Non-monotonic set theory as a pragmatic foundation of mathematics*

Peter Verdée†
peter.verdee@ugent.be
Centre for Logic and Philosophy of Science
Ghent University, Belgium

Abstract

In this paper I propose a new approach to the foundation of mathematics: non-monotonic set theory. I present two completely different methods to develop set theories based on adaptive logics. For both theories there is a finitistic non-triviality proof and both theories contain (a subtle version of) the comprehension axiom schema. The first theory contains only a maximal selection of instances of the comprehension schema that do not lead to inconsistencies. The second allows for all the instances, also the inconsistent ones, but restricts the conclusions one can draw from them in order to avoid triviality. The theories have enough expressive power to form a justification/explication for most of the established results of classical mathematics. They are therefore not limited by Gödel's incompleteness theorems. This remarkable result is possible because of the non-recursive character of the final proofs of theorems of non-monotonic theories. I shall argue that, precisely because of the computational complexity of these final proofs, we cannot claim that non-monotonic theories are ideal foundations for mathematics. Nevertheless, thanks to their strength, first order language and the recursive *dynamic* (defeasible) proofs of theorems of the theory, the non-monotonic theories form (what I call) interesting *pragmatic* foundations.

1 Introduction: founding mathematics

The project of the foundation of mathematics concerns the search for a reduction of mathematics to basic concepts and rules determining the behavior of the concepts. This is an essential aspect of elaborating a philosophy of mathematics. It is an important tool for answering questions concerning the epistemic status of mathematical objects, truths and proofs and thus for giving a justification or an explanation for the reliability of mathematical knowledge. The first systematic foundation of (a part of) mathematics was Euclid's geometry. He reduced geometry to basic kinds of objects: points, lines, circles and planes and a number of relations between these objects. His postulates and axioms determined which statements are accepted about these objects and their relations.

*I am very indebted to Laszlo Kosolosky, Diderik Batens and two anonymous referees who suggested many useful corrections and improvements on a former version of this paper.
†Peter Verdée is a post-doctoral fellow of the Fund for Scientific Research – Flanders.

Pure logic determined which conclusions could be drawn from these postulates and axioms. Assuming that the deductive reasoning done by logic was truth preserving, the truth of the whole domain of results in Euclidian geometry was ensured by the truth of the postulates and axioms. The truth of the axioms and postulates was taken to be self-evident.

Later on, since Frege, the logical part of foundational theories was also formalized and made precise. This made it possible to formulate theories in a fully symbolic way. I shall call such formalized theories *formal theories* and use the following (usual) definition for them: a formal theory is a pair consisting of a recursive set of formulas (axioms) of a formal language and a symbolic logic for this formal language that determines the consequences of the axioms.

The process of founding mathematics by means of a formal theory would ideally proceed in the following way. (1) Select a coherent set of basic objects of a mathematical domain in such a way that all the other objects used by mathematicians can be defined from them. (2) Select a recursive set of accepted truths about those objects. (3) Devise a logic that has a proof system such that every accepted proof in mathematics can (ideally) be translated into a formal proof of the logic. (4) Make sure that the recursive set of truths selected in (2) is strong enough to (ideally) prove all accepted results of the mathematical domain, i.e. all truths about the basic objects. And finally (5): make sure that the obtained theory is provably non-trivial.

Remark that non-triviality is in fact strictly speaking too weak a requirement. A theory is non-trivial iff there is at least one formula of its language that is not a theorem. In principle, it might be the case that a theory is non-trivial when the only formulas that are not theorems are in some small fragment of the theory (e.g. the case where only $\perp$ is not a theorem). In that case, most of the theory is still useless. A theory can only be considered a useful foundation if it succeeds in distinguishing theorems from non-theorems for the entire domain. I shall call a theory that succeeds in making this distinction *discriminatory*. This is obviously a vague term, but for most theories that are structured by a formal logic with a meaningful semantics, being discriminatory and non-trivial are equivalent.

Once one would have obtained such an ideal foundation, the practice of mathematics would be a more controlled, safer and clearer enterprise. We would be more certain of the correctness, the meaningfulness and the coherence of mathematical results. Moreover, the philosophical mission of explaining the meaning and the epistemic status of mathematical results can now be reduced to the set of basic objects and truths.

At the beginning of the 20th century, the formalist and the logicist schools attempted to achieve such an ideal foundation by means of axiomatic calculi. They tried to reduce a large part of mathematics to the consequences of a simple set of logical axioms (for the logicists) or logic and theoretical axioms (for the formalists) formulated in a formal language. Exact rules would determine how to derive new sentences from the axioms. Thus one hoped to obtain an exact axiomatic calculus that would be able to replace the existing informal way of doing mathematics. The calculi would determine the truths of mathematics in an exact way (although many of them would still be unknown, it would only be a matter of discovering the already fixed mathematical universe, not of creating the unfixed universe).

In 1931 Gödel put an end to this foundational dream (cf. [12]). His first incompleteness theorem entails that it is impossible to formulate a non-trivial, complete axiomatic calculus for mathematical domains rich enough to function as a model for arithmetic. The second shows that it is impossible to prove within a given theory its own consistency even for relatively simple static mathematical theories that contain Peano arithmetic.

There are different ways to respond to this drastic result. The most significant categories of alternative (non-ideal) foundations taking Gödel's results into account can be characterized as *partial*, *relative* or *revisionist foundations*. These types of foundations are not ideal foundations but attempt to be optimal foundations. I define a relative foundation as a foundation the non-triviality of which depends on the non-triviality of other infinitistic (mathematical) systems. Hence, there is no real guarantee for the non-triviality of relative foundations. A partial foundation is a foundation which does not cover the whole mathematical domain for which one aims to devise a foundation. From a classical point of view on mathematical semantics, the partiality comes to the negation incompleteness of the foundational theories—a constructivist foundation might be complete and negation-incomplete, as constructivist mathematics does not require a negation-complete foundation. Finally, a foundation is revisionist iff it criticizes part of classical mathematics and therefore only needs to found the part of mathematics that is not criticized.

The most salient and popular foundation nowadays is **ZFC**. The theory **ZFC** is the most commonly used set theory (cf. [11]). It is an extension of Zermelo's set theory (cf. [32]). The underlying logic of **ZFC** is classical logic. Everywhere in this paper, classical logic (henceforth **CL**) is the classical first order consequence relation with the identity symbol and without function symbols. The axioms of **ZFC** are presented in Section 3. It is commonly accepted that most of actual mathematics can be done in **ZFC**. In the project Metamath (cf. `http://us.metamath.org`), for example, many interesting mathematical theorems of very different domains of mathematics are formally proved from the axioms of **ZFC**. So **ZFC** is rather rich and many mathematical results can be expressed in it. Nevertheless it is not an ideal foundation of mathematics. **ZFC** is obviously incomplete, with a large amount of propositions for which it can be proved that the proposition nor its negation is in **ZFC** provided **ZFC** is consistent. Examples are the Continuum Hypothesis, the Axiom of Constructibility, Whitehead's problem and propositions expressing **ZFC**'s consistency (cf. [9, 14]). So **ZFC** is too weak to provide an ideal foundation. But **ZFC** is also too rich for this purpose. It is much richer than a theory like Peano-arithmetic. One has so far not been able to give a convincing consistency proof for Peano-arithematic, without using proof methods the consistency of which is as problematic as the consistency of Peano-arithmetic itself. **ZFC** being even richer, more abstract and further away from intuitions, the consistency of it is an even bigger open problem. The most convincing argument for the consistency of **ZFC** is the following: we have been extensively working with this theory for a long time and have not found an inconsistency, so the theory is unlikely to be inconsistent.

Nevertheless, **ZFC** can still be considered as a foundation, albeit a relative and partial one. As long as **ZFC** does not collapse, it offers a rigorous justification for many mathematical results, with the certainty that these results are not based on unclear proofs and do not rely on implicit intuitions or prejudices. Its relative character is unlikely to be removed in view of Gödel's results, but mathematicians might find sensible extensions of **ZFC** that cover more of mathematics. In that sense, it may be possible to make **ZFC** 'less partial'. Nevertheless it will always remain partial and it will probably always remain as relative as it is now (unless it collapses).

Another option is to restrict classical mathematics, resulting in revisionist foundations. Examples are constructive set theory (which uses an intuitionistic logic and only aims at founding constructive mathematics), strictly finitistic mathematics (which does not accept references to infinity, and therefore does not require a theoretic foundation of infinity), etc. This is a way to avoid partiality and relativity. If one rejects part of classical mathematics, the theories one develops as foundations obviously need not cover the entire classical mathematics and need not be strong enough to be susceptible to Gödel's result. The retained part of classical mathematics may be foundable by means of a complete axiomatic calculus, possibly even one with an absolute non-triviality proof. So Gödel's incompleteness results might be avoided in such approaches. Of course, these foundations are not (and do not want to be) foundations for classical mathematics.

In this paper I propose a new approach to foundation, viz. *pragmatic foundation*. Pragmatic foundations make use of object proofs that do not warrant that certain formulas are finally derived–for a precise definition of final derivability see section 2–from the axioms. Informal mathematical proofs are formalized into proofs with lines that require pragmatic assumptions. These foundations are able to provide maximal metatheoretical elegance, but might at the object level involve the provisional acceptance of pragmatic conclusions, i.e. conclusions that involve provisional uncertainties at some point.

A simple and clear (but rather sterile) example of a pragmatic foundation is the theory **CZFC** defined by the following definition: the axioms of **CZFC** are exactly the same as those of **ZFC**, but the underlying logic is a non-trivializing version of **CL**. This logic accepts all **CL**-consequences of consistent premises, but accepts no consequences of inconsistent premises. Consequently, A is a theorem of **CZFC** iff (A is a **ZFC**-theorem and **ZFC** is consistent). This theory has effectively the same function as plain **ZFC**: if **ZFC** is consistent, **CZFC** is equivalent to it. If **ZFC** is inconsistent, **ZFC** is trivial and none of the **ZFC**-consequences is trustworthy, whence, from a practical point of view, the foundational part of **ZFC** is empty (none of the theorems of **ZFC** has the power to support mathematics in this case). This *foundational part* of **ZFC** is exactly what the theory **CZFC** formalizes.

Formal proofs for the theory **CZFC** could be conceived as follows: consider usual **ZFC**-proofs. Add to every line of such proofs a condition containing the formula $\emptyset \in \emptyset$. Mark lines as 'no longer derived' from the point on that the condition of that line is derived on some other line of the proof. So no lines are marked unless and until the inconsistent formula $\emptyset \in \emptyset$ is derived. Such a proof is dynamic. At no point in time shall we be certain that any formula is a theorem of **CZFC**. Unless an inconsistency is derived from **ZFC**, every conclusion one is able to draw within a dynamic, pragmatic **CZFC**-proof remains defeasible and essentially uncertain.

Of course this uncertainty is a disadvantage. However, compared to **ZFC** itself, **CZFC** is not worse off. The uncertainty is equally present in **ZFC**, for if **ZFC** would be inconsistent, there would be nothing foundational about **ZFC**. In other words, it is unproblematically certain that the conclusions of **ZFC**-

proofs are theorems of **ZFC**. But their foundational strength will always stay uncertain and conditional.

CZFC is as incomplete as **ZFC**, so, as a foundation, **CZFC** is still only partial. However, **CZFC** is evidently non-trivial and can therefore function as an absolute foundation. This makes **CZFC** a partial, pragmatic but absolute foundation.

CZFC is what I call a *non-monotonic theory*, i.e. a theory with a non-monotonic underlying logic. A non-monotonic logic is a logic for which adding more premises may result in less consequences. Why is the underlying logic of **CZFC** non-monotonic? Well, consider the axioms of **ZFC** and add the negation of one of the axioms to it. The resulting set of axioms is inconsistent and hence the set of theorems of this extension of **CZFC** is empty.

Although the non-monotonic theory **CZFC** is an interesting introductory example to present the possibilities of pragmatic foundations, it does not deliver any practical new results. The only difference with **ZFC** is that it is warranted to be non-trivial. What is essentially realized here is that the metatheoretical consideration that **ZFC** would lose its foundational power if it would turn out to be trivial is taken into account at the object level of **CZFC**. So by defining **CZFC**, we already devised an absolute but pragmatic partial foundation out of a relative partial foundation.

Non-monotonic theories generally do not have recursive final proofs for their theorems. Indeed, in order to consider some formulas as theorems of a non-monotonic theory, one needs to know that other formulas are non-theorems of it. In terms of the dynamic proofs: in order to be certain that a conditional line will not be marked in any extension of a dynamic proof, one needs certainty about an infinity of extensions of the proof. In other words: establishing final theoremhood sometimes requires logical omniscience. This makes most non-monotonic theories and pragmatic foundations not semi-recursive. Of course this not an ideal situation for a foundation. But it is precisely this property that makes non-monotonic theories not restricted by Gödel's incompleteness theorem, which depend on the recursiveness of the proofs for demonstrating theoremhood of the theory. Rich enough semi-recursive theories will somehow always be susceptible to Gödel's limitations. If one wants to get around Gödel and design theories rich enough to found mathematics, one needs to give up semi-recursiveness.

It is however far more interesting to devise pragmatic foundations that not only take problems into account, but also solve them. Adaptive logics (henceforth **AL**s) are good tools to create such foundations. They provide the required dynamic proofs plus define a fixed and subtle set of consequences for every premise set. This set can form the set of theorems of the pragmatic foundation. By means of adaptive logics we can make foundations that are stronger than **CZFC** and make sure that the foundation does not disappear when an inconsistency is found in **ZFC**, but rather isolate that inconsistency and enable as much of classical mathematics as possible in parts of the theory where the inconsistency is absent.

Providing a subtle safety net for the case where classical theories would turn out to be trivial is but one aspect of the promising domain of pragmatic foundations. Pragmatic foundations can in fact be much stronger and expressive than usual foundations. They can delineate the sensible parts from intuitive but inconsistent theories, without yielding a weak theory. This will be illustrated in this paper for the inconsistent axioms of naive set theory. It is even realistic that non-monotonic mathematical theories would be negation-complete, whereas usual theories are essentially unable to provide this property. This will be elaborated in a forthcoming paper.

It is generally accepted that the set of all classical mathematical truths (no matter what one philosophically means by the notion 'mathematical truth') is not semi-recursive. Every usual axiomatic calculus, however, is semi-recursive by definition. It should not come as a surprise that the known calculi are essentially insufficiently rich. Non-monotonic theories on the other hand do not have this restriction. They can be far more complex than usual semi-recursive theories and so they are more likely to actually fully capture mathematical truth.

In the present paper I shall give two basic examples of non-monotonic set theories that show different possible ways to devise pragmatic foundations. I shall discuss different advantages of choosing these types of foundations. For my present purposes, I shall not focus on technical details of the systems, but rather explain the possible formal methods to devise adaptive set theories, present some problems one is confronted with when constructing them, and propose possible solutions. I shall also offer a sketch of what the proposed theories may mean for the foundations of mathematics.

The systems I present here are based on the inconsistent comprehension axiom. This axiom states that, for every property $A(\alpha)$, the set defined by $\{x \mid A(x)\}$ exists. This is the most intuitive, most natural way to devise the concept of a set (strong arguments in favor of comprehension are provided in [30] and [15]): every property of the language defines a set. Of course some properties are inconsistent (logically false), but this is not a problem: a set defined by means of an inconsistency is simply an empty set. The problem is that stating the existence of some sets defined by a consistent property (which refers to self-membership) leads to inconsistencies. Consequently, if the underlying logic of a theory with the comprehension axiom is **CL**, the theory is trivial.

There exist several ways out of this without giving up **CL**, such as **ZFC**, Quine's New Foundations (cf. [18]), von Neumann-Bernays-Gödel set theory (cf. [27]) and Morse-Kelley set theory (cf. [13]). All of these restrict the comprehension axiom in such a way that self membership is avoided. This results in more complicated axioms, the non-triviality of which is still an open problem.

I shall sketch different methods to devise non-monotonic theories based on **AL** that allow for instances of the full comprehension axiom whenever this does not lead to problems. The non-triviality of all presented theories is finitisticly provable. If the classical set theories would turn out to be inconsistent, the non-monotonic theories still give sensible results and suggest how to correct the trivial classical set theories. The methods are all variations on two basic approaches to adaptive set theory. The first basic theory, called Maximally Consistent Comprehension Set Theory (**MCC**), only allows for the consistent instances of the comprehension axiom schema. **MCC** contains all the **CL**-consequences of the consistent instances. The second basic theory, called Maximally Rich Universal Set Theory (**MRU**), is an inconsistent set theory that is what I call *universal*: it proves the existence of *all* the sets the existence of which is stated by the comprehension axiom. For the paradoxical sets the logical rules are restricted in such a way that no trivialities are obtained. For the unproblematic sets all **CL**-rules are allowed.

It is important to emphasize that this paper mainly has a programmatic purpose. I want to show that non-monotonic theories can have some surprisingly interesting properties and that adaptive logics are particularly good logics for non-monotonic set theory. I show different techniques one could use when one devises such adaptive set theories. I definitely do not want to prove that this is a superior approach to the foundation of mathematics. Neither do I want to claim that the presented techniques result in properly elaborated and investigated theories, with all optimal properties. I only aim to convince the reader that (i) non-monotonic set theory is a fascinating new approach to foundational questions, (ii) it therefore deserves further investigation and (iii) there are useful techniques available within the adaptive logic framework that might solve obvious issues in non-monotonic set theory.

In the next section, I give a short introduction to **AL**. In the third section I introduce the axioms of the relevant existing set theories, **ZFC** and naive set theory and list a number of paradoxes of naive set theory. The fourth section contains a presentation of the set theory **MCC**. In fifth section I present **MRU**. In section 6, I compare the two theories and I conclude the paper with an overview of the advantages of non-monotonic foundations of mathematics.

2 Adaptive logics

The non-monotonic theories I shall present are based on an **AL** (cf. [1, 2, 3, 5] for some general formal and philosophical introductions to **AL**). **AL**s are excellent tools to formalize defeasible reasoning. There is one elegant formal format for most **AL**s, called the *standard format of* **AL**. A large amount of very different types of defeasible reasoning have been characterized by means of an **AL** in standard format: abductive reasoning, inductive generalization, handling inconsistencies, reasoning with vagueness, reasoning with ambiguity, reasoning about compatibility, question raising, coping with theories where statements are only plausibly true, diagnosis, causal discovery, belief merging and default reasoning. The dynamic proof theory, the semantics and the meta-theory of **AL**s in standard format are generic and intuitive. The dynamic proofs are intuitive explications for actual reasoning processes.

The standard format of **AL** is extremely unifying. Because most kinds of defeasible reasoning can be formalized by means of an **AL** in standard format, the format together with its metatheoretic properties reveals the essential formal structure of defeasible reasoning.

There is a static consequence notion that assigns a fixed set of **AL**-consequences to every premise set. It is this notion that we need for the fixed mathematical theories. The dynamic proofs are used to formalize the informal mathematical proofs. Of course, this is not a straightforward formalization. The informal mathematical proofs are meant to be static: when a conclusion is obtained, it is meant to be a theorem, unconditionally and forever. The dynamic proofs that formalize the informal proofs cannot assure this unconditional validity. Conditional lines of dynamic proofs are accepted as pragmatic certainties as long as they are not revoked. This is why I called **AL**-theories pragmatic foundations. The certainty their proofs deliver is only pragmatic, not irrevocable.

Adaptive logics take formulas of some predefined type, called *abnormalities*, to be false unless and until it is proven that the premises do not allow this presupposition. This 'unless and until' determines the dynamic and defeasible character of adaptive logics. Semantically, taking abnormalities to be false whenever this is compatible with the premises comes to interpreting premise sets as normally as possible. Adaptive logics do this by selecting those models of the premises in which as little abnormalities as possible are true.

An **AL** in standard format is defined as a triple consisting of:

- a **LLL**: a monotonic, reflexive, transitive and compact logic which has a characteristic semantics and contains all the **CL**-symbols.
- a set of abnormalities Ω: a set of **LLL**-contingent formulas Ω, characterized by a (possibly restricted) logical form, and
- a strategy (the most important adaptive strategies are 'Reliability' and 'Minimal Abnormality').

The lower limit logic is the stable part of the **AL**; anything that follows from the premises by the **LLL** will never be revoked. The **LLL** should contain all **CL**-symbols. This is realized by adding, to a logic with its own standard symbols, all the **CL**-symbols with a check '$\check{}$' above the symbol, e.g. $\check{\vee}$, $\check{\neg}$, $\check{\bot}$, $\check{\top}$, etc. They are superimposed on the standard symbols, i.e. they do not occur within the scope of the standard symbols. They are semantically defined by means of their usual classical properties with respect to model verification (e.g $M \models A \check{\vee} B$ iff $M \models A$ or $M \models B$, $M \models \check{\neg} A$ iff $M \not\models A$, $M \models \check{\top}$, and $M \not\models \check{\bot}$) and hence they behave classically. The CL-symbols are added to a usual logic with its own standard symbols, for purely technical reasons. They occur neither in the premises nor in the conclusions. In the concrete logics in this paper, the standard symbols $\neg$,$\vee$,$\wedge$,$\bot$ etc. already behave exactly classically, so there the checked symbols are replaced by their standard counterparts.

The abnormalities are the formulas the **AL** supposes to be false, 'unless and until proven otherwise'. Strategies are ways to cope with derivable disjunctions of abnormalities: an adaptive strategy picks one specific way to interpret the premises as normally as possible. Apart from Reliability and Minimal Abnormality, several strategies were developed mainly in order to characterize consequence relations from the literature in terms of an **AL**. All those strategies can be reduced to Reliability or Minimal Abnormality under a translation.

If the lower limit logic is extended with an axiom that declares all abnormalities logically false, one obtains the *upper limit logic* **ULL**. If a premise set Γ does not require that any abnormalities are true, the **AL**-consequences of Γ are identical to its **ULL**-consequences.

2.1 The proof theory of AL

The proof theory of an **AL** consists of a set of inference rules (determined by the **LLL** and Ω) and a marking definition (determined by Ω and the chosen strategy). A line of an annotated **AL**-proof consists of four elements: (1) a line number i, (2) a formula A, (3) a condition consisting of a set of abnormalities $\Theta \subset \Omega$ and (4) the name of a rule and the line number of the rule's premises. A stage s of a proof is the subproof that is completed up to line number s. The inference rules govern the addition of lines. There are 3 generic rules: a premise rule (PREM), an unconditional rule (RU), and a conditional rule (RC). These rules strongly depend on the rules for **LLL** and only focus on what the **AL** adds to the **LLL**. In the following table Γ refers to the premises of the proof. $Dab(\Delta)$ is a shorthand for the classical disjunction $\check{\bigvee}\Delta$ of the members of a finite $\Delta \subset \Omega$ (such a formula is called a *Dab-formula*).

PREM If $A \in \Gamma$

$$\begin{array}{ll} \dots & \dots \\ \hline A & \emptyset \end{array}$$

RU If $A_1, \dots, A_n \vdash_{\mathbf{LLL}} B$

$$\begin{array}{ll} A_1 & \Delta_1 \\ \dots & \dots \\ A_n & \Delta_n \\ \hline B & \Delta_1 \cup \dots \cup \Delta_n \end{array}$$

RC If $A_1, \dots, A_n \vdash_{\mathbf{LLL}} B \vee Dab(\Theta)$

$$\begin{array}{ll} A_1 & \Delta_1 \\ \dots & \dots \\ A_n & \Delta_n \\ \hline B & \Delta_1 \cup \dots \cup \Delta_n \cup \Theta \end{array}$$

The strategies determine when lines are marked in adaptive proofs and thus they determine when formulas are considered as derived in a proof. The most important strategies are Reliability and Minimal Abnormality. I only mention the Reliability strategy as I shall not consider Minimal Abnormality adaptive logics as the underlying logics of the adaptive theories I shall present. Minimal Abnormality adaptive logics are far more complex than Reliability adaptive logics (cf.[23]) and it is unclear what this strategy would add to the set theories.

$Dab(\Delta)$ is a *minimal* Dab-formula of stage s iff $Dab(\Delta)$ is derived at stage s on the condition $\emptyset$ and no $Dab(\Delta')$ with $\Delta' \subset \Delta$ is derived at stage s on the condition $\emptyset$. Where $Dab(\Delta_1)$, $\dots$, $Dab(\Delta_n)$ are the minimal Dab-formulas derived on the condition $\emptyset$ at stage s of an adaptive proof, the set of *unreliable* abnormalities at this stage, denoted as $U_s(\Gamma)$, is defined as $\Delta_1 \cup \dots \cup \Delta_n$. The abnormalities that are not unreliable at a stage are called *reliable* at that stage.

Definition 1 *Marking definition for Reliability. Where Δ is the condition of line i, line i is marked at stage s iff $\Delta \cap U_s(\Gamma) \neq \emptyset$.*

Two types of derivability are defined for **AL**s. A formula A is *derived at a stage* iff A is derived on an unmarked line at the stage. This notion will formalize the pragmatic results of (conditional) mathematical object proofs. A formula A is *finally derived* at stage s iff A is derived on an unmarked line i at stage s and any extension of the proof in which line i is marked can be further extended to a proof in which line i is unmarked. The finally derivable consequences of a premise set Γ are independent of the stage and constitute the **AL**-consequence sets for Γ, denoted by $Cn_{\mathbf{AL}}(\Gamma)$. If $A \in Cn_{\mathbf{AL}}(\Gamma)$ then I shall write $\Gamma \vdash_{\mathbf{AL}} A$. This final derivability notion will formalize the set of theorems of the adaptive theories.

2.2 The semantics of AL

Semantically, adaptive logics select **LLL**-models of the premises with respect to their normality. In the case of Reliability, we first define when a **LLL**-model is reliable and then we select all the reliable **LLL**-models of the premises. Adaptive logics validate the formulas that are verified by all selected models.

$Dab(\Delta)$ is a *minimal Dab-consequence* of Γ iff $\Gamma \vDash_{\mathbf{LLL}} Dab(\Delta)$ and, for all $\Delta' \subset \Delta$, $\Gamma \nvDash_{\mathbf{LLL}} Dab(\Delta')$. Where $Dab(\Delta_1)$, $Dab(\Delta_2)$, ... are the minimal Dab-consequences of Γ, let $U(\Gamma) =_{df} \Delta_1 \cup \Delta_2 \cup \ldots$. The members of $U(\Gamma)$ are called the *unreliable formulas of* Γ. Finally, where M is a **LLL**-model, $\mathrm{Ab}(M) =_{df} \{A \in \Omega \mid M \models A\}$.

Definition 2 *Reliable model and the corresponding semantical consequence relation* $\vDash_{\mathbf{AL}}$. *A* **LLL**-*model* M *of* Γ *is reliable iff* $\mathrm{Ab}(M) \subseteq U(\Gamma)$. *If* **AL** *uses the Reliability strategy,* $\Gamma \vDash_{\mathbf{AL}} A$ *iff all reliable models of* Γ *verify* A.

2.3 Non-triviality of AL

The following theorem shows that the adaptive consequence set of a premise set Γ is non-trivial whenever the lower limit consequence set of Γ is non-trivial.

Theorem 1 *For every set of formulas* Γ*: if there is a formula* A *such that* $\Gamma \nvdash_{\mathbf{LLL}} A$*, then there is a formula* B *such that* $\Gamma \nvdash_{\mathbf{AL}} B$.

Proof. Suppose that the antecedent is true. Then $\Gamma \nvdash_{\mathbf{LLL}} \check{\bot}$. It suffices to prove that $\Gamma \nvdash_{\mathbf{AL}} \bot$. If a formula A is derived in an adaptive proof on a condition Δ then, by the Derivability Adjustment Theorem from [2], this proof can be extended to a proof in which $A \vee Dab(\Delta)$ is derived on an empty condition. Consequently, every proof from Γ in which a conditional line i occurs on which $\check{\bot}$ is derived on the condition Δ, can be extended to a proof in which $\check{\bot} \check{\vee} Dab(\Delta)$ is derived on the empty condition. Hence, by classical logic, it can be further extended to a proof in which also $Dab(\Delta)$ is derived on the empty condition. In this proof, line i is marked given the marking definition, both in case of the Reliability Strategy and in case of the Minimal Abnormality strategy. Consequently, $\check{\bot}$ cannot be finally derived on an empty condition (otherwise $\Gamma \vdash_{\mathbf{LLL}} \check{\bot}$) nor on a non-empty condition, whence $\Gamma \nvdash_{\mathbf{AL}} \check{\bot}$. ■

3 Axioms of set theory

In what follows I denote a (formal) theory as a pair $\langle \mathsf{A}_1 + \ldots + \mathsf{A}_n, \mathbf{L}\rangle$, where A_i are axioms or axiom schemata and $\mathbf{L}$ is a logic[1]. A is a theorem of the theory $\langle \mathsf{A}_1 + \ldots + \mathsf{A}_n, \mathbf{L}\rangle$ ($A \in Th(\langle \mathsf{A}_1 + \ldots + \mathsf{A}_n, \mathbf{L}\rangle)$ for short) iff $A \in Cn_{\mathbf{L}}(\Gamma)$ where $\Gamma = \Gamma_1 \cup \ldots \cup \Gamma_n$ and for all $i \leq n$, $\Gamma_i = \{\mathsf{A}_i\}$ if A_i is an axiom, and Γ_i is the set of instances of A_i if A_i is an axiom schema.

In this paper, a set theory is a theory that formalizes the use of the membership relation '$\in$'. The only predicate in the language of the set theories presented in this paper is the binary membership predicate $\in$. Consequently, the only formulas without logical symbols will be of the form $\in\alpha\beta$, which I shall always write with the more common infix notation: $\alpha \in \beta$. As I use lowercase letters from the beginning of the Greek alphabet only as metavariables for constants and (object) variables, there should be no confusion with respect to the scope of the logical symbols when combined with formulas of the form $\alpha \in \beta$ (e.g. $\neg\alpha \in \beta$ always means $\neg(\alpha \in \beta)$).

[1] The '+'-symbol is here used as an otherwise meaningless separator in a finite list of strings.

3.1 Naive set theory

Naive set theory is the theory $\langle \mathsf{COMP} + \mathsf{EXT}, \mathbf{CL} \rangle$, where EXT and COMP are defined as follows[2]:

$$\begin{array}{ll} \mathsf{EXT} & \forall x \forall y (\forall z (z \in x \equiv z \in y) \supset x = y) \\ \mathsf{COMP} & \exists x \forall y (y \in x \equiv A(y)) \end{array}$$

Naive set theory is known to be trivial. This is due to the famous paradoxes of naive set theory. Of course, from an inconsistent set theory every single contradiction of the language is **CL**-derivable and so all formulas are actually paradoxical. But we want to correct naive set theory and are therefore interested in the origin of the inconsistency problems. The problems originate in the following primitive paradoxes (among other paradoxes).

We start with Russell's paradox.

$$\exists x \forall y (y \in x \equiv \neg y \in y)$$

Less famous variants of the Russell paradox are the following generalizations, which are discovered by Quine (cf. [19]): for every n,

$$\exists x \forall y (y \in x \equiv (\neg \exists z_1 \ldots \exists z_n (y \in z_1 \wedge z_1 \in z_2 \wedge \ldots \wedge z_n \in y)))\,.$$

One could interpret this axiom as follows. It states that the set of sets $\{x | x$ is not part of a loop of n sets$\}$ exists, where *a loop of m sets* is a series of sets a_1, a_2, $\ldots a_m$ with the property $a_1 \in a_2$, $a_2 \in a_3, \ldots, a_{m-1} \in a_m$ and $a_m \in a_1$. Each **CL**-based theory (in the sense that **CL** is the underlying logic of the theory) of which this axiom is a theorem is trivial because the classical negation of this axiom is a **CL**-theorem [3]. Remark that, for every theory that has **CL** as its underlying in logic, the set of **CL**-theorems is a subset of the set of theorems of the theory.

Curry's paradox (cf. [8]) shows that one can also express Russell's paradox without a negation:

$$\exists x \forall y (y \in x \equiv (y \in y \supset A))\,. \tag{1}$$

The negation of formulas of this form is not always a **CL**-theorem, however $\exists x \forall y (y \in x \equiv (y \in y \supset A)) \supset A$ is a **CL**-theorem, whence **CL** allows us to derive any formula A from an appropriate instance of axiom schema (1).

Also Quine's variants of Russell's paradox are expressible without a negation:

$$\exists x \forall y (y \in x \equiv (\exists z_1 \ldots \exists z_n (y \in z_1 \wedge z_1 \in z_2 \wedge \ldots \wedge z_n \in y) \supset A))\,.$$

I did not find these paradoxes in the literature, but let us call this last series of paradoxes the *Quine-Curry-paradoxes* (QCP). Again, the negations of these paradoxes are not always **CL**-theorems but formulas of the form $\exists x \forall y (y \in x \equiv (\exists z_1 \ldots \exists z_n (y \in z_1 \wedge z_1 \in z_2 \wedge \ldots \wedge z_n \in y) \supset A)) \supset A$ are **CL**-theorems. This means that stating the existence of all of these sets trivializes a **CL**-based set theory. All of the former paradoxes can be considered as instances of (QCP).

[2] I sometimes use a simplified notation for axioms and axiom schemas: the formulas I write represent the axioms that are the universally closed versions of the written formulas.

[3] A **L**-theorem is a formula A such that $\vdash_{\mathbf{L}} A$.

3.2 ZFC-set theory

As explained in the first section, mathematicians have proposed several solutions for the problems of naive set theory. The most famous and most popular (but probably not the best) solution is **ZFC**-set theory.

ZFC-set theory is the theory $\langle \mathsf{EXT} + \mathsf{UNION} + \mathsf{POWER} + \mathsf{INF} + \mathsf{REPL} + \mathsf{FOUND} + \mathsf{CHOICE}, \mathbf{CL}\rangle$, where EXT, UNION, POWER, INF, REPL, FOUND, and CHOICE are defined as follows (I follow the rather elegant formulation of the axioms that is found in the Metamath project, cf. page 3):

EXT	$\forall x(x \in y \equiv x \in z) \supset y = z$
REPL	$\forall w \exists y \forall z(\forall v A \supset z = y) \supset \exists y \forall z(z \in y \equiv \exists w(w \in x \wedge \forall v A))$
UNION	$\exists y \forall z(\exists u(u \in x \wedge z \in u) \supset z \in y)$
POWER	$\exists y \forall z(\forall w(w \in z \supset w \in x) \supset z \in y)$
INF	$\exists y(x \in y \wedge \forall z(z \in y \supset \exists w(z \in w \wedge w \in y)))$
FOUND	$\exists y(y \in x) \supset \exists y(y \in x \wedge \forall z(z \in x \supset \neg z \in y))$
CHOICE	$\exists y \forall z \forall w((z \in w \wedge w \in x) \supset$ $\exists v \forall u(\exists t((u \in w \wedge w \in t) \wedge (u \in t \wedge t \in y)) \equiv u = v))$

4 Maximally Consistent Comprehension Set Theory

The first adaptive set theory, *Maximally Consistent Comprehension Set Theory* (**MCC**), is a theory that selects a maximal consistent set of instances of the comprehension schema. Its theorems are the **CL**-consequences of the consistent selection together with extensionality. I start by presenting the underlying adaptive logic **AM**.

4.1 The adaptive logic AM

I first define the logic **M**, which will serve as the **LLL** of the logic **AM**. It is a rather simplistic modal logic. It can be characterized by means of only two possible worlds, but here I give an even simpler characterization. The language of **M** is the **CL**-language to which a $\lozenge$-symbol is added with to restrictions: $\lozenge$-symbols are not nested and no $\lozenge$ occurs in the scope of a quantifier. For example, where A is a **CL**-formula, A, $\lozenge A$, $\neg\lozenge A$, and $\lozenge A \vee \neg\lozenge\neg A$ are **M**-formulas, but $\lozenge\lozenge A$ and $\forall x \lozenge A$ are not **M**-formulas. $\mathcal{W}$ is the set of closed formulas of **CL** and $\mathcal{W}_{\mathbf{M}}$ is the set of closed formulas of **M**.

Let a **M**-model be defined as a triple $\langle v, D, \Psi\rangle$, where $\langle v, D\rangle$ is a **CL**-model and Ψ is a set of **M**-formulas. All **CL**-formulas get the same truth values in a **M**-model $\langle v, D, \Psi\rangle$ as in the corresponding **CL**-model $\langle v, D\rangle$. Where $M = \langle v, D, \Psi\rangle$, $M \models \lozenge A$ iff $A \in \Psi$ or A is true in $\langle v, D, \Psi\rangle$. For the formulas A that have another logical symbol than $\lozenge$ as their outmost connective, $M \vdash A$ is recursively as usual (for example $M \vdash A \vee B$ iff $M \vdash A$ or $M \vdash B$) Finally, we can define the consequence relation $\Gamma \vDash_{\mathbf{M}} A$ iff $M \models A$ for every **M**-model such that $M \models \Gamma$.

Proof theoretically, this logic can be characterized by adding the axiom $A \supset \lozenge A$ to an axiomatic characterization of classical logic, restricting it to the formulas of the language[4].

The adaptive logic **AM** is defined as the triple $\langle \mathbf{M}, \{\Diamond A \wedge \neg A \mid A \in \mathcal{W}_{\mathbf{M}}\}, \text{Reliability}\rangle$. The abnormalities are formulas of the form $\Diamond A \wedge \neg A$, so they enable the conditional derivation of A from $\Diamond A$.

4.2 The tentative theory WMCC

I first present a tentative adaptive set theory, which will turn out to be too weak.

Where $\mathsf{COMP}_{\Diamond}$ is the axiom schema

$$\mathsf{COMP}_{\Diamond} \quad \Diamond \exists x \forall y (y \in x \equiv A(y))\,,$$

the tentative adaptive set theory **WMCC** is defined as follows:

Definition 3 $\mathbf{WMCC} = \langle \mathsf{COMP}_{\Diamond} + \mathsf{EXT}, \mathbf{AM}\rangle$

In this theory, one is able to derive instances A of the normal comprehension schema COMP (without the modality) on the condition that $\Diamond A \wedge \neg A$ is not derivable. If the negation of an instance of COMP or the disjunction of such negations is **CL**-derivable (and therefore also **M**-derivable) and there is no smaller disjunction derivable of this form, the involved instances are paradoxical and the corresponding abnormalities are unreliable. This causes the marking of the lines on which the $\Diamond$-free versions of the problematic axioms and their consequences are derived. Because they are derived on marked lines (that are bound to stay marked in every extension of the proof), these formulas do not count as **AL**-consequences of the axioms. The instance of comprehension that states the existence of the Russell set is a typical example of a problematic instance and the corresponding abnormality is unreliable. The theorems of this simple tentative definition of an adaptive set theory would be the **CL**-consequences of the extensionality axiom together with all the unproblematic instances of comprehension.

This theory is non-trivial. The set of theorems of the theory $\langle \mathsf{COMP}_{\Diamond} + \mathsf{EXT}, \mathbf{M}\rangle$ is not trivial, because there is a finite model for the axioms, more particularly the (finite) **M**-model $M = \langle v, D, \Psi\rangle$, where $\langle v, D\rangle$ is some arbitrary **CL**-model and Ψ is the set of all instances of COMP. By Theorem 1, we know then that also the adaptive theory **WMCC** is non-trivial.

Given that Ψ is obviously an infinite set, one may object that the **M**-model M is not really finite. It is true that there is an infinite aspect to it, but the domain is finite and one does not need to rely on theoretic properties of infinite sets to prove that this model is a model for the theory. Compare this to **ZFC**. We can only give infinite models of **ZFC**, for example the cumulative hierarchy. Defining this model and proving that it is a model for the axioms, requires a theory at least as strong as **ZFC** itself. Suppose **ZFC** would be inconsistent, then the model would be incoherent. Hence, the existence of this kind of models does not give us arguments for the consistency of the theory. The **M**-model presented here, on the other hand, is presentable without strong theoretic instruments. To see this, consider the fact that the elements of Ψ

[4] Proving completeness for this proof theory is safely left to the reader in view of the fact that **M** can be translated into **CL** by translating every (sub)formula $\Diamond A$ into $A \vee p_A$, where p_A is an atomic propositional letter.

could be seen as empty strings. The only requirement is that these strings are exactly the instances of COMP. Defining the (uninteresting) model M for $\langle \mathsf{COMP}_\Diamond + \mathsf{EXT}, \mathbf{M}\rangle$ is therefore not harder than defining the comprehension axiom schema itself.

This way of making naive set theory non-trivial could be quite sensible and applicable to many other potentially inconsistent axiomatic systems. However, there is a problem. Suppose A is a harmless instance of the comprehension axiom schema (for example the instance stating the existence of the empty set). Consider the following proof, in which B abbreviates $\exists x\forall y(y \in x \equiv (y \in y \supset \neg A))$. Slightly abusing notation, I introduce an auxiliary constant in line 5 (in a natural deduction style). This is harmless, as the constant disappears in lines 9-11.

1	$\Diamond A$	$\emptyset$	PREM	
2	A	$\{\Diamond A \wedge \neg A\}$	RC; 1	$\checkmark$
3	$\Diamond\exists x\forall y(y \in x \equiv (y \in y \supset \neg A))$	$\emptyset$	PREM	
4	$\exists x\forall y(y \in x \equiv (y \in y \supset \neg A))$	$\{\Diamond B \wedge \neg B\}$	RC; 3	$\checkmark$
5	$\forall y(y \in o \equiv (y \in y \supset \neg A))$	$\{\Diamond B \wedge \neg B\}$	RU; 4	$\checkmark$
6	$(o \in o \equiv (o \in o \supset \neg A))$	$\{\Diamond B \wedge \neg B\}$	RU; 5	$\checkmark$
7	$o \in o \supset \neg A$	$\{\Diamond B \wedge \neg B\}$	RU; 6	$\checkmark$
8	$o \in o$	$\{\Diamond B \wedge \neg B\}$	RU; 6, 7	$\checkmark$
9	$\neg A$	$\{\Diamond B \wedge \neg B\}$	RU; 8	$\checkmark$
10	$A \wedge \neg A$	$\{\Diamond A \wedge \neg A, \Diamond B \wedge \neg B\}$	RU; 2,9	$\checkmark$
11	$(\Diamond A \wedge \neg A) \vee (\Diamond B \wedge \neg B)$	$\emptyset$	RU; 1,3	

At the end of this proof all conditional lines are marked and there is no extension of the proof in which the lines can become unmarked, for $\Diamond B \wedge \neg B$ nor $\Diamond A \wedge \neg A$ are themselves unconditionally derivable (by the harmlessness of A). Nevertheless, we did not put any conditions on the axiom A. Hence, this can be done for every sensible instance of the comprehension axiom, which results in the complete idleness and uselessness of our tentative adaptive set theory.

Obviously comprehension instances of the form $\exists x\forall y(y \in x \equiv (y \in y \supset \neg A))$ are responsible for this behavior. However, by itself an axiom of this type does not lead to triviality. It only leads to problems when combined with another comprehension instance A, which makes the abnormality corresponding to A unnecessarily unreliable.

What is special about comprehension instances of the form $\exists x\forall y(y \in x \equiv (y \in y \supset \neg A))$ that they cause this problem? In other words, what kind of reliable abnormalities do we need to make unreliable in order to prevent them from infecting the other axioms? Indeed, if we could make $\Diamond B \wedge \neg B$ unreliable by itself, the abnormalities in its weakening $(\Diamond A \wedge \neg A) \vee (\Diamond B \wedge \neg B)$ are no longer considered unreliable.

4.3 Solving the weakness of WMCC

One elegant solution concerns the observation that comprehension instances like the Curry paradox can allow for the derivation of arbitrary formulas independent of the truth of these formulas, simply by introducing the arbitrary formulas in the comprehension axioms themselves. Although the negation of axioms of this form is not always derivable, the negation of a very similar comprehension instance is derivable. $\neg A$ follows from $\exists x\forall y(y \in x \equiv (y \in y \supset \neg A))$ and therefore, if we substitute A by $\top$, then the negation of the result of the substitution is derivable ($\vdash_{\mathbf{CL}} \neg\exists x\forall y(y \in x \equiv (y \in y \supset \neg\top))$). The fact that a very similar comprehension instance is explicitly **CL**-inconsistent, justifies the claim that we better do not trust the instance itself.

Let Insts(A), where A is an axiom schema, be the set of instances of A. Let $\mathrm{Var}_{\mathsf{A}}(A)$, where $A \in$ Insts(A), be the set of formulas B such that $B \in$ Insts(A) and B is the result of substituting in A one or more open or closed subformulas of A by $\top$ or $\bot$. For example, the following set is the set $\mathrm{Var}_{\exists x\forall y(y\in x\equiv A(y))}(\exists x\forall y(y \in x \equiv (y \in y \supset \exists z(z \in y))))$:

$$\begin{gathered}
\{\exists x\forall y(y \in x \equiv \bot), \exists x\forall y(y \in x \equiv \top), \\
\exists x\forall y(y \in x \equiv (y \in y \supset \bot)), \exists x\forall y(y \in x \equiv (y \in y \supset \top)), \\
\exists x\forall y(y \in x \equiv (y \in y \supset \exists z\bot)), \exists x\forall y(y \in x \equiv (y \in y \supset \exists z\top)), \\
\exists x\forall y(y \in x \equiv (\bot \supset \exists z(z \in y))), \exists x\forall y(y \in x \equiv (\top \supset \exists z(z \in y))), \\
\exists x\forall y(y \in x \equiv (\top \supset \exists z\bot)), \exists x\forall y(y \in x \equiv (\top \supset \exists z\top)), \\
\exists x\forall y(y \in x \equiv (\bot \supset \exists z\bot)), \exists x\forall y(y \in x \equiv (\bot \supset \exists z\top)), \\
\exists x\forall y(y \in x \equiv (\top \supset \bot)), \exists x\forall y(y \in x \equiv (\top \supset \top)), \\
\exists x\forall y(y \in x \equiv (\bot \supset \bot)), \exists x\forall y(y \in x \equiv (\bot \supset \top))\}
\end{gathered}$$

The idea is now the following: if $\lozenge B \wedge \neg B$ is unreliable, where $B \in \mathrm{Var}_{\mathsf{A}}(A)$ for some formula $A \in$ Insts(A), then also consider the abnormality $\lozenge A \wedge \neg A$ as unreliable.

The effect of this is that, where Δ is a set of abnormalities, $\lozenge B \wedge \neg B$ is unconditionally derivable, $B \in \mathrm{Var}_{\mathsf{COMP}}(A)$ and $\lozenge A \wedge \neg A \in \Delta$, even if $Dab(\Delta)$ is a minimal Dab-consequence of our theory, Δ does not contribute to the set of unreliable abnormalities. By means of this mechanism, all lines on which Quine-Curry-paradoxical comprehension instances are introduced become marked and our problem is solved. There are different ways to formalize this idea. I present the three most important ones.

4.3.1 Explicitly making abnormalities unreliable by adding axioms: the theory $\mathbf{MCC^a}$

One solution for this problem is making the infecting axioms explicitly derivable. This can be done by adding the extra axiom schema $(\lozenge A \wedge \neg A) \supset (\lozenge B \wedge \neg B)$, for every $A \in \mathrm{Var}_{\mathsf{COMP}}(B)$, where $B \in$ Insts(COMP) (call this schema SPR). The resulting adaptive theory is defined by

Definition 4 $\mathbf{MMC^a} = \langle \mathsf{COMP}_{\lozenge} + \mathsf{EXT} + \mathsf{SPR}, \mathbf{AM}\rangle$.

By adding this schema, the innocent abnormalities will no longer be a part of a minimal disjunction of abnormalities, because the infecting abnormality is unconditionally derivable and therefore forms a shorter Dab-consequence. Hence, the innocent abnormalities are no longer unreliable and the problem is solved.

One might, however, object against the fact that axioms are added without a motivation for their truth, only to better localize the problems of the original theory.

4.3.2 A different adaptive strategy: the theory $\mathbf{MCC^s}$

A second solution, without adding extra axioms, involves changing the adaptive strategy. Basically, we define the same set of unreliable abnormalities as in proposal 1, without adding extra axioms. This requires some technicalities. We define the alternative Reliability strategy as follows. Let an *alternative for an abnormality* $\Diamond A \wedge \neg A$ be an abnormality $\Diamond B \wedge \neg B$ such that $A \in \{B\} \cup \mathrm{Var}_{\mathsf{B}}(B)$, where $\Diamond \mathsf{B}$ is an axiom schema and $B \in \mathrm{Insts}(\mathsf{B})$. Let an *alternative for a set of abnormalities* Δ be the smallest set Δ' such that for each member of Δ there is in Δ' exactly one alternative. Where ξ_s is the set of all sets Δ such that $Dab(\Delta)$ is unconditionally derived at stage s, define ξ^a_s as the set containing all the alternatives for members of ξ_s. Let ξ'^a_s be the set of all members of ξ^a_s for which there is no proper subset in ξ^a_s. Let $U^a_s(\Gamma)$ be the set $\bigcup \xi'^a_s$.

Analogously define $U^a(\Gamma)$: where $\xi(\Gamma)$ is the set of sets such that $\Gamma \vDash_{\mathbf{LLL}} Dab(\Delta)$, define $\xi^a(\Gamma)$ as the set containing all the alternatives for members of $\xi(\Gamma)$. Let $\xi'^a(\Gamma)$ be the set of all members of $\xi^a(\Gamma)$ for which there is no proper subset in $\xi^a(\Gamma)$. Let $U^a(\Gamma)$ be the set $\bigcup \xi'^a(\Gamma)$. Replace in the definition of the normal semantic and proof theoretic Reliability strategy $U(\Gamma)$ by $U^a(\Gamma)$ resp. $U_s(\Gamma)$ by $U^a_s(\Gamma)$ to obtain the *Reliability strategy for Axiom schemata*. The underlying logic for our adaptive theory would then be the adaptive logic $\mathbf{AM^a} = \langle \mathbf{M}, \{\Diamond A \wedge \neg A \mid A \in \mathcal{W}_{\mathbf{M}}\}, \text{Reliability for Axiom schemata}\rangle$.

The adaptive theory is then defined by

Definition 5 $\mathbf{MCC^s} = \langle \mathsf{COMP}_{\Diamond} + \mathsf{EXT}, \mathbf{AM}^a\rangle$.

This solution avoids adding new weakly motivated axioms, but involves a non-standard strategy. This strategy is non-standard but by no means ad hoc or idiosyncratic. A variant of the strategy can be applied to every case where the unreliability of abnormalities needs to be spread to some specific set of similar abnormalities in order to avoid them from making too many abnormalities unreliable. For different cases just replace $\mathrm{Var}_{\Diamond \mathsf{A} \wedge \neg \mathsf{A}}$ in the definitions of $U^a_s(\Gamma)$ en $U^a(\Gamma)$, by the appropriate set of abnormalities similar to A if A is the abnormality that has to be made unreliable.

4.3.3 Making the abnormalities weaker: the theory $\mathbf{MCC^b}$

The last (and probably preferable but more technical) solution involves making the abnormalities more complex.

The underlying logic for the resulting adaptive theory is the adaptive logic $\mathbf{AM^b} = \langle \mathbf{M}, \{\Diamond A \wedge \neg \bigwedge(\{A\} \cup \mathrm{Var}_{\mathsf{A}}(A)) \mid A \in \mathrm{Insts}(\mathsf{A}); \Diamond \mathsf{A} \text{ is an axiom schema}\},$ Reliability$\rangle$.

The adaptive theory is then defined by

Definition 6 $\mathbf{MCC^b} = \langle \mathsf{COMP}_{\Diamond} + \mathsf{EXT}, \mathbf{AM^b}\rangle$.

This solution works because the abnormalities are a lot weaker. From every $\Diamond A$ (an instance of an axiom schema A) and $\neg B$ where $B \in \mathrm{Var}_{\mathsf{A}}(A)$, we can $\mathbf{M}$-derive $\Diamond A \wedge \neg \bigwedge(\{A\} \cup \mathrm{Var}_{\mathsf{A}}(A))$. Hence, the inference from $\Diamond A$ to A (on the condition that exactly contains this abnormality) is blocked, independent of the question whether $\neg A$ is $\mathbf{M}$-derivable.

Let me illustrate this with an example. Consider the adaptive proof above, where an arbitrary comprehension instance A was marked, because the disjunction of $\Diamond A \wedge \neg A$ and $\Diamond B \wedge \neg B$, where $B = \exists x \forall y (y \in x \equiv (y \in y \supset \neg A))$, was unconditionally derivable whereas $\Diamond B \wedge B$ was not. In the logic $\mathbf{AM^b}$, however, A is derived on a weaker condition: the singleton containing the abnormality $\Diamond A \wedge \neg \bigwedge(\{A\} \cup \mathrm{Var}_{\mathsf{COMP}}(A))$. We know that $\exists x \forall y(y \in x \equiv (y \in y \supset \neg\top))$ is an element of $\mathrm{Var}_{\mathsf{COMP}}(B)$. Moreover, one can easily prove that $\exists x \forall y(y \in x \equiv (y \in y \supset \neg\top))$ **CL**-entails $\bot$, whence $\neg\exists x \forall y(y \in x \equiv (y \in y \supset \neg\top))$ is a **M**-theorem and therefore $\Diamond B \wedge \neg \bigwedge(\{B\} \cup \mathrm{Var}_{\mathsf{COMP}}(B))$ is unconditionally derivable in an $\mathbf{AM^b}$-proof. As a consequence the disjunction of $\Diamond A \wedge \neg \bigwedge(\{A\} \cup \mathrm{Var}_{\mathsf{COMP}}(A))$ and $\Diamond B \wedge \neg \bigwedge(\{B\} \cup \mathrm{Var}_{\mathsf{COMP}}(B))$ is not a minimal *Dab*-formula and so $\Diamond A \wedge \neg \bigwedge(\{A\} \cup \mathrm{Var}_{\mathsf{COMP}}(A))$ is not an unreliable abnormality. Consequently, B does not prevent A from being derivable, which is what we wanted.

4.3.4 Can we warrant that the presented solutions solve all problems?

All three solutions solve the problems caused by all Curry-like paradoxes concerning infective abnormalities. Moreover one can easily prove that these solutions do not affect the guaranteed non-triviality of the adaptive theory.

It is not entirely excluded that there would exist yet other infections that make the adaptive theory too weak and are not solved by the presented solutions. However this is rather unlikely given the fact that the presented solutions are generally applicable solutions for problems of infective abnormalities.

5 Non-trivially rich universal set theory

The second approach (elaborated in detail in [24]) to adaptive set theory is fundamentally different from the first. Whereas for the first theory we were interested in a maximally large consistent sub-theory of naive set theory, for the second theory, we explicitly accept all the instances of comprehension. Of course this must result in an inconsistent set theory and in order to avoid triviality the underlying logic must be paraconsistent and have an implication or equivalence weaker than the classical variants.

Many systems have been proposed to give a non-trivial version of the comprehension or the abstraction axiom. Most of the proposed set theories have a relevant underlying logic that does not validate the rule of contraction to avoid Curry's paradox (cf. [6], [7], [28] and [29]). One can also use other logics without contraction (cf. [16], [10], [21] and [31]) or a weak paraconsistent logic like **LP** (cf. [17] and [20]).

All set theories of this type are universal (they use full comprehension), but are often too weak to serve as a foundation of mathematics or they are strong enough to formalize basic arithmetic, in which case they have the same problem as **ZFC** with respect to proving non-triviality.

I start by semantically introducing the logic **EL** that adds a special equivalence symbol $\equiv\!\!\!\!\!\!\!\!\rightarrow$ to classical logic. The equivalence will be non-commutative: classical in the left-right direction, but it shows a glut in the right-left direction. It is the only primitive symbol that does not behave classically[5], but a paraconsistent negation can be defined. The logic is four valued. The special

[5] As **EL** contains full classical logic, a material equivalence connective $\equiv$ is also definable, cf. the list of definitions in the next subsection.

equivalence will be used in the comprehension axiom schema. This logic will function as the **LLL** of the **AL** that will be the underlying logic of our set theory. By means of this logic, I shall define a tentative adaptive set theory which will turn out to be too weak. Finally, I present two adaptive set theories that solve the weakness problem of the tentative theory.

5.1 The logic EL

Let the language of **EL** contain the logical symbols $\neg$, $\vee$, $\equiv$, $\forall$, $\perp$ and $=$, a set of constants $\mathcal{C}$, a set of variables $\mathcal{V} = \{x, y, z, \ldots\}$, and sets of predicates $\mathcal{P}^0$, $\mathcal{P}^1$, ..., where $\mathcal{P}^r$ contains the r-ary predicates. All the sets in the previous sentence are supposed to be pairwise disjoint. Formulas are constructed in the usual way with the restriction that $\equiv$ cannot be nested. Let $\mathcal{F}_{\mathbf{EL}}$ and $\mathcal{W}_{\mathbf{EL}}$ denote respectively the set of formulas and the set of closed formulas of $\mathcal{L}_{\mathbf{EL}}$. Let $\mathbb{P} \subset \mathcal{F}_{\mathbf{EL}}$ denote the set of primitive formulas and let $\mathbb{P}^{\neg} = \{\neg A \mid A \in \mathbb{P}\}$.

In order to simplify the characterization of the semantics, I introduce a pseudo-language—this method also occurs in [4]. Let $\mathcal{O}$ be a set of *pseudo-constants*; $\mathcal{O}$ should have at least the cardinality of your largest set and its elements do not occur in premises or conclusions. The pseudo-language ${}^{+}\mathcal{L}_{\mathbf{EL}}$ is defined by adding the members of $\mathcal{O}$ to the constants. Let ${}^{+}\mathcal{F}_{\mathbf{EL}}$ and ${}^{+}\mathcal{W}_{\mathbf{EL}}$ denote respectively the set of formulas and the set of closed formulas of ${}^{+}\mathcal{L}_{\mathbf{EL}}$.

The semantics makes use of four truth values: the values T (can be interpreted as full truth), F (can be interpreted as full falsehood), B (can be interpreted as both true and not true), and D (can be interpreted as both false and not false).

Let, for every $r > 0$, $D^{(r)}$ denote the r-th Cartesian product of D and let $D^{(0)} = \{\emptyset\}$, i.e. a 0-tuple will be identified with $\emptyset$. $\wp(A)$ denotes the powerset of A.

An **EL**-model M (for the language $\mathcal{L}_{\mathbf{EL}}$) is a couple $\langle D, v\rangle$ in which D is a non-empty set and the assignment function v is restricted as follows:

C1.0 v is the union of v_1, v_2 and v_3.
C1.1 $v_1\colon \mathcal{C} \cup \mathcal{O} \to D$ (where v_1 is a surjection)
C1.2 $v_2\colon (\{\mathrm{B}, \mathrm{T}, \mathrm{F}, \mathrm{D}\} \times \mathcal{P}^r) \to \wp(D^{(r)})$ (for every $r \in \mathbb{N}$)
C1.3 For any $\pi \in \mathcal{P}^r$, $\bigcap\{v_2(\mathfrak{V}, \pi) | \mathfrak{V} \in \{\mathrm{B}, \mathrm{T}, \mathrm{F}, \mathrm{D}\}\} = \emptyset$
C1.4 For any $\pi \in \mathcal{P}^r$, $\bigcup\{v_2(\mathfrak{V}, \pi) | \mathfrak{V} \in \{\mathrm{B}, \mathrm{T}, \mathrm{F}, \mathrm{D}\}\} = D^{(r)}$
C1.5 $v_3\colon \{\perp\} \to \{\mathrm{F}, \mathrm{D}\}$

The following clauses define how a model M determines the truth values $\mathcal{V}_M$ formulas receive in that model[6].

C2.1 $\mathcal{V}_M(\pi\alpha_1 \ldots \alpha_r) = \mathfrak{V}$ iff $\langle v(\alpha_1), \ldots, v(\alpha_r)\rangle \in v(\mathfrak{V}, \pi)$ where $\mathfrak{V} \in \{\mathrm{B}, \mathrm{T}, \mathrm{F}, \mathrm{D}\}$.
C2.2a $\mathcal{V}_M(\neg A)$, $\mathcal{V}_M(A \equiv B)$ and $\mathcal{V}_M(A \wedge B)$ are determined according to the truth values in table 1.
C2.2b $\mathcal{V}_M(\perp) = v(\perp)$
C2.3a $\mathcal{V}_M(\forall\gamma A(\gamma)) = \mathrm{T}$ iff ($\mathcal{V}_M(A(\alpha)) = \mathrm{T}$ for all $\alpha \in \mathcal{C} \cup \mathcal{O}$).
C2.3b $\mathcal{V}_M(\forall\gamma A(\gamma)) = \mathrm{B}$ iff ($\mathcal{V}_M(A(\alpha)) \in \{\mathrm{B}, \mathrm{T}\}$ for all $\alpha \in \mathcal{C} \cup \mathcal{O}$ and $\mathcal{V}_M(A(\alpha)) = \mathrm{B}$ for at least one $\alpha \in \mathcal{C} \cup \mathcal{O}$).
C2.3c $\mathcal{V}_M(\forall\gamma A(\gamma)) = \mathrm{F}$ iff ($\mathcal{V}_M(A(\alpha)) = \mathrm{F}$ for at least one $\alpha \in \mathcal{C} \cup \mathcal{O}$).

[6]So, where π is an 0-ary predicate, $\mathcal{V}_M(\pi) = \mathrm{T}$ iff $v(\mathrm{T}, \pi) = \{\emptyset\}$.

$\mathcal{V}_M(A)$ \ $\mathcal{V}_M(B)$	$\mathcal{V}_M(A \Rrightarrow B)$				$\mathcal{V}_M(A \vee B)$				$\mathcal{V}_M(\neg A)$
	B	T	F	D	B	T	F	D	
B	T	T	F	F	B	T	B	B	D
T	T	T	F	F	T	T	T	T	F
F	F	F	T	T	B	T	F	D	T
D	T	T	T	T	B	T	D	D	B

Table 1: Matrices for the propositional symbols.

C2.3d $\mathcal{V}_M(\forall\gamma A(\gamma)) = \text{D}$ iff ($\mathcal{V}_M(A(\alpha)) = \text{D}$ for at least one $\alpha \in \mathcal{C} \cup \mathcal{O}$ and $\mathcal{V}_M(A(\alpha)) \neq \text{F}$ for all $\alpha \in \mathcal{C} \cup \mathcal{O}$).
C2.4a $\mathcal{V}_M(\alpha = \beta) \in \{\text{T}, \text{F}\}$.
C2.4b $\mathcal{V}_M(\alpha = \beta) = \text{T}$ iff $v(\alpha) = v(\beta)$.

Some symbols are defined from the other symbols. The defined symbols function as mere abbreviations of more complex formulas.

D1 $A \supset B =_{\text{df}} \neg A \vee B$
D2 $A \wedge B =_{\text{df}} \neg(\neg A \vee \neg B)$
D3 $A \equiv B =_{\text{df}} (A \supset B) \wedge (B \supset A)$
D4 $\exists\alpha A(\alpha) =_{\text{df}} \neg\forall\alpha\neg A(\alpha)$
D5 $\sim A =_{\text{df}} \neg A \Rrightarrow (A \vee \neg A)$
D6 $\top =_{\text{df}} \neg\bot$

Definition 7 **EL**-*satisfaction. Where $A \in \mathcal{W}_{\mathbf{EL}}$, $\Gamma \subseteq \mathcal{L}_{\mathbf{EL}}$ and $M = \langle v, D\rangle$ is an* **EL**-*model, $M \models A$ iff $\mathcal{V}_M(A) \in \{B, T\}$ and $M \models \Gamma$ iff $M \models A$ for every $A \in \Gamma$.*

Definition 8 **EL**-*consequence. Where $\Gamma \cup \{A\} \subseteq \mathcal{L}_{\mathbf{EL}}$, A is an* **EL**-*consequence of Γ, in symbols $\Gamma \vDash_{\mathbf{EL}} A$, iff $M \models A$, for every* **EL**-*model M such that $M \models \Gamma$.*

Definition 9 **EL**-*equivalence. A is* **EL**-*equivalent to B (abbreviated as $A \approx B$) iff $\{A\} \vDash_{\mathbf{EL}} B$ and $\{B\} \vDash_{\mathbf{EL}} A$.*

One can easily devise a sound and complete proof theory for this logic, but for my current purpose, this is not needed. The reader can check the correctness of the following theorem with the help of tables 2 and 1.

Theorem 2 *The following are important properties of this logic:*

F1 $A \Rrightarrow B \approx (A \supset B) \wedge (B \supset \sim\neg A)$
F2 $\sim\neg A \nvDash_{\mathbf{EL}} A$
F3 $A \wedge \sim A \approx \neg A \Rrightarrow A$
F4 $\neg A \wedge \sim\neg A \approx A \Rrightarrow \neg A$
F5 $\neg A, \sim\neg A \vDash_{\mathbf{EL}} A \Rrightarrow B$

Although the equivalence symbol $\Rrightarrow$ is the only non-classical symbol in this logic, an alternative negation $\sim$ can be defined by means of this equivalence. Because, for every $A \in \mathcal{W}_{\mathbf{EL}}$, there is a $B \in \mathcal{W}_{\mathbf{EL}}$ such that $A, \sim A \nvDash_{\mathbf{EL}} B$,

$\mathcal{V}_M(A)$ \ $\mathcal{V}_M(B)$	$\mathcal{V}_M(A \wedge B)$				$\mathcal{V}_M(\sim A)$	$\mathcal{V}_M(\sim\neg A)$	$\mathcal{V}_M(A \equiv \neg A)$
	B	T	F	D			
B	B	B	F	D	T	T	F
T	B	T	F	D	F	T	F
F	F	F	F	F	T	F	F
D	D	D	F	D	T	T	T

$\mathcal{V}_M(A)$ \ $\mathcal{V}_M(B)$	$\mathcal{V}_M(A \supset B)$				$\mathcal{V}_M(A \wedge \sim A)$	$\mathcal{V}_M(\neg A \equiv A)$	$\mathcal{V}_M(\neg A \wedge \sim\neg A)$
	B	T	F	D			
B	B	T	D	D	B	T	D
T	B	T	F	D	F	F	F
F	T	T	T	T	F	F	F
D	B	T	B	B	D	F	B

Table 2: Truth functionality of some important formulas

$\sim$ is paraconsistent. The logic **EL** however is not paraconsistent because it is explosive (for all $A, B \in \mathcal{W}_{\mathbf{EL}}$, $A, \neg A \vDash_{\mathbf{EL}} B$ holds).

Formulas get one out of four truth values in an **EL**-model. One can easily understand the meaning of the truth values by means of the negation $\sim$. The value T stands for pure truth, F for pure falsity, B for both A and $\sim A$ true and D for A false, $\sim A$ true, but $\sim\neg A$ also true. The fourth unusual value is necessary because the classical negation $\neg$ can occur within the scope of the paraconsistent negation $\sim$. There are no other values needed because the definition of the language does not allow $\sim$ to be nested (remember that it is defined in terms of $\equiv$, which cannot be nested). Every $\sim$-inconsistency boils down to $\sim$-inconsistencies at the level of primitive formulas or on the level of the classical negation of primitive formulas.

5.2 A tentative adaptive version of naive set theory: the theory WMRU

We start again with a simple adaptive set theory that will turn out to be too weak.

Let the language of the set theory be based on the language of **EL**. Restrict the set of predicates so that it only contains the binary predicate $\in$ and augment the language with set terms of the form $\{\alpha \mid A(\alpha)\}$, where $\alpha \in \mathcal{V}$ and $A(\alpha)$ is a formula in which only α occurs free. Formulas of the form $\{\alpha \mid A(\alpha)\} \in \{\beta \mid A(\beta)\}$ are considered as primitive formulas and are therefore elements of $\mathbb{P}$.

For this theory, I use the abstraction axiom schema rather than the comprehension schema[7]. Abstraction is here defined as follows.

$$\textsf{ABS} \qquad \forall x(x \in \{y \mid A(y)\} \equiv A(x))$$

[7] This allows me to easily express the inconsistency of particular sets in the abnormalities. The only difference between comprehension and abstraction is the availability of names for sets. Contextually defining constants also allows this if one uses the comprehension schema, but this requires more technicalities. The comprehension schema is a **CL**-consequence of the abstraction schema.

The adaptive theory I shall present has the adaptive logic **AEL** as its underlying logic. **AEL** is the adaptive logic defined by the following standard format triple: $\langle \mathbf{EL}, \{\exists(A \equiv \neg A) \mid A \in \mathbb{P} \cup \mathbb{P}^{\neg}\}, \text{Reliability}\rangle$ (for the explanation of the abnormalities, see below).

Of course one could use another lower limit logic, and there may even be much stronger or more sensible logics available, but this logic is sufficiently strong, quite elegant, and there exists a finite **EL**-model for $\mathsf{ABS} + \mathsf{EXT}$ (see below). I am not aware of any existing logic with similar or better properties.

The strategy of **AEL** is Reliability and the abnormalities are existentially closed sentences that express the (existentially closed) inconsistency of primitive formulas or of their classical negations (such as $\alpha \in \beta$, $\alpha = \beta$, $\neg\alpha \in \beta$, $\neg\alpha = \beta$, where α and β are variables or set terms)[8].

We come to the definition of the provisional theory **WMRU**.

Definition 10 **WMRU**$= \langle \mathsf{EXT} + \mathsf{ABS}, \mathbf{AEL}\rangle$

This theory is provably non-trivial. The reason for this is simple. There is a finite **EL**-model for EXT+ABS, viz. the model $\langle D, v\rangle$, where $D = \{o\}$, where o is an arbitrary object, $\langle o, o\rangle \in v_{\mathrm{D}}(\in)$ but $v_{\mathrm{B}}(\in) \cup v_{\mathrm{T}}(\in) \cup v_{\mathrm{F}}(\in) = \emptyset$, and finally $v(\{y|A(y)\}) = o$ for every $A(y)$. The truth functionality of $\neg$ and $\sim$ and the definition of **EL**-satisfaction ensure that for every $\alpha, \beta \in \mathcal{C} \cup \mathcal{O}$, $M \models \neg\beta \in \alpha$ and $M \models {\sim}\neg\beta \in \alpha$. F5 warrants that $M \models \beta \in \alpha \equiv A$, for every formula A. Using the definition of $\forall$, one obtains $M \models \mathsf{ABS}$. Because $M \models \forall x \forall y x = y$ (D is a singleton), also $M \models \mathsf{EXT}$. So there is a finite **EL**-model for the axioms. So the **EL**-consequence set of $\mathsf{EXT} + \mathsf{ABS}$ is non-trivial, whence the **AEL** consequence is also non-trivial, in view of Theorem 1. Consequently, **WMRU** is a non-trivial set theory.

This theory is also universal in a rather strong sense. For every formula $A(\alpha)$, (1) the set $a = \{x \mid A(x)\}$, for which it holds that $y \in a \equiv A(y)$, exists in this theory (and this existence is provable as a theorem) and (2) for every constant α and every formula $A(\alpha)$, $A(\alpha)$ **EL**-entails $\alpha \tilde{\in} \{x \mid A(x)\}$, where $\tilde{\in}$ is defined by $\alpha \tilde{\in} \beta =_{df} {\sim}\neg\alpha \in \beta$. $\tilde{\in}$ is an interesting weak paraconsistent membership relation. One may interpret this membership relation as follows: α is weakly a member of β, denoted by $\alpha \tilde{\in} \beta$, iff α is in β but it may also be in the complement of β. This is a weaker notion than full membership $\in$, which holds when an object is in a set and not in the complement of the set. Many other universal alternative set theories only have property (1) and not property (2), because they do not enable the possibility of defining a weak set membership predicate like $\tilde{\in}$.

In which sense does the adaptive logic add theorems to the theory $\langle \mathsf{EXT} + \mathsf{ABS}, \mathbf{EL}\rangle$? This is quite intuitive: in the logic **EL**, the inference from ${\sim}\neg A$ to A is not valid. If this inference would be generally valid, a theory containing the abstraction axioms would be trivial. Nevertheless, not all instances of this inference rule are problematic. With the logic **EL**, we are able only to derive ${\sim}\neg\emptyset \in \{x \mid x = \emptyset\}$. It is evidently harmless to strengthen this to $\emptyset \in \{x \mid x = \emptyset\}$. The idea behind the adaptive theory is to block only the instances of this rule that are problematic and to allow all the harmless ones. Ideally we would obtain an adaptive theory which is as strong as **ZFC**, if **ZFC** is consistent. If **ZFC** is consistent, its axioms are harmless.

[8] Remember that $\neg A \equiv A$ is **EL**-equivalent to $A \wedge {\sim}A$.

For this theory, the only relevant abnormalities are of the form $\alpha \in \beta \equiv \neg\alpha \in \beta$ or of the form $\exists\gamma(\gamma \in \beta \equiv \neg\gamma \in \beta)$, where α and β are set terms of the form $\{\gamma \mid A(\gamma)\}$. This is due to the fact that from every abstraction axiom of the form $\forall x(x \in \{y \mid A(y)\} \equiv A(x))$ the following is **EL**-derivable: $\forall x(x \in \{y \mid A(y)\} \equiv A(x)) \vee B$, where B is an abnormality of the mentioned form. Because $\equiv$ is only meant to occur in the abstraction axioms instead of the classical equivalence, the mentioned fact about **EL**, enables all **CL**-consequences of all abstraction axioms. Note that I shall sometimes write abnormalities as $\sim\neg A \wedge \neg A$ instead of the **EL**-equivalent $A \equiv \neg A$, where this is clearer.

This theory does not do what it intuitively is expected to do; it does not add anything to the **EL**-consequences of the axioms. Suppose C is a useful consequence that the adaptive logic should add to the **EL**-consequences of $\mathsf{ABS} + \mathsf{EXT}$. Suppose it is derived on the condition $\{A\}$, where A is an **AEL**-abnormality. In the proof below, let a abbreviate $\{y \mid y \in y \supset A\}$.

1	C	$\{A\}$	RC	✓
2	$\forall x(x \in a \equiv (x \in x \supset A))$	$\emptyset$	PREM	
3	$a \in a \equiv (a \in a \supset A)$	$\emptyset$	2; RU	
4	$a \in a \supset A$	$\emptyset$	3; RU	
5	$\sim\neg(a \in a)$	$\emptyset$	4; RU	
6	$a \in a$	$\{\sim\neg a \in a \wedge \neg a \in a\}$	5; RC	✓
7	A	$\{\sim\neg a \in a \wedge \neg a \in a\}$	6; RC	✓
8	$A \vee (\sim\neg a \in a \wedge \neg a \in a)$	$\emptyset$	2; RU	

5.3 Solving WMRU's weakness: the theories MRU and MRU$^+$

Solutions similar to the ones provided for solving the weakness of **WMCC** are applicable to this problem. One could for example add the following axiom schema to the axioms: $\forall x((x \in \alpha \equiv \neg x \in \alpha) \supset (x \in \beta \equiv \neg x \in \beta))$, where α abbreviates $\{y \mid B(y)\}$, α abbreviates $\{y \mid A(y)\}$ and $A(\alpha) \in \mathrm{Var}_{\mathsf{ABS}}(B(\alpha))$. This solves the problem because for all of the sets of the infecting type, there is a variant a in which $\top$ or $\bot$ is substituted such that $\gamma \in a \equiv \neg\gamma \in a$ occurs in a minimal *Dab*-consequence of the axioms, for some constant γ. This is the case because every ABS-instance that states the existence of a Quine-paradoxical set, will allow us to **EL**-derive a disjunction of $\sim$-contradictions. In its turn, this disjunction of $\sim$-contradictions entails a disjunction of abnormalities of the given form.

The other solutions I have presented to solve the weakness of **MCC** also apply to **WMRU**. But there are two other solutions that allow for an even more specific localization of the abnormalities.

5.3.1 Adding the axiom of foundation: the theory MRU

This approach brings us closer to **ZFC**. Consider the axiom FAF:

$$\mathsf{FAF} \qquad \bigvee\{\bigwedge\{\neg x_i \in x_j \mid i \leq n\} \mid j \leq n\} \qquad \text{where } n \in \mathbb{N}\,.$$

The adaptive set theory is now defined by the following:

Definition 11 **MRU**$=\langle \mathsf{EXT} + \mathsf{ABS} + \mathsf{FOUND} + \mathsf{FAF} + \mathsf{CHOICE}, \mathbf{AEL}\rangle$.

Actually, FAF is a finitistic version of the axiom of foundation. The axiom of foundation FOUND and its finitistic version does not state the existence of some set, but rather the non existence of loops of sets. For every Russell-Curry-Quine set b we can easily show that there is a loop of sets for which $a_1 \tilde{\in} a_2, a_2 \tilde{\in} a_n$, and finally $a_n \tilde{\in} a_1$ are **EL**-consequences of ABS, such that b is one of $a_1, \ldots, a_n$. If we would have an axiom that states that for every sequence of sets $a_1, \ldots, a_n, a_1$ there is at least one $i < n$ such that $\neg a_i \in a_{i+1}$ or $\neg a_n \in a_1$, then the disjunction of abnormalities $(a_1 \in a_2 \mathrel{\equiv\!\!\!\!\neg} \neg a_1 \in a_2) \vee \ldots \vee (a_{n-1} \in a_n \mathrel{\equiv\!\!\!\!\neg} \neg a_{n-1} \in a_n) \vee (a_n \in a_1 \mathrel{\equiv\!\!\!\!\neg} \neg a_n \in a_1)$ would be derivable for every Russell-Curry-Quine set a_1. This disjunction makes every Russell-Curry-Quine set unreliable which is exactly what was required.

FAF is simply the result of instantiating the universal quantifier $\forall x$ in FOUND with every possible finite set, i.e. with the sets $\{x_0\}$, $\{x_0, x_1\}$, $\{x_0, x_1, x_2\}$ and so on, for every possible sets x_0, x_1, x_2, and so on (and afterwards simplifying the obtained expression). Hence, in combination with the POWER-axiom and the UNION-axiom of **ZFC**, FAF is a **CL**-consequence of FOUND (but FAF is not an **EL**-consequence of FOUND). I use FAF in addition to full FOUND because this allows for the most specific *Dab*-consequences. In order to obtain a set theory that is as close to **ZFC** as possible, all axioms of **ZFC** are added to the adaptive set theory that are not falsified by the finite **EL**-model M for $\mathsf{EXT} + \mathsf{ABS}$ mentioned above.

This theory may very well validate all **ZFC**-theorems, if **ZFC** is not trivial. Indeed, the **ZFC**-theorems that are not **CL**-consequences of $\mathsf{EXT} + \mathsf{FOUND} + \mathsf{FAF} + \mathsf{CHOICE}$, are consequences of a subset of the instances of the abstraction axiom schema. If **ZFC** is non-trivial, then this subset of abstraction axioms is also unproblematic. Conditionally, we are able to derive the full classical variants of the axioms under consideration. If **ZFC** is not trivial, these classical variants will not lead to triviality. Consequently, a disjunction of the relevant **ZFC**-abnormalities will not be derivable. So the relevant abnormalities will not be unreliable. However, there might exist mechanisms similar to the problems that made our first attempt useless. These mechanisms may result in the infection of harmless **ZFC**-sets by paradoxical sets or even by reasonable non-**ZFC** sets that are internally consistent and meaningful but that cause problems in combination with certain **ZFC**-sets.

5.3.2 Expressing the preference for ZFC: the theory MRU$^+$

In this paper we are interested in providing a foundation of classical mathematics. There might be harmless selections of abstraction axioms that are philosophically more interesting *and* strong enough to formalize results from most of classical mathematics. However, for now, **ZFC** seems to be the best studied theory with respect to formalizing classical mathematical results. If we want our theory to inherit the nice foundational or formalizing properties of **ZFC**, we need a method to formally prefer the **ZFC**-axioms to other axioms, whenever they are incoherent with **ZFC** and thus risk excluding each other.

One may formalize this preference for **ZFC**-axioms and sets by means of a prioritized adaptive logic (cf. [3, 25, 22, 26]). However, this would involve too many technicalities for the present paper. There is another way to obtain the same result. I present the basic idea behind this method.

The crucial aspect lies in distinguishing what I call a **ZFC**-abnormality[9]. The precise way to do this involves many technicalities that are beyond the scope of this paper. I give an informal sketch. Consider the axioms of **MRU**. For every set $\{y \mid A(y)\}$ that exists according to the axioms of **ZFC**, one can unconditionally derive an ABS-instance $\forall x(x \in \{y \in A(y)\} \equiv\!\!\!\!/\; A(x))$ in **MRU**. From this ABS-instance, **MRU** allows us to derive $\forall x(x \in \{y \in A(y)\} \equiv A(x))$ on the condition $\{\exists x(x \in \{y \in A(y)\} \equiv\!\!\!\!/\; \neg A(x))\}$. Let a **ZFC**-abnormality be the abnormality in such a condition.

Let $\Omega_{\mathbf{ZFC}}$ denote the set of **ZFC**-abnormalities. Define the logic **AEL**$'$ as the logic that is exactly like **AEL**, except that the set of abnormalities is restricted to $\Omega_{\mathbf{ZFC}}$. Finally, define the logic $\mathbf{AEL}^{\mathbf{ZFC}}$ as the combined adaptive logic for which $Cn_{\mathbf{AEL}^{\mathbf{ZFC}}}(\Gamma) = Cn_{\mathbf{AEL}}(Cn_{\mathbf{AEL}'}(\Gamma))$.

A possible semantics of $\mathbf{AEL}^{\mathbf{ZFC}}$ uses the following strategy: the **ZFC**-reliability strategy. Where $Dab(\Delta_1),\ Dab(\Delta_2),\ \ldots$ are the minimal Dab-consequences of Γ,

$$\Delta_i' = \begin{cases} \Delta_i & \text{if } \Delta_i \subset \Omega_{\mathbf{ZFC}} \\ \Delta_i - \Omega_{\mathbf{ZFC}} & \text{otherwise,} \end{cases}$$

Let $U^{\mathbf{ZFC}}(\Gamma) = \Delta_1' \cup \Delta_2' \cup \ldots$ The set $U^{\mathbf{ZFC}}(\Gamma)$ comprises the abnormalities that are **ZFC**-*unreliable* with respect to Γ. Where M is a **LLL**-model, $\mathrm{Ab}(M)$ is the set of abnormalities verified by M.

Definition 12 *An* **EL**-*model* M *of* Γ *is* **ZFC**-reliable *iff* $\mathrm{Ab}(M) \subseteq U^{\mathbf{ZFC}}(\Gamma)$.

Definition 13 $\Gamma \vDash_{\mathbf{AEL}^{\mathbf{ZFC}}} A$ *iff* A *is verified by all* **ZFC**-*reliable models of* Γ.

One can easily transform this semantics of $\mathbf{AEL}^{\mathbf{ZFC}}$ into a proof theory.

The definition of the enriched set theory $\mathbf{MRU}^+$ based on this non-standard adaptive logic is as expected.

Definition 14 $\mathbf{MRU}^+ = \langle \mathsf{EXT} + \mathsf{ABS} + \mathsf{FOUND} + \mathsf{FAF} + \mathsf{CHOICE}, \mathbf{AEL}^{\mathbf{ZFC}} \rangle$

If **ZFC** is consistent, no disjunction of **ZFC**-abnormalities is derivable. Hence, it is provable that $\mathbf{MRU}^+$ is an extension of **ZFC**, if **ZFC** is consistent.

Although the **ZFC**-reliability strategy is not the most elegant strategy, it effectively realizes what we want: it enables the definition of a theory $\mathbf{MRU}^+$ that is not only provably non-trivial but also contains all the theorems of **ZFC**, if **ZFC** is consistent. Apart from the **ZFC**-theorems, it also proves the existence of many more paradoxical and innocent sets like the universal set and the Russel-Curry-Quine-sets. Even for the paradoxical sets sensible theorems are derivable by means of the weaker paraconsistent logic and membership relation $\tilde{\in}$. Some readers might protest against the admittedly ad hoc character of the strategy. For those readers: consider that (a) there is a completely general prioritized variant (omitted here), which works fine but requires some technicalities plus extra axioms (to express the preference for **ZFC**-sets) and (b) in the definition of the **ZFC**-reliability strategy every **ZFC**-reference can be replaced by a reference to any other preferred set theory.

[9] Although **ZFC** is obviously not recursive, being a **ZFC**-abnormality is recursive, as this only depends on the set of axioms and not on its consequences.

6 Discussion of the presented theories

I have presented two different approaches to adaptive set theory as pragmatic foundations for mathematics. Both theories start from the axioms of naive set theory, i.e. comprehension and extensionality. Both are constructed from the basic assumption that the comprehension axiom still is the most natural and philosophically most attractive characterization of the intuitive notion of a set. In most informal mathematics or meta-logic one does not prove that the sets one constructs actually exist, one simply defines a set by writing down the property shared by all (and only) its members. Both set theories maximally validate the comprehension axioms and the classical consequences of these axioms, unless and until this leads to triviality. This makes them both provably non-trivial.

They are also discriminatory. Given that the theorems of an adaptive theory are those formulas that are verified by all selected ('all reliable models' in the case of the here used Reliability strategy) lower limit logic models of the axioms. Each particular lower limit logic model verifies some formulas and falsifies others (it is a regular logic with a regular notion of models). The non-theorems of an adaptive logic are all those formulas that are falsified by some selected model. So this set of non-theorems is definitely not too small to make a useful distinction between theorems and non-theorems. Moreover, this distinction is coherent as there is a formal logic that structures the set of theorems. Even if one does not trust adaptive logic for that purpose, one can argue that adaptive theories are also structured by the more regular lower limit logic of the adaptive logic. If one adds formulas to the axioms of an adaptive theory that express that the Reliable abnormalities of the adaptive theory are (classically speaking) false and one uses the lower limit logic as the underlying logic of a new theory based on the enriched axioms, one obtains exactly the same set of theorems as for the adaptive logic. So all the formulas that are not lower limit logic consequences of this enriched axiom set are also non-theorems of the adaptive theory, whence the adaptive theories are coherently discriminatory.

However, there are important differences between the approaches. Here I briefly compare the theories thematically.

Let us start with the lower limit logic and the meaning of the logical symbols. The lower limit logics of the theories are both extensions of classical logic. The **MCC** set theory has a lower limit logic that in itself does not add any useful consequence to the modalized version of the comprehension axioms. One needs conditional derivations for every single non-modalized consequence. Remark, however, that this does not mean that the theory would be as simplistic as the theory **CZFC**, introduced in Section 1. If **ZFC** would turn out to be inconsistent, many harmless instances of comprehension are still perfectly consistent. The adaptive theory will still validate those consistent instances.

By contrast, **MRU** has a quite rich unconditional basis. The logic **EL** is designed in such a way that this basis is as rich as possible. The comprehension axioms are completely classical apart from the right to left direction of the crucial equivalence symbol. This obviously means that also the properties that construct the sets are phrased in the classical part of the language. Even for the right to left direction of the equivalence symbol, we still have a nice paraconsistent alternative [10]. This can be considered as an advantage of **MRU** in

[10]Remember that $B, A \Rrightarrow B \vDash_{\mathbf{EL}} \sim\neg A$ and that $\sim$ is a strong paraconsistent negation, for

comparison with **MCC**.

MCC has the advantage that the meaning of the logical symbols is universal throughout the whole theory. The meaning of the non-classical equivalences in **MRU** is contextual: it depends on the consistency of the particular axioms in which the symbol occurs. If it is inconsistent the meaning and the behavior of the equivalence symbol are paraconsistent and if it does not lead to inconsistencies, the meaning and the behavior are classical. Some logicians might consider this a disadvantage.

Classical logicians that appreciate the adaptive approach presented here will definitely prefer the **MCC**-approach as this theory stays very close to classical logic. All symbols behave classically and the resulting theory is perfectly consistent, whereas **MRU** is inconsistent and uses non-classical symbols. Is it possible that convinced classical logicians (who reject all non-classical logics) appreciate non-monotonic theories? I strongly believe this is perfectly possible. One need not consider the underlying formal tool of the theory (in this case, an adaptive logic) as a full blown Logic[11]. Basically, I only provide a means that is able to (1) collect a set of formulas into an interesting foundational theory and (2) relate the formulas in this collection, by means of the dynamic proofs, to the usual possibly informal mathematical object proofs. In my opinion, there is nothing irrational about using a formal tool to collect formulas and at the same time hang on to the one true classical Logic as the ultimate standard of deduction.

I now come to a different topic in the comparison: the generalizability of the used techniques to other axioms and domains. The means used to develop the theory **MCC** are quite universally applicable. The methods applied do not rely on any particular property of the comprehension schema.

This obviously does not hold for **MRU**. Because the behavior of the alternative equivalence symbol that occurs in the abstraction schema is central for the theory, it seems unlikely that the ideas from this theory are applicable to completely different domains. However it is likely to be applicable to all similar domains. I focused my approach around **ZFC**, but similar techniques are likely to be successful for any usual set theory. Moreover, the methods are even likely to be applicable to non-set theoretic domains that are susceptible to paradoxes similar to the set theoretic paradoxes, e.g. truth theories, property theories and other domains where self-reference is allowed.

One aspect both theories have in common is the problem that paradoxical sets that are not inconsistent by themselves may infect otherwise perfectly innocent sets. This is the case for sets that lead to the Curry paradox and similar paradoxes. This phenomenon is strongly related to the way in which the Curry paradox (and the related paradoxes) leads to triviality. The Russell paradox (and similar paradoxes) leads to a plain inconsistency. Classical logic trivializes this inconsistency. This trivializing effect of classical logic can be blocked by choosing a paraconsistent logic and the inconsistency can easily be localized. In case of the Curry paradox, the arbitrary formula that can be derived is already in the axiom itself, which makes it a lot harder to localize and block it. However, both adaptive theories present a simple but effective means to localize them:

---

which the Double Negation rule and the De Morgan rules hold.

[11] I write Logic here with a capital to contrast it with the concept of logic I use: for me logic is nothing more than a formal tool that explicates reasoning and defines a consequence relation.

link them to the related Russell-like paradoxical sets. For the theory **MRU**, I did this by focusing on a common property Russell-like and Curry-like sets share: they typically are members of themselves or enable a loop of sets. In the case of the theory **MCC** sets are linked to their variants that are the result of substituting subfomulas by $\top$ or $\bot$. Comprehension axioms for the existence of Curry-like sets are thus formally linked to related axioms for the existence of Russell-like sets.

The advantages shared by both approaches are their flexibility, their foundational strength and their ability to get the most out of the comprehension axiom schema.

In this paper, I saw **CL** as the standard of deduction, **ZFC** as the standard practical set theory and the axioms of comprehension and extensionality as the ideal axioms of set theory. Of course one is able to develop other adaptive set theories that start from other standards and other ideals. This might result in better foundations. One could, for example prefer Quine's New Foundations as a standard practical set theory and devise an adaptive set theory that both enriches this where possible and provides a safety net for in case it would turn out be trivial. The resulting theory would probably be more elegant and philosophically more justifiable, but has the disadvantage that Quine's set theory is not as thoroughly studied as **ZFC**. A similar situation holds for the choice of an intuitionistic or a relevant logic as the standard of deduction: maybe this choice is philosophically less problematic, but the logics are also less studied as the underlying logic for set theory.

In conclusion, it was not my aim to develop one finished and optimal mathematical theory that can serve as a foundation of mathematics. There is many work to be done among which (i) carefully investigating the mathematical properties of adaptive theories and (ii) elaborating the metaphysical implications and justification of pragmatic foundations and adaptive theories. I only aimed to show in this paper that there are many interesting possibilities to realize a pragmatic foundation. The domain of adaptive mathematical theories is an almost unexplored research domain, in which many interesting subtle foundational theories may be found that are able to get around Gödel's incompleteness results.

References

[1] Diderik Batens. The need for adaptive logics in epistemology. In Dov Gabbay, S. Rahman, J. Symons, and J. P. Van Bendegem, editors, *Logic, Epistemology and the Unity of Science*, pages 459–485. Kluwer Academic Publishers, Dordrecht, 2004.

[2] Diderik Batens. A universal logic approach to adaptive logics. *Logica Universalis*, 1:221–242, 2007.

[3] Diderik Batens. *Adaptive Logics and Dynamic Proofs. A Study in the Dynamics of Reasoning.* 200x. Forthcoming.

[4] Diderik Batens and Kristof De Clercq. A rich paraconsistent extension of full positive logic. *Logique et Analyse*, 185–188:227–257, 2004. Appeared 2005.

[5] Diderik Batens, Kristof De Clercq, Peter Verdée, and Joke Meheus. Yes fellows, most human reasoning is complex. *Synthese*, 166:113–131, 2009.

[6] Ross T. Brady. The simple consistency of a set theory based on the logic csq. *Notre Dame Journal of Formal Logic*, 24:431–449, 1983.

[7] Ross T. Brady and Richard Routley. The non-triviality of extensional dialectical set theory. In Graham Priest, Richard Routley, and Jean Norman, editors, *Paraconsistent Logic: Essays on the Inconsistent*, pages 415–436. Philosophia Verlag, 1989.

[8] Haskell Curry. The inconsistency of certain formal logics. *Journal of Symbolic Logic*, 7:115–117, 1942.

[9] Paul C. Eklof. Whitehead's problem is undecidable. *The American Mathematical Monthly*, 83(10):775–788, 1976.

[10] Hartry Field. *Saving Truth from Paradox.* Oxford University Press, Oxford, 2008.

[11] Abraham Adolf Fraenkel, Yehoshua Bar-Hillel, and Azriel Lvy. *Foundations of set theory.* North Holland, Amsterdam, 1973.

[12] Kurt Gödel. Über formal unentscheidbare Sätze der *Principia Mathematica* und verwandter Systeme. I. *Monatshefte für Mathematik und Physik*, 38: 173–198, 1931.

[13] John L. Kelley. *General Topology.* Springer, 1975.

[14] Juliette Cara Kennedy and Mark van Atten. Gödel's modernism : On set-theoretic incompleteness. *Graduate Faculty Philosophy Journal*, 25(2): 289–349, 2004.

[15] Thierry Libert. *More Studies on the Axiom of Comprehension.* PhD thesis, University Libre de Bruxelles, Facult des Sciences, 2004.

[16] Uwe Petersen. A contraction free approach to logic based on inclusion and unrestricted abstraction. first draft.

[17] Graham Priest. *In Contradiction: A Study of the Transconsistent.* Martinus Nijhoff, The Hague, 1987.

[18] Willard V. Quine. New foundations for mathematical logic. *American Mathematical Monthly*, 44(2):70–80, 1937.

[19] Willard V. Quine. On the theory of types. *Journal of Symbolic Logic*, 3 (4):125–139, 1938.

[20] Greg Restall. A note on naïve set theory in *LP*. *Notre Dame Journal of Formal Logic*, 33(3):422–432, 1992.

[21] Greg Restall. How to be *Really* contraction free. *Studia Logica*, 52:381–391, 1993.

[22] Guido Vanackere. Preferences as inconsistency-resolvers: the inconsistency-adaptive logic **PRL**. *Logic and Logical Philosophy*, 8:47–63, 2000. Appeared 2002.

[23] Peter Verdée. Adaptive logics using the minimal abnormality strategy are Π^1_1-complex. *Synthese*, 167:93–104, 2009.

[24] Peter Verdée. Strong, universal and provably non-trivial set theory by means of adaptive logic. *Logic Journal of the IGPL*, 201x. URL `http://logica.ugent.be/centrum/preprints/adaptive_set_theory_foundation.pdf`. Submitted.

[25] Liza Verhoeven. All premisses are equal, but some are more equal than others. *Logique et Analyse*, 173–174–175:165–188, 2001. Appeared 2003.

[26] Liza Verhoeven. Proof theories for some prioritized consequence relations. *Logique et Analyse*, 183–184:325–344, 2003. appeared 2005.

[27] John von Neumann. An axiomatization of set theory. In *From Frege to Gödel: a source book in mathematical logic, 1879–1931*, pages 393–413. Harvard University Press, Cambridge, Mass., 1967.

[28] Zach Weber. Extensionality and Restriction in Naive Set Theory. *Studia Logica*, 94(1), 2010.

[29] Zach Weber. Transfinite Numbers in Paraconsistent Set Theory. *Review of Symbolic Logic*, 3(1):71–92, 2010.

[30] Alan Weir. Naïve set theory is innocent! *Mind*, 107(428):763–798, 1998.

[31] Richard White. The consistency of the axiom of comprehension in the infinite valued predicate logic of łukasiewicz. *Journal of Philosophical Logic*, 8(1):503–534, 1979.

[32] Ernst Zermelo. Untersuchungen ber die grundlagen der mengenlehre. I. *Mathematische Annalen*, 65(2), 1908.