000) and one due to Leeds (2003). And maybe the very notion of separation and isolation of a system allows for a formal or conceptual derivation of the TD nature of that system—perhaps along the lines of Reichenbach's (1956) theory of branch systems. (Note, however, that Reichenbach himself labels his assumptions about the lattice of branch systems not as conceptual truths, but as "empirical hypotheses which are convincingly verified" (Reichenbach, 1956, p. 136).) We will remain neutral with respect to these options. Indeed, the logical framework that we will describe in §4 is general enough to accommodate them as well.

different senses of "possible" are involved. Neither option will help to support the status of the Second Law as a law.[5]

Second, one may doubt that a law such as the Second Law has to be bound to a kind of entities, and dismiss this assumption as overly essentialistic. Certainly there are diverging metaphysical views about laws of nature, with different implications for the role of kinds of entities in grounding laws. However, a discussion of the status of the Second Law only makes sense before the background of a metaphysics that allows for laws of the type to which the Second Law purports to belong—and the Second Law is explicitly about thermodynamic systems. The challenge to the Second Law that comes with the reversibility objection has got nothing to do with a critique of essentialism. Rather, it is built on the assumption of the existence of laws of exactly the type that the Second Law, if successful, would instantiate.

The strategy of merely explaining issues (A) and (B), while undoubtedly worthwhile, thus amounts to giving up on the status of the Second Law as a law. The aim of the present paper is to try and find a way to *hold on to the Second Law as a law for a specific kind of systems*—thermodynamic systems—while, at the same time, acknowledging that these systems are nothing but systems of particles, and that for systems of particles, motion reversal is a physical possibility. In the context of this project, an explanation of issues (A) and (B) then provides a natural-historical account of why systems of the specific kind in question are so widespread—why there are so many TD systems around. That is helpful. But the modal difference between thermodynamic systems and mere systems of particles, which has to be presupposed for a difference in laws, has to be explained in a different way. In order to find that way, we will analyze the reversibility objection from a formal logical point of view, asking which assumptions about systems, their identity over time, and what is possible for them, have to be made in order to formulate the objection. This will allow us to asses the objection's logical force, and will provide hints as to how the Second Law can retain its special status in a formally coherent way.

## 3. The reversibility objection formalized

The crucial move in pressing the reversibility objection is to infer from actual, entropy-increasing behavior of a system to the possibility of entropy-decreasing behavior of the very same system. It is, therefore, crucial to spell out the objection with explicit recourse to assumptions about possibilities and identity. A formal presentation of the objection thus has to employ the framework of quantified modal logic, which allows one to talk about different objects and their possibilities.

Before we start to formalize, we give an informal summary of the argument. Thus, assume that the Second Law is a law of TD, and take a TD system, *s*, subject to Newtonian micro-dynamics. When isolated, *s* is in one of its possible states. Since *s* is a TD system, we know that state must give rise to a time evolution of the system in the course of which the entropy of *s* will not decrease over time. In order to have anything to prove, we will assume that the state of *s* is such that its entropy is actually increasing over time. Now we know that for *s*, which is a system of particles, it is possible to be in any of the states in its phase space—the phase space just *is* the set of all possible dynamic states of the system of particles. The operation of motion reversal is an involution in the system's phase space: For every possible state of *s* there is a motion-reversed state (and that state's motion-reversal is the state we started with). Therefore, for *s* it is possible to be in the state that is the motion-reverse of its actual state. That state will, however, be such that the system's entropy will be decreasing over time, which means that the system *s* will evolve anti-thermodynamically, in violation of the Second Law. And so, the Second Law is not a law that holds of *s* after all, since behavior contrary to what is described by the law is possible for *s*. Since *s* is a TD system, it follows that the Second Law cannot be a law of TD.

We take it that this argument is convincing, and that it is based on at least prima facie plausible assumptions. But its conclusion seems too hard to swallow, given the Second Law's generality, high degree of empirical support, and technological and conceptual fundamentality. And in fact, the status of the Second Law as a law is not given up.[6] So the reversibility argument stands as a conceptual challenge.

In what follows, we provide a formal reconstruction of the reversibility argument, using standard resources of quantified modal logic. We will proceed in a way that is neutral with respect to many of the fine differences between available logical frameworks. Given the overall aim of this paper, which is to provide a formally coherent account of the Second Law *as a law of TD*, note that the following formalization will ultimately be rejected. Its dialectical import here is two-fold. First, we underscore the common assessment that the reversibility objection is a strong argument: in our formalization we will derive a contradiction using plausible premises and valid formal reasoning steps, exactly in line with the informal argument given above. Second, formalizing forces us to make explicit assumptions about the modal reasoning involved, thereby offering hints for a representation that could retain the status of the Second Law as a law of TD.

Quantified modal logic combines resources for expressing possibility and necessity with resources for quantifying over objects. Syntactically, this means that, first, we can use variables, identity, the quantifiers ("$\forall x$", for all $x$, and "$\exists x$", there is an $x$), etc., as in standard predicate logic. Second, there has to be an operator $\Diamond$ to stand for "it is physically possible that", and its dual, the operator $\Box$, for "it is physically necessary that". Duality here means that we can define "$\Diamond\phi$" as "$\neg\Box\neg\phi$": $\phi$ is physically possible iff (if and only if) it is not physically necessary that non-$\phi$ holds.

Semantically, there are many choices to interpret formulae of quantified modal logic. Following standard practice, here we read the modal operators as quantifying over physically possible worlds $w \in W$, where the actual world is denoted $w_@$. A formula $\Diamond\phi$ is true at a world $w$ iff $\phi$ is true at some world accessible from $w$, and dually, $\Box\phi$ is true at $w$ iff $\phi$ is true at all accessible worlds. The argument we will give is valid already in the minimal normal modal logic with the characteristic axiom K, $\Box(p \to q) \to (\Box p \to \Box q)$, which does not constrain the accessibility relation among the physically possible worlds.[7] Variables are assumed to range over systems of particles—but here we need to be more precise, because

[5] It is true that anti-thermodynamic behavior of systems of particles, while physically possible, is rarely observed and thus untypical. Explaining issue (A) will help to clarify and perhaps even quantify the relevant sense of "untypical". But the commonly accepted status of the Second Law is *not* just one of typicality. That law appears to be different in kind from, e.g., laws about biological species, such as "horses have four legs", which refer to typical (or healthy, well-formed) individuals. So we do not think that playing the "two senses of possibility" card in this simple way will help. Our own solution in §5 will differentiate the incompatible possibilities by explicit reference to the sortal predicates "… is a system of particles" and "… is a TD system".

[6] Famously, Einstein said about TD: "It is the only physical theory of universal content concerning which I am convinced that, within the framework of applicability of its basic concepts, it will never be overthrown" (Einstein, 1949, p. 32).

[7] The argument therefore remains valid if the accessibility relation is constrained, e.g., to be an equivalence relation as in the modal logic S5, which is plausible for physical possibilities. Garson (2013) provides a useful overview of modal logical systems.—A variant of our argument, with a strengthened premise (8), is valid in the modal logic T, and therefore, also in S5. See note 11 for details.

we need to say in which way possibilities for objects are expressed, and how the domains of discourse at different possible worlds interrelate. Again following standard practice, we assume that all possible worlds have the same domain of discourse ("constant domain"). Since the physically possible worlds are to represent the possibilities for any given system of particles, the denotation of a variable (or of any other term) at a world is plausibly taken to be a point $(Q, P)$ in the system's phase space $S$.[8] Such a point specifies both the positions $(Q)$ and the momenta $(P)$ of all the particles making up the system, and therefore allows one to classify a possibility for the system as entropy-increasing, entropy-decreasing, or neither.

We will use the following mnemonic predicates: $TD(x)$ for "$x$ is a thermodynamic system", $SP(x)$ for "$x$ is a system of particles", $SL(x)$ for "the Second Law holds of $x$", $MR(x)$ for "the system $x$ is in the motion-reversed state relative to its actual state", $EI(x)$ for "the entropy of $x$ is increasing", and $ED(x)$ for "the entropy of $x$ is decreasing".[9]

Our formalization of the argument starts with two factual premises about our system, $s$, which are assumed to hold in the actual world. We make the world-dependence of premises explicit by prefixing them with "$w_@ \vDash \ldots$", which means "the actual world satisfies the formula …".

(1) $w_@ \vDash TD(s)$

(2) $w_@ \vDash EI(s)$

Thus, we assume that our system $s$ is a thermodynamic system whose entropy is actually increasing. (As we said, if $s$'s entropy were constant, there would be nothing to prove.) Now we need to lay down the general principles invoked in the argument, which we formulate as follows:

(3) $w_@ \vDash \Box\, \forall x\, [SL(x) \rightarrow \neg ED(x)]$

(4) $w_@ \vDash \forall x\, [TD(x) \rightarrow \Box SL(x)]$

(5) $w_@ \vDash \forall x\, [SP(x) \rightarrow \Diamond MR(x)]$

(6) $w_@ \vDash \forall x\, [EI(x) \rightarrow \Box [MR(x) \rightarrow ED(x)]]$

(7) $w_@ \vDash \forall x\, [ED(x) \rightarrow \Box [MR(x) \rightarrow EI(x)]]$

(8) $w_@ \vDash \forall x\, [TD(x) \rightarrow SP(x)]$

These principles say that (3) necessarily, if the Second Law holds for a system, then that system's entropy is non-decreasing, that (4) for a TD system, the Second Law holds as a law, with necessity, and that (5) any system of particles could be in the motion-reverse of its actual state. Then, (6) and (7) state that the operation of motion reversal also reverses the direction of change of entropy over time: if a system's entropy is actually increasing, then necessarily, if the system is in the motion reverse of its actual state, its entropy is decreasing (6), and vice versa (7). Finally, principle (8) says that actually, all thermodynamic systems are systems of particles.

For clarity's sake, we comment on the use of modal operators in these premises explicitly. In (3), the necessity operator ranges over the whole quantified conditional: In any possible world, for any system, if the Second Law holds (in that world), then that system's entropy is non-decreasing (in that world). This just gives the content of the Second Law as ruling out entropy-decreasing behavior. In (4), which expresses the status of the Second Law as a law, the necessity attaches to the consequent only: we say that if a system $x$ is actually a thermodynamic system, then in any possible world, the Second Law holds of $x$—i.e., the Second Law is a law and thus holds in all physically possible worlds, it is not just a contingent truth that only holds in some worlds. In (5), the possibility also attaches to the consequent only: for any system $x$ that is actually a system of particles, it is possible that that system is in the motion reverse of its actual state. Thus, there is a possible world at which the state of $x$ is motion-reversed with respect to $x$'s state at the actual world. This expresses a plain truth about the phase space of systems of particles that summarizes the possibilities for these systems, as we said: the phase space of a system of particles is closed under motion reversal.[10] And in (6, 7), the necessity attaches to the inner conditional to allow for the following reading: if a system's entropy is actually increasing, then in any possible world in which that system's actual motion is reversed, the entropy is decreasing (and vice versa). The physical basis for these claims lies in the assumed Newtonian micro-dynamics. No modality is involved in (8).[11]

Formulae (3)–(8) thus capture plausible assumptions about TD systems. But as in the informal argument with which we started, the assumptions cannot be innocent, because we can derive a contradiction from them. Here is how:

(9) $w_@ \vDash \Box SL(s)$ 1, 4

(10) $w_@ \vDash \Box [SL(s) \rightarrow \neg ED(s)]$ 3

(11) $w_@ \vDash \Box \neg ED(s)$ 10, 9, modal logic

(12) $w_@ \vDash \Box [MR(s) \rightarrow ED(s)]$ 2, 6

[8] Assuming conservation of energy (the First Law mentioned above), the physical possibilities will be constrained to the system's orbit in phase space, a proper subset of $S$. We will not need to make this consideration explicit for our formalization, and therefore leave it aside.

[9] In line with what we said in note 1 above, we read the latter two predicates in such a way that, e.g., small fluctuations of entropy are unproblematic.

[10] In line with the informal statement of the reversibility objection, we need to formulate the possibility of motion reversal as a possibility *for the system under discussion*, $x$. Here we have opted for the simplest formulation in terms of a predicate $MR(x)$, which predicates motion reversal relative to $x$'s *actual* state. Thus, if $MR(x)$ were to hold at the actual world $w_@$, then the actual state of $x$ (at $w_@$) would have to be stationary. But for $\Diamond MR(x)$ to hold at the actual world, $MR(x)$ may hold at a different world $w \neq w_@$, and then the actual state of $x$ is not constrained: all that is said is that at that world $w$, system $x$ is in the motion-reverse of its state at $w_@$.—We add two pertinent notes. (A) It is possible to set up our argument with a function that maps states to their motion-reverse, and with a relation attributing states to systems. As this would require significant additional formal machinery without providing additional benefits, we stick to our simpler formalization. (B) One may also think of motion reversal in terms of an intervention, so that "$\Diamond MR(x)$" says that it is possible to intervene on $x$ in such a way that the actual motion of all constituent particles is reversed. This reading is innocent as long as one reads the possibility involved here as physical, not as technological possibility. Of course it is not claimed that we can actually intervene on any system in such a way that the motion of all of its constituent particles is reversed. That is technologically infeasible for most systems—but there are exceptions. Spin echoes are often claimed to show that in some spin systems, we can actually intervene such as to effect motion reversal. The original spin echo experiment was conducted by Hahn (1950); see, e.g., Ridderbos and Redhead (1998) for discussion.

[11] It is possible to strengthen (8) to have law-like status, via a "necessarily" prefix as in (3): $\Box\, \forall x\, [TD(x) \rightarrow SP(x)]$. Our argument then remains valid, given that the modal axiom $T$ holds, which licenses the inference "ab necesse ad esse", from $\Box \phi$ to infer $\phi$. That axiom certainly holds for our interpretation of $\Box$ as physical necessity: What is physically necessary is also actually true. Semantically, the $T$ axiom holds if the accessibility relation is reflexive, which it certainly is for physical necessity: any world is physically possible relative to itself.

| | | |
|---|---|---|
| (13) | $w_@ \vDash SP(s)$ | 1, 8 |
| (14) | $w_@ \vDash \Diamond MR(s)$ | 5, 13 |
| (15) | $w_@ \vDash \Diamond ED(s)$ | 12, 14, modal logic |
| (16) | $w_@ \vDash \neg\Box\neg ED(s)$ | definition of "$\Diamond$" |
| (17) | contradiction | 11, 16 |

So we have derived that it is both necessary that $s$'s entropy be non-decreasing (11) and that it is not the case that it is necessary that $s$'s entropy be non-decreasing (16). This is a straightforward contradiction of the form "$p$ and not-$p$". What are we to make of this? At a first glance, we seem to be stuck. The factual claims (1) and (2) are what we start with, they cannot be denied. (There actually *are* thermodynamic systems whose entropy is increasing; when in doubt, pour some milk in your coffee.) And our general principles (3)–(8) are also plausible. Furthermore, the two steps of modal reasoning employed above reflect really minimal assumptions about the underlying logic: In the derivation of (11), we use an inference of the form, "from $\Box(p \to q)$, derive $(\Box p \to \Box q)$", and in the derivation of (15), we use an inference of the form, "from $\Box(p \to q)$, derive $(\Diamond p \to \Diamond q)$". Both inferences are valid already in the weakest normal modal logic, K. (In fact, the first just captures the characteristic K axiom of modal logic, and the second follows directly from that axiom.)

We are facing a contradiction arrived at through valid reasoning, and something has to go. In the informal version of the argument, the blame was on the status of the Second Law as a law, but the formalization offers more options—including the option to reject the formalization. In fact we will show that an alternative, richer framework of quantified modal logic, case-intensional first order logic (CIFOL), allows one to formalize the premises of the informal argument in a way that does not allow for the derivation of a contradiction. In that way, the status of the Second Law as a law can be retained. We will present our resolution of the reversibility problem, via a precisification of premise (8), in §5. But first we need to sketch the alternative framework.

## 4. Case-intensional first order logic

In this section we describe a logical framework that offers the resources for properly representing what is right about the reversibility objection without leading to inconsistency: case-intensional first order logic (CIFOL). We do not claim that this framework is necessary for making sense of Loschmidt's thought experiment, only that it is a useful framework for doing so. A detailed exposition of CIFOL can be found in Belnap and Müller (2014a, b).[12] The three main features important in the present context are that CIFOL allows for the explicit formal representation of the possible temporal developments of individual objects, that it allows for the differentiation of plain identity and necessary identity, and that it allows for a perspicuous representation of sortal predication that can anchor modal differences.

[12] The main background for CIFOL is Bressan (1972). Bressan, a physicist, was in turn influenced by Carnap's ideas about intensional logic, taking a lead from (and improving upon) Carnap's "method of intension and extension", which was propounded in Carnap (1947) and criticized by Quine already as guest author in that very book (p. 196f.). Bressan developed a higher-order case-intensional modal logic in order to tackle issues in the foundations of classical mechanics. The simpler first-order system CIFOL employed here, due to Belnap and Müller (2014a), provides all the resources we need.

**Table 1**
A simple example illustrating intensions and extensions.

| Term \ Case | $\gamma_0$ | $\gamma_1$ | $\gamma_2^1$ | $\gamma_2^2$ |
|---|---|---|---|---|
| $s$ | $a$ | $b$ | $c$ | $*$ |
| $s'$ | $a$ | $b$ | $c$ | $\overline{a}$ |

In this section we provide a brief, informal sketch of CIFOL as a background for our analysis in §5. A more formal overview of the semantics is given in Appendix A.

CIFOL starts with two basic notions: A set, $\Gamma$, of *possible cases*, and a set, $D$, of so-called *extensions* for singular terms. An object, such as a system of particles that may or may not be subject to the laws of thermodynamics, is represented in CIFOL not via an extension (as is customary), but via an *intension*, which is a function specifying for each possible case $\gamma \in \Gamma$ a case-relative extension $d \in D$. Two numerically different intensions can still coincide (specify the same extension) in a given case, which is to say that identity of extension (in that one case) does not have to imply identity of intension (which is identity of extensions across all cases). An *extensional* context is one in which extensional identity licenses substitution, while in an *intensional* context, substitution is allowed only given identity of intensions. In CIFOL, predication need not be extensional, i.e., whether a predication $P(\alpha)$ is true in a case $\gamma$ may depend not only on the extension of term $\alpha$ in case $\gamma$, but on the full intension of term $\alpha$. Sortal predication applies to objects represented via intensions, and is properly intensional: whether a term falls under a sortal, e.g., "… is a thermodynamic system", in a given case is *not* settled by that term's extension in that case alone, and therefore, substitution in a sortal predication requires intensional, not just extensional identity. This behavior allows sortal predication to take into account not just what is so about an object in a specific case, but also what may be so about an object in different possible cases, e.g., how the object can develop over time.

With a view to our application of CIFOL to the reversibility objection in §5, one may think of the cases in $\Gamma$ as possible moments in time, and of the extensions in $D$ as possible momentary states of a system of particles. An intension will then specify a system via a function from possible moments to possible system states, i.e., via a suitable collection of possible temporal developments of the state of a system. And sortal predication will depend on such intensions, not just on extensions, as we will show in §5.

Formally, a term of the CIFOL language—such as a variable $x$ or a constant $s$—has as its intension a function

$$int(s) : \Gamma \mapsto D$$

and has as its extension in case $\gamma$ the value

$$ext_\gamma(s) = [int(s)](\gamma) \in D$$

that the intension-function $int(s)$ assigns for the argument $\gamma$.

A simple illustration, which will resurface in §5, is given in Table 1. We consider just four cases, $\Gamma = \{\gamma_0, \gamma_1, \gamma_2^1, \gamma_2^2\}$, and a domain of extensions $D = \{a, b, c, \overline{a}, *\}$ that contains four regular extensions and the special extension $*$ that signifies non-existence in a case. The two terms, $s$ and $s'$, have the pattern of case-relative extensions and thus, the intensions, specified in Table 1. To abbreviate, we write intensions as strings of extensions in the natural order, e.g., $int(s') = abc\bar{a}$.[13]

CIFOL semantics takes the notion of *extensional identity* to be basic. Thus, the basic semantic clause for identity statements says that a sentence of the form $\alpha = \beta$ is true in a case $\gamma$ iff $ext_\gamma(\alpha) = ext_\gamma(\beta)$. Identity at a case is thus identity of extensions in that case. For example, $s = s'$ holds in case $\gamma_1$, as $ext_{\gamma_1}(s) = b = ext_{\gamma_1}(s')$, but in case $\gamma_2^2$, we have $s \neq s'$, because $ext_{\gamma_2^2}(s) = * \neq \bar{a} = ext_{\gamma_2^2}(s')$. This extensional non-identity in one case witnesses the *intensional* non-identity of the terms; we have $abc* = int(s) \neq int(s') = abc\bar{a}$. To express this fact via a formula in the CIFOL object language, we need the modal operator $\Box$ ("necessarily", or "in any case"). CIFOL defines that operator via quantification over all cases, much like in the standard modal logic used above: $\Box\phi$ is true at a case $\gamma$ iff $\phi$ is true at all cases.[14] Given that way to express necessity, intensional identity, which is extensional identity across all cases, can then be formalized as $\Box(\alpha = \beta)$. In our example, we have, in any case, $\neg\Box(s = s')$. Thus, we have a clear and logically well-founded distinction between plain (case-relative, extensional) identity and necessary (intensional) identity. Accordingly, we can make precise the two mentioned types of contexts for substitution: In an extensional context, substitution is allowed given an extensional identity $\alpha = \beta$, while in an intensional context, only intensional identity $\Box(\alpha = \beta)$ licenses substitution.

Intensions also play a fundamental role in CIFOL's semantics for predication, which is intensional: whether a predicate $F$ applies to a term in a case, may depend not just on that term's extension in that case, but also on that term's intension. That is, in CIFOL in general, predication creates an intensional context requiring necessary identity to license substitution. It is therefore consistent for the following formula to be true in some case:

$$s = s' \wedge \neg\Box(s = s') \wedge F(s) \wedge \neg F(s').$$

Sortal predication is a special case of intensional predication for which that behavior is quite typical. A sortal predicate applies to a term $s$ iff the intension function $int(s)$ satisfies the modal profile of the sortal in question.[15] Typical qualitative predicates are, however, extensional. For example, the number of particles in a system in a given case depends only on what is so in that single case. Accordingly, for the predicate "… consists of $10^{23}$ particles", substitution is licensed already by extensional identity. If that predicates holds of $s$ in a case, and we have $s = s'$ in that case, then the predicate also holds of $s'$ in that case. On the other hand, the predicate "… consists of a constant number of particles" is intensional, as the possible time course of development of the system needs to be considered, and so, extensional identity will not allow for substitution.

The CIFOL treatment of identity and sortal predication offers resources that allow us to avoid the contradiction derived in §3, as we will now show.

[13] Note that the single letters "$a$", "$b$", etc., are not terms, and do not stand for things. They stand for extensions, which in our application will be momentary system states. A system, which the terms of our language such as "$s$" and "$s'$" stand for, is represented not by an extension, but by an intension, i.e., by a function from cases (possible moments) to extensions (possible system states).

[14] Given these semantics, CIFOL does not require an accessibility relation for the basic modal operator $\Box$, and the resulting modal logic for $\Box$ is S5. Additional relational modalities with different properties can be introduced as well, but these are not needed for the purposes of this paper.

[15] Again, more formal details are provided in Appendix A.

## 5. The contradiction avoided

In a nutshell, we can express the challenge of the reversibility objection in the following way: The motion reversal of the entropy-increasing behavior of a thermodynamic system must be both possible (because the motion-reverse behavior is physically possible for any system of particles) and impossible (because the motion-reverse behavior leads to entropy decrease, which violates the Second Law). This is a contradiction. The only thing that could yield, it seems, is the Second Law. But we are not comfortable giving it up.

The framework of CIFOL just laid out, with its resources for differentiating plain, extensional from necessary, intensional identity, and with its definition of a special class of sortal predicates, allows for a different analysis that avoids the contradiction. The key point is to clarify the interrelation between the two sortal predicates that are involved, being a thermodynamic system (*TD*) and being a system of particles (*SP*). The contradiction in what is possible for a given system can be avoided if the conflicting possibilities can be suitably linked to these sortal predicates. The main challenge is to acknowledge that the possibility of motion reversal is a possibility for systems of particles but not for thermodynamic systems—while at the same time holding on to the plain truth that thermodynamic systems *just are* systems of particles. Here is a biological example of the same combination of identity with a difference in possibilities: A cat is a system of particles. (Where a cat is, there is, at any moment, just a system of particles.) But while it is possible for such a system of particles to continue to exist for 100 years, this is not possible for the cat. When two sortals apply, the possibility or impossibility of certain developments or changes can depend on the sortal in question. CIFOL allows us to make formal sense of this idea.

Our formalization relies on the predicates used above in §3: the sortal predicates $TD(x)$ ("$x$ is a thermodynamic system") and $SP(x)$ ("$x$ is a system of particles") and the further predicates $SL(x)$ ("the Second Law holds of $x$"), $EI(x)$ ("the entropy of $x$ is increasing"), $ED(x)$ ("the entropy of $x$ is decreasing"), and $MR(x)$ ("the system $x$ is in the motion-reverse state relative to its actual state"). We use the constant $s$ for the thermodynamic system under discussion, as above. We cannot, however, retain the above formulation of premises (1)–(8) from §3 exactly, because CIFOL allows for the same modal reasoning as above, and would still allow us to derive a contradiction. But we have additional resources at hand: We can be more precise about the interrelation of the sortal predicates in our formalization of the fact that thermodynamic systems *are just* systems of particles. Here is our CIFOL formalization, in which the formulae need to be satisfied in the actual case, $\gamma_@$:

(1') $\gamma_@ \models TD(s)$

(2') $\gamma_@ \models EI(s)$

(3') $\gamma_@ \models \Box\,\forall x\,[SL(x) \rightarrow \neg ED(x)]$

(4') $\gamma_@ \models \forall x\,[TD(x) \rightarrow \Box SL(x)]$

(5') $\gamma_@ \models \forall x\,[SP(x) \rightarrow \Diamond MR(x)]$

(6') $\gamma_@ \models \forall x\,[EI(x) \rightarrow \Box[MR(x) \rightarrow ED(x)]]$

(7') $\gamma_@ \models \forall x\,[ED(x) \rightarrow \Box[MR(x) \rightarrow EI(x)]]$

(8') $\gamma_@ \models \Box\,\forall x\,[TD(x) \rightarrow \exists y\,[SP(y) \wedge x = y]]$

Apart from the change from worlds $w$ to cases $\gamma$, the only change with respect to §3 is in the formulation of the link between the two sortals that are involved, premise (8'). Premises (1')–(7') remain as above. The change to premise (8') is crucial. The old premise (8) said that for any system $x$, if the sortal $TD(x)$ applies, then the sortal $SP(x)$ applies as well. It followed that all $SP$-possibilities have to be $TD$-possibilities, and this way lay trouble, because motion reversal, which is an $SP$ possibility, is impossible for systems for which the Second Law holds. The new premise (8') is formulated in a more cautious way: it only says that in any case, if $x$ is a thermodynamic system, then there is a system of particles $y$ that is, *in that case*, identical to the thermodynamic system. Extensionally, thermodynamic systems *just are* systems of particles. This is plainly true: in any case, if you look at where a thermodynamic system is, all you find is a system of particles. But the resources of CIFOL allow for an *intensional* difference between thermodynamic systems and systems of particles. In this way, not everything that is possible for systems of particles has to be possible for thermodynamic systems, and the ascription of contradictory possibilities can be avoided.

We will go on to provide an explicit model that satisfies all the above premises (1')–(8'). If we are able to provide a model satisfying all the formulae, they have to be consistent, i.e., it can no longer be possible to derive a contradiction. And indeed, if we look at the proof from §3, it turns out that there is one step that is now blocked: we are not licensed to infer $SP(s)$ (line 13 of that proof), which allowed us to infer $\Diamond MR(s)$ (line 14), which in turn was crucial in deriving the contradiction. The extensional identity of TD systems and systems of particles expressed through (8') does not give us intensional identity, and is therefore not sufficient to infer $SP(s)$ from premise (1'), $TD(s)$. We would, however, need $SP(s)$ in order to invoke premise (5') to derive $\Diamond MR(s)$. The inference fails because we have clarified the interrelation of the two sortal predicates, using the additional expressive resources of CIFOL.

In a way, it is not surprising that the inference from §3 can be blocked. We said that there was widespread agreement now that the Second Law is not a law of analytic mechanics, i.e., it is not a law of systems of particles generally, exactly because motion reversal is a possibility for systems of particles generally. The question is just how to square this fact with the undeniable fact that thermodynamic systems, which are subject to the Second Law, just are systems of particles. Our explicit semantic model answers this question by supplying full formal details.[16]

In order to specify a CIFOL model, we have to spell out the set of cases, Γ, and the set of extensions, $D$, as well as the interpretation of the predicates and constants that we use. As we are considering possibilities for the time development of systems (such as the decrease or increase of entropy over time), the set of cases has to be temporal-modal in some sense. A CIFOL theory of temporal-modal

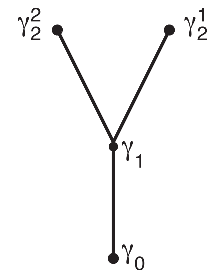


**Fig. 1.** A discrete branching structure with four moments. Time increases in the upward direction; branching indicates alternative possibilities.

cases is laid out in detail in Belnap and Müller (2014b), but that theory is quite complex, and not all of it is needed here. In this paper we will work with a simplified set of temporal-modal cases represented by a discrete, finite branching structure of moments:[17] We posit a set of four cases: $\gamma_0$, which we take to be the actual case, $\gamma_@ = \gamma_0$; $\gamma_1$, the next case after $\gamma_0$; and $\gamma_2^1$ and $\gamma_2^2$, the two possible next cases after $\gamma_1$. The Y-shaped structure of cases is pictured in Fig. 1.

The two branches above $\gamma_1$ are to represent two physical possibilities for the system of particles that we will model: one in which the system follows its Newtonian micro-dynamics undisturbed, and one in which motion reversal is applied.[18] For example, a system of particles that starts in some state $(Q_0, P_0)$ at moment $\gamma_0$ will evolve to state $(Q_1, P_1)$ at moment $\gamma_1$, and can then either evolve to state $(Q_2^1, P_2^1)$ at moment $\gamma_2^1$, or have its motion reversed, which results in state $(Q_2^2, P_2^2) = (Q_0, \overline{P_0})$, the motion reverse of the actual state $(Q_0, P_0)$ at $\gamma_0$, at moment $\gamma_2^2$. All these states have to be available as extensions of terms at these moments. In fact, it makes sense to assume that the whole phase space of the system in question is included in the set of extensions, $D$. For the simple finite model that we provide here in order to prove consistency of our premises (1')–(8'), however, we can do with a much smaller set. We will use single-letter abbreviations for the states mentioned above in order to shorten the exposition:[19] $a$ for $(Q_0,P_0)$, $b$ for $(Q_1,P_1)$, $c$ for $(Q_2^1,P_2^1)$, and $\overline{a}$ for $(Q_2^2,P_2^2) = (Q_0,\overline{P_0})$. As stated in §4, $D$ also needs to contain the $*$-extension as a mark of non-existence.[20] So here is our definition of Γ and of $D$:

$$\Gamma =_{df} \left\{\gamma_0, \gamma_1, \gamma_2^1, \gamma_2^2\right\}; \qquad D =_{df} \{a, b, c, \overline{a}, *\}.$$

Next we need to specify the interpretation of the terms and predicates. We will have two constant terms, $s$ for our thermodynamic system, and an additional constant term, $s'$, for the system of

[16] Here the Poincaré recurrence theorem, mentioned in note 3 above, is pertinent again: For Newtonian systems one can prove from rather weak assumptions that given any system state at a time $t_0$, under the micro-dynamical laws of evolution the system will sooner or later return (almost exactly) to the state it was in at $t_0$. Now, doesn't this imply that any epoch of thermodynamical entropy-increase will later on have to be followed by an epoch of entropy-decrease, such that any system whatsoever will sooner or later have to violate the Second Law, meaning that *no* system can fall under the predicate, *TD*? Our answer to this charge is simple: No isolated system of non-trivial size exists for times long enough to make Poincaré recurrence practically relevant. Thermodynamics is a theory for real systems. Thermodynamic systems are, as Schrödinger (1950) and Reichenbach (1956) have pointed out, "branch systems": They are isolated from the rest of the universe only for some limited time and then recombine with their environment again. The reversibility objection is crucial because it could also pertain to real systems (see the end of note 10 above), so we need to find a way to meet that objection. Considerations of Poincaré recurrence, on the other hand, pertain to idealized, absolutely isolated systems that, by assumption, we could not know anything of. And even very small systems would have to exist for times much longer than the age of the universe for Poincaré recurrence to play any role. Semantically, we can meet the problem by asuming that TD systems cease to exist after some long but finite time.

[17] The framework of Belnap and Müller (2014b) uses Ockhamist cases, which are needed to discuss future contingents in full detail. They also employ continuous, unending branching time models. The framework that we use here is much simpler, but it still reflects the main characteristics of branching models that are relevant for our discussion in this paper.

[18] As mentioned in note 10 above, we do not assume that motion reversal is technologically possible. We just picture the physically possible time developments of a system as presupposed in Loschmidt's thought experiment, and thus, in the discussion of the reversibility objection.

[19] Note again that these letters are not terms of the language; see note 13 above.

[20] In our handling of non-existence, we deviate from Belnap and Müller (2014b) in order to simplify the exposition. The issue is whether to count an intension as falling under a sortal in a case in which the respective extension is $*$. Here, in constrast to the general theory of Belnap and Müller (2014b), we answer that question in the negative. See the " – " entry in Table 2 for $TD/\gamma_2^2$ and our definition of *MConst* in Appendix A.

particles that constitutes $s$ in case $\gamma_0$. The intensions of these terms will be as in Table 1 of §4:

$$int(s) = abc*; \qquad int(s') = abc\overline{a}.$$

Here, as above in §4, we write an intension, which specifies a system via a function that assigns case-relative extensions to terms, as a string of extensions, one per possible case, in their natural order. Thus, we have that $ext_{\gamma_1}(s) = ext_{\gamma_1}(s') = b$, and $ext_{\gamma_2^2}(s') = \overline{a}$. As to identity of terms, extensional and intensional, the following holds, as one can read off from the intensions $int(s)$ and $int(s')$:

$$\gamma_0 \vDash s = s'; \quad \gamma_1 \vDash s = s'; \quad \gamma_2^1 \vDash s = s'; \quad \gamma_2^2 \vDash \neg(s = s');$$

$$\gamma_0 \vDash \neg\Box(s = s'); \quad \gamma_1 \vDash \neg\Box(s = s'); \quad \gamma_2^1 \vDash \neg\Box(s = s'); \quad \gamma_2^2 \vDash \neg\Box(s = s').$$

The extensions of the terms $s$ and $s'$ differ in case $\gamma_2^2$, and thereby, $s$ and $s'$ are intensionally distinct. In case $\gamma_2^2$, the system of particles $s'$ has the non-trivial extension $\overline{a}$—its state in that case is the motion-reverse of its actual initial state in case $\gamma_0$, which was signified by the extension $a$. The thermodynamic system $s$, by contrast, has the extension $^*$ in case $\gamma_2^2$, which signals non-existence. This represents the fact that in case $\gamma_2^2$, *there is no thermodynamic system any more*. Motion reversal, we may say, has destroyed the thermodynamic system, leaving only a system of particles behind. This way of looking at Loschmidt's thought experiment is made possible by CIFOL's explicit treatment of sortal predicates.

It remains to specify the interpretation of the sortal predicates *TD* and *SP*, and of the other predicates used. For the one-place predicates that we are dealing with here, we have to specify which intensions fall under these predicates in which cases. Since there are $5^4 = 625$ possible intensions (each of the four cases from $\Gamma$ can be assigned any one of the five extensions from $D$), a full specification would be too long. We abbreviate by just stating whether the terms $s$ and $s'$ fall under the predicates in the respective cases, as shown in Table 2. This will be enough: We can check by inspection that all our premises (1')–(8') are satisfied in our model, proving their consistency.

The truth of (1') and (2') is obvious from the $\gamma_0$ column of Table 2. For (3'), we have to check that there is no case in which both *SL* and *ED* apply to the same things, and this can be checked by comparing the entries in the rows for *SL* and for *ED*, which indeed have no overlap in any case. For (4'), the antecedent is fulfilled for $s$ in case $\gamma_0$ $(= \gamma_@)$, and so *SL* has to hold of $s$ in all cases, which it does. For (5'), as $SP(s')$ is true at $\gamma_0$, we need to establish that $\Diamond MR(s')$ holds at $\gamma_0$, i.e., that there is some case in which $MR(s')$ holds—and $\gamma_2^2$ is available as a witness. That case is the only one for which we need to check the consequent of (6'), and in fact, $ED(s')$ holds at $\gamma_2^2$. Formula (7') holds vacuously, since nothing satisfies $ED(x)$ in case $\gamma_0$. For (8'), we have to check the three cases in which $TD(s)$ holds, $\gamma_0$, $\gamma_1$, and $\gamma_2^1$; in all these cases, the equation $s = s'$ is satisfied, and therefore, $s'$ provides a witness for the existence claim.

**Table 2**
Interpretation of the predicates in our model.

| | $\gamma_0$ | $\gamma_1$ | $\gamma_2^1$ | $\gamma_2^2$ |
|---|---|---|---|---|
| *TD* | $s$ | $s$ | $s$ | – |
| *SP* | $s'$ | $s'$ | $s'$ | $s'$ |
| *EI* | $s, s'$ | $s, s'$ | $s, s'$ | – |
| *ED* | – | – | – | $s'$ |
| *SL* | $s, s'$ | $s, s'$ | $s, s'$ | $s$ |
| *MR* | – | – | – | $s'$ |

We have hereby shown that CIFOL has the expressive resources to formalize and to satisfy all of the assumptions made in the reversibility thought experiment. These assumptions therefore have a consistent interpretation, and do not necessitate to drop the status of the Second Law as a law. What one needs to do, when following our formalization, is just to acknowledge the specific role of sortal predicates in judgments of possibility and necessity.

## 6. Conclusion

In this paper we have advocated a fresh look at the well-known Loschmidt reversibility objection against the Second Law of thermodynamics. Our approach was to adopt a logical point of view and to try to formulate the objection as a deductive argument. We spelled out a formalization in a standard framework of quantified modal logic in §3, showing how a seemingly innocuous list of premises allows one to derive a contradiction. This we take to be the logical heart of the reversibility objection: There are plausible assumptions about physical possibility and necessity that are inconsistent with the Second Law. In §4 we presented case-intensional first order logic (CIFOL) as a framework offering resources for a different formal representation of the reversibility objection. Given these resources we could show, in §5, that it is possible to formalize the premises of the reversibility objection in a consistent way by highlighting the role of sortal predicates, especially in identity judgments. The CIFOL formulation blocks the contradiction from §3. In fact, we constructed an explicit model that proves consistency of our assumptions (1')–(8'). According to the CIFOL formalization of the reversibility objection, motion reversal of a system of particles is a general possibility, and this blocks the Second Law from holding generally for systems of particles. For a thermodynamic system, therefore, motion reversal amounts to destruction: take a thermodynamic system and reverse the motion of all of its particles, and you are left with a system of particles that no longer fulfills the laws of thermodynamics.

Such anti-thermodynamic systems are physically possible—coffee and cream might unmix rather than mix, given a carefully manufactured brew. We offer no explanation for the actual natural prevalence of thermodynamic systems, nor for the technological difficulty of motion reversal, in this paper. As we said in §2, these are interesting and relevant questions of natural history. We hold, however, that it is useful to keep these questions separate from the logical question of how to deal, in a consistent way, with the possibility of motion reversal for systems of particles vis-à-vis the Second Law of thermodynamics. That logical question has been our target, and we have offered a formally explicit answer.

Our result is, in one sense, neither new nor surprising. Everybody knows that one needs to make some distinction between thermodynamic systems and mere systems of particles, and that the systems' possible temporal development is crucial for making that distinction. Employing the resources of case-intensional logic, we can represent this distinction as a difference in kinds of systems, which allows for different laws, while at the same time holding on to the fact that TD systems are nothing over and above systems of particles. In our view, this constitutes a significant step forward in our understanding of the reversibility objection. Our formalization shows precisely what is right, and what is wrong, with the quick informal argument that we know so well. Intricate issues of identity and possibility lie at the heart of the matter, and case-intensional first order logic provides a framework for disentangling them.

## Acknowledgments

The author acknowledges support by the European Research Council (ERC) under the European Community's Seventh Framework Programme (FP7/2007–2013), ERC Grant agreement nr 263227, and by the Templeton World Charity Foundation, grant TWCF0078/AB46.

Thanks to Roman Frigg, Jenann Ismael, and Charlotte Werndl for feedback on an earlier version of the material, to an anonymous referee for helpful comments on a previous draft, and to audiences at Barcelona, Bern, Bielefeld, Düsseldorf, Gent, Innsbruck, Kraków, München, Osnabrück, Padova, Salzburg, and Utrecht for discussion.

## A. CIFOL's formal semantics

In §4 we sketched the semantic clauses of CIFOL in an informal way. Here we list all relevant clauses for completeness. We also make fully explicit the dependence of truth values (extensions of sentences) on all semantic parameters: the model $\mathscr{M}$, assignment $\delta$, and case $\gamma$.

A CIFOL model is a triple $\mathscr{M} = \langle \Gamma, D, \mathscr{I} \rangle$, where $\Gamma$ is a non-empty set of cases, $D$ is a set of extensions containing the special extension $*$ for non-existence in a case, and $\mathscr{I}$ is an interpretation function that assigns intensions to the constants and predicates of our language. The interpretation of a constant term is an intension, i.e., a function from cases to extensions:

$$int_{\mathscr{M}}(c) = \mathscr{I}(c) : \Gamma \mapsto D,$$

and the interpretation of a one-place predicate is a function from cases to functions from intensions to truth values,

$$int_{\mathscr{M}}(P) = \mathscr{I}(P) : \Gamma \mapsto ((\Gamma \mapsto D) \mapsto \{\mathbf{T}, \mathbf{F}\}),$$

and so on for higher arities.

The intension of a term (variable or constant), given a model and an assignment $\delta : Var \mapsto (\Gamma \mapsto D)$, is defined as follows:

$$int_{\mathscr{M},\delta}(\tau) = \begin{cases} \mathscr{I}(c) & \text{iff} \quad \tau \text{ is a constant } c, \quad \text{and} \\ \delta(x) & \text{iff} \quad \tau \text{ is a variable } x. \end{cases}$$

Accordingly, the extension of a term (variable or constant), given a model, an assignment, and a case, is defined as follows[21]:

$$ext_{\mathscr{M},\delta,\gamma}(\tau) = \left(int_{\mathscr{M},\delta}(\tau)\right)(\gamma) = \begin{cases} (\mathscr{I}(c))(\gamma) & \text{iff} \quad \tau \text{ is a constant } c, \quad \text{and} \\ (\delta(x))(\gamma) & \text{iff} \quad \tau \text{ is a variable } x. \end{cases}$$

There are two types of basic formulae, viz., identity statements and predications, which have the following extensions (truth values):

$$ext_{\mathscr{M},\delta,\gamma}(\tau_1 = \tau_2) = \begin{cases} \mathbf{T} & \text{iff} \quad ext_{\mathscr{M},\delta,\gamma}(\tau_1) = ext_{\mathscr{M},\delta,\gamma}(\tau_2), \\ \mathbf{F} & \text{otherwise;} \end{cases}$$

$$ext_{\mathscr{M},\delta,\gamma}(P(\tau)) = \left(\left(int_{\mathscr{M},\delta}(P)\right)(\gamma)\right)\left(int_{\mathscr{M},\delta}(\tau)\right).$$

Thus, an identity statement is true in a case iff the relevant extensions are the same. For the truth value of a predication $P(\tau)$ in a case $\gamma$, first take the intension of $P$, $int_{\mathscr{M},\delta}(P)$, which is a function from cases to functions from intensions to truth values, then apply that function to the given case $\gamma$, which results in a function from intensions to truth values $((int_{\mathscr{M},\delta}(P))(\gamma))$, and then apply that function to the intension $(int_{\mathscr{M},\delta}(\tau))$ of the term $\tau$, which results in a truth value.

Complex formulae are built up from basic ones using the truth functional operators $\neg$ and $\wedge$ (adding parentheses where appropriate), the necessity operator $\Box$, and universal quantification $\forall x$ for a variable $x \in Var$:

$$ext_{\mathscr{M},\delta,\gamma}(\neg\phi) = \begin{cases} \mathbf{T} & \text{iff} \quad ext_{\mathscr{M},\delta,\gamma}(\phi) = \mathbf{F}, \\ \mathbf{F} & \text{otherwise;} \end{cases}$$

$$ext_{\mathscr{M},\delta,\gamma}(\phi_1 \wedge \phi_2) = \begin{cases} \mathbf{T} & \text{iff} \quad ext_{\mathscr{M},\delta,\gamma}(\phi_1) = ext_{\mathscr{M},\delta,\gamma}(\phi_2) = \mathbf{T}, \\ \mathbf{F} & \text{otherwise;} \end{cases}$$

$$ext_{\mathscr{M},\delta,\gamma}(\Box\phi) = \begin{cases} \mathbf{T} & \text{iff} \quad \text{for all } \gamma' \in \Gamma, \quad ext_{\mathscr{M},\delta,\gamma'}(\phi) = \mathbf{T}, \\ \mathbf{F} & \text{otherwise;} \end{cases}$$

$$ext_{\mathscr{M},\delta,\gamma}(\forall x\phi) = \begin{cases} \mathbf{T} & \text{iff for all intensions } \bar{z}, \quad ext_{\mathscr{M},[\bar{z}/x]\delta,\gamma}(\phi) = \mathbf{T}, \\ \mathbf{F} & \text{otherwise.} \end{cases}$$

In the clause for the universal quantifier, we have used the notion of a shifted assignment $[\bar{z}/x]\delta$, which is defined in the standard way, as follows: Given an assignment $\delta : Var \mapsto (\Gamma \mapsto D)$, an intension $\bar{z} : \Gamma \mapsto D$, and a variable $x \in Var$,

$$([\bar{z}/x]\delta)(y) = \begin{cases} \bar{z} & \text{iff} \quad y = x, \\ \delta(y) & \text{otherwise.} \end{cases}$$

We also use the material conditional "$\rightarrow$" and the material biconditional "$\leftrightarrow$" in the standard way, and we introduce the duals of the necessity operator and the universal quantifier, the possibility operator "$\Diamond$" and the existential quantifier "$\exists x$", as abbreviations:

$$\Diamond\phi \quad \text{is short for} \quad \neg\Box\neg\phi; \qquad \exists x\phi \quad \text{is short for} \quad \neg\forall x\neg\phi.$$

For the existence predicate $E(x)$, we introduce the special constant term "$*$", with $ext_{\mathscr{M},\delta,\gamma}(*) = *$ for all $\gamma \in \Gamma$, and then introduce the abbreviation

$$E(x) \quad \text{is short for} \quad \neg(x = *).$$

We also use the satisfaction relation:

$$\mathscr{M}, \delta, \gamma \models \phi \quad \text{iff} \quad ext_{\mathscr{M},\delta,\gamma}(\phi) = \mathbf{T},$$

where we drop $\mathscr{M}$ if it is clear from context, and $\delta$ if it plays no role (e.g., for closed sentences).

Based on these semantic definitions, we can lay down axioms characterizing certain types of predicates (or, more generally, contexts for substitution). Some predicates are *extensional*, meaning that whether they apply to a term in a case only depends on the case-relative extension:

$$Ext(P) \Leftrightarrow_{df} \Box\forall x\forall y\, [x = y \rightarrow (P(x) \leftrightarrow P(y))].$$

Two other important properties of predicates are "modal constancy":

[21] See Belnap and Müller (2014a) for the somewhat lengthier clause for definite descriptions. As explained in that paper, the uniform treatment of all terms in CIFOL means that the controversial notion of a "rigid designator", which has become famous following Kripke (1980), is not needed in CIFOL.

$$MConst(P) \Leftrightarrow_{df} \forall x \left[\Diamond(E(x) \wedge P(x)) \rightarrow \Box(E(x) \rightarrow P(x))\right]$$

and modal separation:

$$MSep(P) \Leftrightarrow_{df} \Box \forall x \forall y \left[(P(x) \wedge P(y)) \rightarrow \left(\Diamond(E(x) \wedge E(y) \wedge x = y) \rightarrow \Box(x = y)\right)\right].$$

Predicates having these properties are not extensional. A modally constant predicate applies to a term in all cases in which its referent exists, given just that it applies in one case in which the referent exists. And a modally separated predicate is such that if two intensions fall under it and coincide non-vacuously in one case (i.e., in some case, *Ex* and *Ey* and $x = y$), then these intensions are the same (i.e., $\Box(x = y)$). In this way, a single non-trivial extension is enough to pick out a whole intension.

A predicate that fulfills both modal constancy and modal separation allows one to trace a thing through all possible cases given just an extensional identification in one case. Such a predicate is called a *CIFOL sortal predicate:*[22]

$$Sortal(P) \Leftrightarrow_{df} MConst(P) \wedge MSep(P).$$

## References


Albert, D. Z. (2000). *Time and chance*. Cambridge, MA: Harvard University Press.

Belnap, N., & Müller, T. (2014a). CIFOL: Case-intensional first oder logic. (I) Toward a theory of sorts. *Journal of Philosophical Logic, 43*, 393–437. https://doi.org/10.1007/s10992-012-9267-x.

Belnap, N., & Müller, T. (2014b). BH-CIFOL: Case-intensional first oder logic. (II) Branching histories. *Journal of Philosophical Logic, 43*, 835–866. https://doi.org/10.1007/s10992-013-9292-4.

Bressan, A. (1972). *A general interpreted modal calculus*. New Haven, CT: Yale University Press.

Carnap, R. (1947). *Meaning and necessity: A study in semantics and modal logic*. Chicago, IL: University of Chicago Press. Enlarged edition, 1956.

Einstein, A. (1949). Autobiographical notes. In P. A. Schilpp (Ed.), *Albert Einstein: Philosopher-scientist, pages 1–96. Open court, La Salle, IL*.

Garson, J. W. (2013). *Modal logic for philosophers* (2nd ed.). Cambridge: Cambridge University Press.

Hahn, E. L. (1950). Spin echoes. *Physical Review, 80*, 580–601.

Ismael, J. (2016). Could statistical mechanical probabilities have a quantum mechanical grounding? Assessing Albert's proposal. In B. Weslake, & E. Winsberg (Eds.), *Time's arrow and the probability structure of the world*. Cambridge, MA: Harvard University Press (in press).

Kripke, S. (1980). *Naming and necessity*. Oxford: Blackwell. Originally published in 1972.

Leeds, S. (2003). Foundations of statistical mechanics—two approaches. *Philosophy of Science, 70*(1), 126–144.

Loewer, B. (2004). David Lewis's Humean theory of objective chance. *Philosophy of Science, 71*(5), 1115–1125.

Loschmidt, J. (1876). Über den Zustand des Wärmegleichgewichtes eines Systems von Körpern mit Rücksicht auf die Schwerkraft. *Sitzungsber. Kais. Akad. Wiss. Wien, Math. Naturwiss. Classe, 73*, 128–142.

Reichenbach, H. (1956). *The direction of time*. Berkeley, CA: University of California Press.

Ridderbos, T. M., & Redhead, M. L. G. (1998). The spin-echo experiments and the second law of thermodynamics. *Foundations of Physics, 28*(8), 1237–1270.

Schrödinger, E. (1950). Irreversibility. *Proceedings of the Royal Irish Academy, 53*, 189–195.

Winsberg, E. (2008). Laws and chances in statistical mechanics. *Studies in the history and philosophy of modern physics, 39*(4), 872–888.


[22] Bressan (1972) speaks of "absolute predicates". See Belnap and Müller (2014a) for a more detailed introduction to the properties of predicates briefly discussed here. Note that our definition of *MConst* deviates slightly from the definition in Belnap and Müller (2014a), as mentioned in note 20.