ACTA UNIVERSITATIS CAROLINAE

PHILOSOPHICA ET HISTORICA 1/2015

MISCELLANEA LOGICA X

ACTA UNIVERSITATIS CAROLINAE
PHILOSOPHICA ET HISTORICA 1/2015

MISCELLANEA LOGICA X

CHARLES UNIVERSITY IN PRAGUE • KAROLINUM PRESS • 2016

Editor: Michal Peliš (Charles University in Prague)

http://www.karolinum.cz/journals/philosophica-et-historica

ISSN 0567-8293 (Print)
ISSN 2464-7055 (Online)

CONTENTS

2015 ACTA UNIVERSITATIS CAROLINAE PAG. 7–21
PHILOSOPHICA ET HISTORICA 1 / MISCELLANEA LOGICA X

MODEL THEORY AND FOUNDATIONS OF LOGIC

PAVEL ARAZIM*
Dept. of Logic, Faculty of Arts, Charles University in Prague
* E-mail: pavel.arazim@centrum.cz

ABSTRACT
Despite its popularity, model theory based on Tarski's insights is in need of deeper philosophical reflection. A wide range of stances towards it was proposed, some seeing it as project based on fundamental misconceptions, some asserting it reveals the very essence of logic. I would like to balance these extreme views. Of particular importance will be its connection to the problem of logical constants. Identifying logical constatnts enables us to identify logical forms of statements and thus brings us close to demarcating logic. We will see that solving this issue in ways suggested by model theory has its considerable costs, while the gains are rather modest.
Keywords: model-theory, completness, substitution, interpretation, representation

It would not be but a pure folly to doubt that the Tarskian semantics and model theory is a discipline of great importance which contributed significantly to the development of logic and mathematics. Thus it is only natural that it belongs among the most important parts of introductory lectures on formal logic and this should not change. So far, so good. Yet powerful as it is, it remains unclear how we should see it, how to interpret it philosophically. What is actually its place in logic? Is it the core of logic and other ways how to approach it, most prominently from the perspective of proof-theory, belong to logic only derivatevely? Or is the other way round? Or are they perhaps on a par, as far as their logicality is concerned?

This question would be, of course, uninteresting had there not been significantly different outcomes in logic and in philosophy of logic which depend on which answer to it we prefer. The topic of this article will be one of these answers, which depends exactly on seeing the model theory as essential for logic and which was taken by some significant figures of this discipline, including most prominently Tarski himself (in his particular case this was a later turn in his thought, but I will have more to say about this later). The idea is that in the model theory we can precisely specify what makes something a part of logic and thus we can delineate this discipline, ensuring that we will let neither too much, nor too little in. And although Tarskian models are essential to the semantics of classical first-order logic, these model-theoretic demarcations typically have it that logic is actually a much broader discipline. To be sure, taking model theory as the core of logic has got many further consequences, yet it is mainly on those regarding demarcation of the discipline I will focus on.

We will have to distinguish our main topic from a different, albeit related and actually more general one. The general topic which we will touch, as well, is the adequacy of model theory for the study of logic in general. The main point of reference in this regard is Etchemendy (1990). Now, if we reject the radical criticism towards the model theory presented in this Etchemendy's book, we will have settled at least that the model theory can be indeed of some use in logic. Such a situation will naturally call for a closer specification of what we can indeed use it for. And the bold thesis which Tarski presented later, is that in the model-theoretical framework we can say what the bounds of logic are, i.e. demarcate it as a specific discipline and thus, among perhaps other things, show something important about its relationship with mathematics.

We will see that there is both a debate about this kind of demarcations (which I will from now on call The Tarskian demarcation) in general, as well as an internal debate about how it should be exactly spelled out. I will present the gist of both these debates and take a stance towards them.

1. The Tarskian semantics

Let us now begin with recalling what the shape of Tarskian semantics is and see what its main virtues are. Locus classicus of his approach is the 1936 article Tarski (1936) in which he claims the inadequacy of merely proof-theoretic approach to logic. Among other things, Gödel's incompleteness theorems show that it is bound to undergenerate, i.e. to fail to display the relation of consequence in its completeness. Thus a quite different approach is required, one which pays more attention to the meanings of the expressions used in inferences. In other words, a (more)[1] semantic approach. And one which is concerned with what these expressions stand for.

I pressupose that the reader is familiar with standard semantics of classical first-order logic, that is the predicate logic with existential and general quantifier. This antecedent knowledge should serve as a common ground, even if I redescribe it in potentially controversial manners. First of all, Tarski wants to generalize a substitutional approach, which was heralded already by Bolzano. According to a perhaps somewhat anachronic interpretation of Bolzano – and according to Tarski as well – we have to identify a group of elements of our language as *logical constants*, that is as a specifically logical part of our language.[2] It is a separate issue, which elements these should exactly be and we will come to this peculiar problem later. But let us suppose, at least for the sake of this exposition, that it is the standard connectives (conditional, conjunction, disjunction and negation) and the two classical quantifiers.

Now we have to define the notion of a correct substitution, as it is essential for Bolzano's and for Tarski's approach. Here again, I rely on reader's knowledge of this notion. The idea is that in a given argument or an inference, we consider all its substitutional variants, where the variant is defined by mostly obvious restrictions, such as that only a predicate

[1] Although there were such tendencies among the logical positivists, we do not have to say that proof-theoretical approaches abstract from meanings completely. This is not so from the standpoint of inferentialism and also the later development of proof-theoretic semantics shows that the positivist views were somewhat hasty.

[2] It is anachronic because Bolzano spoke rather of extralinguistic *ideas*.

can be substituted for a predicate, an individual term for an individual term and so on. Not that there is no room for significant differences in opinions about what counts as a correct substitution, let us remind ourselves that for example Carnap in Carnap (1931) instisted that rules for substitution should be much more restrictive, such as to prevent the substitution of *prime number* for *emperor* in *Julius Ceasar was an emperor*. Obviously enough, such restriction would be difficult to formulate in any systematic way. But there are viable options of more fine-grained substitution rules than those for classical first-order logic which we deal with right now. But this is a seperate issue. Let us pressupose the notion of substitution operative in classical first-order logic.

The notion which we are after in logic is, of course, that of a valid argument. So, according to the substitutional account, an argument from a set of sentences to a sentece is valid if and only if all for all its legitimate substitutional variants it holds that either one of the premises is false or the conclusion is true. A legitimate substitutional variant of a given sentence – and then of a given a argument – is one, in which we substitute only for the expression which are not logical constants (so far, once again, we countenance the classical two quantifiers and the truth-functional connectives) and we substitute only according to the settled restrictions.

It is well known that this approach is problematic because it makes the inference dependent on the language we employ. When the language is not large enough, i.e. if it does not have a large enough vocabulary, it may well happen that we do not have enough substitutional variants of some plausibly invalid arguments to be able to declare them as actually invalid. Obvious examples can be found which are not much weakened even by the vagueness of the notion of intuitively valid/invalid arguments.

Tarski's approach seeks to circumvent this excessive dependence on language. John Etchemendy brings a highly controversial and, as we will see, not completely fair portrayal of the Tarskian endeavour. He describes Tarski's attempt at improving the substitutional approach as proceeding by binding the relation of logical entailment not just to a given actual language but to all possible languages. Even this is inaccurate and leads to Etchemendy's overall inaccurate interpretation, but it is useful heuristically. Thus Etchemendy takes Tarski as trying to step outside the actual language by considering its relation to the world. We do not consider as much substitutional variants of a given argument, but rather we imagine that the word–world relations might change. In our example, individual terms might refer to different objects than they do and predicates may refer to different relations, where relation is understood extensionally. This leads to the familiar notion of a model or structure, the reader's acquaintance with which I pressupose.

So much now for introducing the Tarskian semantics. I would like to note that there is an interesting discussion about whether this standard form it received is what Tarski actually intended. There are authors who think that he did not countenance a plurality of models as we do today, but rather just one univesal model. This would make his position generally much more vulnerable. But the most serious forms of criticism of Tarskian semantics are aimed at the nowadays usual form and thus we can afford to put this historical issue aside.

2. Pressupositions of Tarkian semantics and varieties of critique

Despite the fact that Tarski's analysis has become a part of standard logical curriculum, there has been serious criticism, which nevertheless always respected the achievements of the model theory. Given its proven usefulness, the issue is not whether or not we should somehow accept and use it but rather what to make of its use, how to interpret it. I will like to contribute to the view that the model-theory should be given maybe a little more modest interpretation than usual. But first, let us see the critique of it which is perhaps most direct and most prominent. The one which is due to John Etchemendy.

2.1 Etchemendy – exposition

Etchemendy claims that in logical semantics we have to choose between two basic alternatives, namely doing the interpretational or the representational semantics. Tarski is then supposed to be doing the first one. Let us explain these terms briefly.

In the interpretational semantics the models do not model the ways the world might be but only how the linguistic items might relate to it. Thus a singular term such as the president of the Czech Republic might be interpreted differently in different models, not because a different person can in fact have that political function, but merely because it can mean e.g. what the term *the highest mountain in the world* means in our actual language. The world is thus taken as it is. In the representational semantics, we model rather the ways the world itself might be, thus differing in the reference of the aforementioned individual term simply because different individuals (perhaps more than one at a time or none at all) may be the president. Now, it is obvious that both approaches are dependent on taking some vocabulary as logical, i.e. such that we do not consider either different interpretations of it or the ways the world might make it refer to something else.[3]

Now, according to Etchemendy, Tarski adheres to the interpretational semantics. This claim was much disputed by various authors, but Etchemendy claims that it is indeed the very core of Tarski's approach and his criticism thus cannot be seen as merely a historically interesting thesis about potential confusion of Tarski himself but has to be regarded as an attack on the whole tradition which it produced. Let us begin with marking Etchemendy's basic objections.

On the one hand, the interpretational semantics, even if it would give extesionally acceptable results, i.e. declare as valid all and nothing but the logically valid arguments as logically valid, it would, according to Etchemendy, succeed only by chance, as it reduces the logical validity of some sentences to the material validity of some other ones. Thus we claim that the argument

$$\frac{\text{John is a man and John is single}}{\text{John is single}}$$

[3] It may be good to note right now that it is not reasonable to see the sematics of classical logic as either represenational or interpretational but rather as both at the same time. But we will get to this later.

is logically valid because of the merely material validity (that is the truth of the conclusion or falsity of one of the premises) of all the instances of the following scheme

$$\frac{\phi \land \psi}{\psi}$$

or perhaps of

$$\frac{P(j) \land S(j)}{S(j)}$$

If we take just the $\land$ as logical, then this is plausibly materially valid for all the arguments of (one of the) indicated forms. But that gives us no assurance that the universal claim about all the arguments of this form is not true only accidentally. Etchemendy invites us to consider an argument of the form

$$\frac{\text{John was the president of the USA}}{\text{John was a man}}$$

Now if we allow just the individual term John to vary in the interpretation, then we will still get a an argument such that all its variants are materially valid. And this makes the argument logically valid, despite the fact that it obviously should not be declared as such. It has no materially invalid variants only by chance, so to say. Yet the Tarskian approach cannot distinguish it from the previous, presumably logically valid, one. Or so Etchemendy claims (it immediately comes to mind that the problem here lies in not letting the right expression vary in the interpretation but let us delay this a little bit yet).

If we, just for simplicity's sake, now shift the focus on logically true sentences instead of logically valid arguments, it is obvious that universally quantified sentence's being true does not guarantee the logical (or necessary, a priori or perhaps formal) truth of its instances, but only that they are simply true.

Etchemendy calls this alleged Tarski's step *Tarski's fallacy*. Tarski wanted to reduce the complicated and unclear notion of logical consequence, which was traditionally seen as involved in the difficult epistemological issues, by proposing a relatively clear-cut technical criterion. But his attempt is ultimately fallacious.

I have to say that I am a little bit puzzled by the most basic suppositions of Etchemendy's attack. It appears that the problem is supposed to be hidden in the fact the logical truth of one sentence, say A, is founded in the plain truth of all its variants and therefore in the plain truth of the general statements about the variants (i.e. *all the interpretational variant of A are true*), not latter's being logically true, as well. But I think that any special epistemological status of any claim can potentially be formulated by another sentence which itself is in fact just true. This explaining sentence has to be formulated in some metalanguage which is stronger with regard to the targeted epistemological features of claims. Or should we perhaps demand that the general sentence be true logically and not just simply true in the metalanguage? I am not sure whether such a demand makes sense, just as it hardly makes sense to speak of "trying to try" or "believing that I believe" (as is nicely shown in Brandom (1994)). Maybe the very question whether the metasentence claiming that the original sentence has got only true interpretational variants is true logically or just materially does not really make a good sense at all. And it was not an issue in the first place, as we were examining the logical status of just the original one.

Nevertheless, Etchemendy claims that the interpretational semantics gets the extension wrong, which I think would be, should it be indeed the case, a good reason for refusing it. But let us see some of the alleged instances of extensional inadequacy Etchemendy has in mind. Before that, it should be mentioned that it is in general far from clear what the talk of extension of the logical consequence relation being right or wrong is supposed to mean. I suspect it is simply too naive to suppose that our intuitions hide a totally clear-cut set of logically valid arguments and logically true sentences and that there are no border exemplars, which we ultimately have to simply choose whether to declare or not as logical.

Now for the examples of overgeneration of interpretational semantics. Let us say that there are at least two things (whatever that means) which can be denoted by the individual terms. This means that the following formula(and by applying this analysis, also the sentences which are its equivalents in natural languages)

$$\exists x \exists y((x \neq y))$$

will be declared as logically true. Now, such a sentence can apparently be true only by accident and therefore it makes no sense to declare it as a logical truth[4]. The way the Tarskian analysis escapes having to make this false decision is by varying the domain we quantify over, but that amounts, in Etchemendy's view, to changing the meaning of the existential quantifier, which compromises the original claim that it will be treated as a logical constant. Etchemendy brings more examples of overgeneration, but they are mostly of the same spirit, so we can confine our attention to this simple one.

Etchemendy thus has to interpret Tarski as saying that there must be just one universe of discourse, it must be somehow given what there is in the most general sense. Even if there might be some indications that Tarski might have wanted to head in this direction, it is hardly understandable how such a position can be supposed to be held by the proponents of the Tarskian semantics. Indeed, if his attack is not supposed not to be directed merely at Tarski and thus not to be of mainly historical interest (and I have already noted that there are debates regarding what Tarski originally had in mind), it is hard to see who it is supposed to be aimed at. Nobody is against using different models and it does not make much sense to see them as submodels of one great supermodel (which might perhaps lead to the charge of changing the meaning of the existential quantifier, as Etchemendy formulates it). When adopting a model, there is in principle no claim that the members of the domain have to be existent in the sense of being real. We are free to choose a domain containing pegases and treat it as an intended model of our theory of imaginery creatures. And by doing so, we do not anyhow claim their existence, just suppose it, thus modelling some contexts of argumentation.

4 It should rememebered, though, that some authors would not oppose such a verdict. Frege though that logic guarantees existence of objects, such as natural numbers.

A problem arises though, namely how many models are we supposed to be using. The problem of possible overgeneration of logical consequences is that we might not have enough models. And it depends on the set theory we use as a background which models there are. This seems to be first of all a problem of indeterminacy, because we cannot say which class of models is somehow the right one. There are more ways to react to this. One of them is, I believe, to regard this as reflection of some genuine vagueness of the notions which are being formalized and thus rendered more precise. This apologetic stance might not appeal to every one, but it is at least not obviously wrong and further discussion is needed here. I find Etchemendy's claim that the axiom of infinity is ad hoc as worthy of attention, though. Be it as it may, it certainly helps to get more plausible verdicts about the logical entailment relation and about the logical truth of statements (thus no statement of the form *there are at most n objects* for a finite *n* will be declared as logically true[5]). So the most counterexmples Etchemendy considers do not arise. Or better, they do not arise in the case of classical first-order logic, which we were considering so far. This is, to remind us, not just because of the happy choice of the underlying set theory, but also because of the particular choice of the logical vocabulary.

Etchemendy follows Tarski in believing that in case of first order logic we have the problem of undergeneration. The invalidity of the ω-rule is claimed to be something in need of a remedy both by Tarski in 1936 and Etchemendy. And thanks to Gödel's incompleteness theorems there is no way we can hope to solve the problem at the level of the classical logic. Yet here we can dispute whether it is a genuine example of undergeneration. Tarski and Etchemendy thus in fact seem to favor a logic which will validate the logicist thesis that arithmetic is a part of logic, as it was presented in Frege (1884). But as interesting as this project was, it is not clear that its fulfillment is so desirable. It is still possible to stick with the more traditional Kantian view that mathematics and logic are indeed separate disciplines. The omega-rule is thus an argument whose validity is not purely logical, but involves our mathematical faculties as well. We have to use the pure intuition to *see* its validity. Depending on one's philosophical background, one can see the (in-)validity of the ω-rule either as an asset or as a problem for a given logic.

Be it as it may with the problem of undergeneration, the overgeneration appears to be much more of a threat. But only so if we interpret the model-theoretic semantic in the interpretational manner suggested by Etchemendy. And that interpretation is rather a straw man for him to attack, as I will try to show.

3. Representational semantics

Another possibility to interpret the Tarskian semantic is, according to Etchemendy, to interpret it representationally. We have already seen a sketch of what that would involve. A given argument is declared valid in case it remains valid under any changes in the world. Described in this way, as it is described by Etchemendy in his 1990 book, it means

5 To make sense of Etchemendy's claim, we cannot say that the axiom of infinity prevents such statements from being logically true by enabling infinite models, as they would not be true even if we had just finite models, though of unbounded finite cardinality. Etchemendy pressuposes the whole time that these finite models must be taken from one big universal model – the world – which then has to be infinite.

merging logic with some sort of general metaphysics. Clearly this does not look like a promising explanatory strategy, as it would involve claryfying the logical notions by means of perhaps even more obscure methaphysical ones. As difficult as it might be to decide about the logical validity of arguments, should it be determined by these criteria, we can see that this would most likely mean running the logical validity together with the analytical one. The inference from

John is a bachelor

to

John is an unmarried man

obviously has to remain valid, no matter how the world changes, as far as the changes do not involve our language (whatever that means). But I think that since at least the appearence of Quine's *Two dogmas of empiricism* we should beware of such a construct. Does Etchemendy's distinction between the interpretational and representational semantic make really sense? And can we accept taking all the analytical entailments as logical ones? I suspect that it was one of the tasks of logic to distinguish precisely between analytically and purely logically valid entailments.

Yet Etchemendy eventually refrains then from this concept of representational semantics and uses the term differently in Etchemendy (2008), partly perhaps as a reaction to criticism, which was issued by Gila Sher in her article Sher (1996). There she accuses Etchemendy of presenting us with a false dilemma, having to choose between the two basic kinds of semantics. Indeed, I think most people acquainted with Tarskian semantics will say, when forced to decide whether it is interpretational or representational, that it is somewhere in between[6]. Indeed, sentence such as

Every bachelor is unmarried

is not declared as a logical truth, perhaps mostly because of the fact that the actual language could have worked differently, many other senteces are not declared logical truths rather because of the way the world could have been, but we cannot in general allot the responsibility just to the language or to the world.

It should be noted that Etchemendy refuses the Quinean attack at the synthetic/analytic distinction, claiming that the attack is based on too narrow a conception of logic. And here we come to the meaning Etchemendy later gives to the represantional semantics[7]. Under this new description he actually endorses it. Logicians, according to this view, always study the inferential properties only of certain expressions, for example the classical connectives and the two classical quantifiers and consider the situations when

6 As should be clear already, we are putting aside the disputes about Tarski's opionons in the 30's and talking about the model theory in its modern shape. Saying that it is somewhere in between the two approaches is a somewhat simplifying expression of what is better expressed in MacFarlane (2000), namely that various models model different contexts. That is, not necessarily interpretations or states of affairs.

7 Though it is perhaps a little dubious why he calls it so. His exposition can be found in Etchemendy (2008).

the members of the other parts of the vocabulary, such as *bachelor* or *unmarried man* change their denotation. And herewith we come to the problem of logical constants.

4. Logical constants

Logical constants can be characterized as the elements of language which determine the logical properties of sentences, the only ones, which, as Quine puts it in Quine (1986), occur in logical truths or logical entailment relations essentially. The problem is that it is not clear which elements of the language should be counted as logical constants. Tarski himself expresses in the 1936 article the opinion that the division between logical and non-logical constants cannot be completely arbitrary, but it might be impossible to demarcate the logical constants quite principally, as well.

Etchemendy thinks that the choice of logical constants is indeed arbitrary, because every element of language has some logical properties and it is only up to us, which collection of linguistic items we want to study from the logical point of view. He calls the problem of finding the right set of logical constants a red herring. Every element of language has got some specific logical properties and it is up to us whether we find it useful to study them. It is thus very well possible to study for example the logic of "color-words", which typically involves inferences such as

$$\frac{\text{This apple is red all over its surface}}{\text{This apple is not green all over its surface}}$$

Or inferences such as

$$\frac{\text{This apple is red}}{\text{This apple is coloured}}$$

The traditional logical constants were historically given special treatment only because their logical properties are particularly important or particularly amenable for logical analysis. Now, this approach is of course a possible one, but it obviously makes the very notion of logic very vague. Or rather very broad. Logic is thus transferred into a general study of inference. It is thus important that Etchemendy does not regard inferences such as

$$\frac{\text{Socrates is a man}}{\text{Socrates is mortal}}$$

as an enthymeme. This brings him close to positions of Robert Brandom. But what drives him far away from Brandom's position is that he does not endorse logical expressivism, which is a corollary of the fact that he does not think that logic has got a specific vocabulary. He probably also does not agree with Brandom's identification of meaning of an expression with its inferential properties. Meaning can hardly be, according to Etchemendy's picture, constituted by a position a given sentence – and derivatively also its constituents – have in the overall inferential web. Without logical expressivism it is mysterious how inferentialism could work. Thus his view of meaning seems to be irreducibly representationalist. And such a view has faced many problems in the recent decades of philosophy of meaning.

Keeping that aside for now, we should say that there are ways how to characterize logical constants in the Tarskian framework, ways which are vulnerable to criticism but are not completely arbitrary, thus at least partly fulfilling Tarski's original desideratum. The core of the proposal comes from Tarski himself, from his lecture, given well after his original articles about semantics, namely Tarski (1986). In it he generalizes the Klein's Erlangen programme of demarcating notions of various geometries.

5. Invariance criterion

The key notion in this Tarskian enterprise of demarcating logical constants is that of invariance. For example, the notions of Euclidian geometry, such as being an isosceles, are invariant under permutations of the universe of points which preserve similarity. A permutation of the points naturally induces a permutation of the sets of points, of sets of sets of points and so forth (the permutations on higher levels). Thus a similarity-permutation is one which maps a given triangle onto another triangle, which can be proportionally smaller but remains an isosceles if and only if the original one was such and so forth.[8]

Now the first attempt to define logical notion is to say that they are the ones which are invariant under all the permutations. The argument starts off with the premise that logical notions should be the most general ones. Now, when we relax our demands on the class of permutations, under which the notions of a given discipline are supposed to be invariant, we get an increasingly general discipline. Logic therefore goes as far as is possible in this setting. Here it actually seems that Tarski is speaking about a single universe (the "world"), which might give support to the earlier mentioned Etchemendy's interpretation of his endeavour. Yet this approach needs to be ammended, as it would allow for example the quantifier $\exists_{\forall}$, which would behave as an existential quantifier in case there are some cats and as the universal one otherwise. It would be thus indeed invariant under all permutations, but it indeed feels strange to accept it as a logical notion.

This problem was nevertheless fixed, as later authors, such as Sher, began to consider not just all the permutations of a given domain, but rather bijections between various domains. Sher calls this typically an isomorphism, but she does not mean that it respects the interpretation of non-logical symbols in a given model, that is preserves the properties of the members of the universe, rather it just preserves the properties of higher level-objects (sets of objects, sets of sets and so on), i.e. those which are induced from the original domain of the given structure. Let us see which notions thus get counted as logical.

To begin with, no individual constant passes the test. If we have e.g. the constant 0 in the language of arithmetics, it can be mapped e.g. to 1 even in the same structure of natural numbers, that is in the standard model of Peano arithemetic. Of the first-level

8 A more systematic and less hasty exposition can be found in Tarski (1986).

predicates (or sets) we get counted the universal relation and the empty relation, from the first-level binary relations identity, the non-identity (that is the complement relation to that of identity) and so forth.

When we start talking about the quantifiers, understood as second-order predicates (the predicates of predicates, or sets of sets), the list gets significantly extended. First of all the two classical quantifiers, that is the existential and the universal one clearly pass the test. A non-empty set get clearly mapped on a non-empty set by any bijection between two structures, as well as an empty one. The same consideration holds for the universal set. But we can go futher and consider any quantifiers regarding the cardinality. Thus any quantifier demanding that a set of objects satysfying a given formula has got a certain cardinality is declared as logical. Just for illustration consider

$$\aleph_1 x \phi(x)$$

But we can generalize even more. The traditional quantifiers are, from our point of view, sencond-level unary predicates. But we can consider also second-order predicates of higher arity, for example the relation *most*(thus being able to formalize such propositions as *Most A's are B's.*). And we can also consider unary second-order predicates, which are applied to first-order relations of higher arity than one, say the binary ones. Or we can have hybrid relations, which are applied for example to an individual and a predicate, such as the relation of membership, understood not as a relation between elements of the universe but between the elements of the universe and sets of elements thereof.

Gila Sher in her Sher (1991) presents the results of this approach in a very comprehensive manner. The book is thus reccomendable for those who want to get a more exact idea of the results of this demarcation. Yet we have now seen what might at least give the basic flavor of what we get. Now, can we be happy with such a result? When it comes to the problem of extensional adequacy, it is clear that overgeneration is much more of a danger than undergeneration in this case. It can even be shown that any structure can be characterized by the means of the bijection-invariant operators, among other the standard model of Peano arithmetic, see Bonnay (2008). What are we to make of this? There are authors who see this as a mark of adequacy of this demarcation, such as Sher, and also ones who see it also a clear mark of problem, such as Dennis Bonnay.

5.1 Virtues of the demarcation

This approach gives a demarcation which is very precise and systematic. From a certain point of view, given by Sher, the classical logic, confined to its two quantifiers, appears to contain a relatively arbitrarily small fragment of what the whole system, which she calls *universal logic* has to offer. For example the cardinality-quantifiers are very similar to the two classical ones from the set-theoretic point of view. The system she countenances is actually still a first-order logic, as it does not contain the second-order quantifiers, it might be called the generalized first-order logic[9]. Or rather generalized

[9] Of course, this demarcation based on invariance can be exted to higher-orders and Tarski originally does exactly this. Yet Sher shows that it is actually quite enough to consider just the first-order generalized quantifiers.

first-order logics, as we may choose to work in smaller systems, such as the classical logic, or the classical logic enhanced just by the quantifier "there are infinitely many" or "there are uncountably many" etc. As we speak in the case of merely propositional logic of the completness of the connectives, i.e. that e.g. the negation together with disjuction are capable of expressing all the boolean functions, so we can speak in a slightly figurative manner also of the completness of the first-order logic. The universal logic which we just sketched can be thus seen as complete with respect to what can be expressed by the means of first-order quantifier-operations defined over structures.

Despite the mentioned similarity between the two cases, the contrast is great, as well. As we content ourselves with just two (or, of course, one) connectives in the case of completness of the propositional logic, in the case of first-order logic we need a host of quantifers which is very difficult to oversee (at least as difficult as to oversee the set theory). Sher claims that although this approach blows logic up to an unprecedented degree and makes it thus immensly complex, it compensates for this fact by being principled, i.e. by being based on a single and clear principle. Informally said, logic is a discipline which abstracts from the identity of objects, all objects are equal for it. This might remind us of Kant's conception of logic. Kant claimed that logic abstracts from the relationship of cognition to its object. This probably cannot be said of the universal logic Sher proposes. This logic treats of a relationship of cognition to objects, though in a very general way, surely not of relationship to any concrete objects.

But even this might be slightly doubtful. Of course, we have to accept the specific understanding of *object*, i.e. the member of domain of some model-theoretic structure and not, for example, a set theoretical construction on these objects. An existential quantifier or any of the generalized ones can be seen as operation on the structures and as such perhaps also as an abstract object. This is not a refutation but it shows that this approach, not much surprisingly, pressuposes that the notion of an object is already settled. It is up to the reader to decide whether logic can be build upon such a presuposition.

Anyway, Sher praises general logic for displaying the form of our reasoning. Logic becomes the discipline of the formal. Certainly we can choose to understand formality as what is captured by this generalized logic. I suspect that any non-mathematical or intuitive notion undergoes some changes when it gets treated mathematically, at least in the sense of being made more precise and thus bereft of its vagueness which might have contributed to its importance and vivacity.

But there is one larger problem. Or perhaps two related ones. The first one might be the concern with possible overgeneration. Again, such an issue typically cannot be decided definitely, as it is not clear with respect to what the given logic is supposed to over- or undergenerate (some sort of "right" relation of logical consequence). Yet in this case we see that a lot of set theory has creeped in. Indeed, the situation recalls the second-order logic and Quine's dictum that it is a *set theory in sheep's clothing*. Perhaps logic should not be able to speak of such things as various infinite cardinalities. For one thing, it might jeopadize a status which is often attributed to it, namely being topic-neutral. I believe that the set theory is a topic and a large one! Logic is thus to contain vocabulary, which is relevant only to one specific discipline, which also seems hardly acceptable. Yet of course, it will depend on broader philosophical stances towards logic, whether one sees this as problematic. It is possible to renounce the topic-neutrality as a desideratum of logic.

The more acute, though related worry is that should this all be logic, then it would somehow lack real foundation. After all, the set theory does not seem to be a safe foundation for lots of reasons. In a way, we do not really understand what the quantifier $\aleph_1$ means, since we do not know whether the continuum hypothesis is true. Actually, since it was shown to be independent of the axioms of ZF in (classical) first order logic, it is even quite reasonable to say that we cannot declare the continuum hypothesis neither as true nor as false. By this I do not mean only that we cannot know whether it is true or false but rather that it itself is neither true nor false. In order to say either, we would have to know what the *real* model of set theory is, which I believe is not the case. We actually have no way of veryfying that there is even any model at all.

Again, similarly to the case of the second-order logic, we can even formulate in the purely logical language of Sher's universal logic a sentence which is equivalent to the continuum hypothesis. Logic thus has to declare CH either as true or false, which is very hard to swallow. Sher tries to defend her system in a similar way in which Shapiro tries to defend the second-order logic, charging its opponents of "foundationalism" in Shapiro (1991). Sher claims in her article Sher (1999) that when we try to explicate logic, it is bound to lose its character of foundation of all knowledge, it has to be made partly dependent of something, which helps to explain it. This much, I believe, is true. Yet of course the question remains how complicated can the tool, e.g. the set theory, which we use to explicate logic be. In general, it is up to us and our preferences, though founding logic on something as complicated as the set-theory seems to be too much. We might be prepared to revise some of our intuitions about logic as a foundation of cognition, but this amounts rather to changing the subject that offering a novel account of logic.

Furthermore, this criterion does not rule out some very dubious quantifiers, because it pays attention, so to speak, only to the quantifier's good behaviour on structures of every cardinality separately. We can thus think of a quantifier, which behaves as an existential one on finite models and as a universal one on the models of infinite cardinality. Furthermore. We can think of quantifiers which are extensionally equivalent to, say, the existential quantifier, but have obviously a different meaning. For example a quantifier, which is – taken as a second-order predicate – true of a set under the conditions that it is non-empty and water is H_2O. Nice overview of these examples of overgeneration can be found in MacFarlane (2009).

Indeed, this criterion is, as John Macfarlane calls it in MacFarlane (2000), actually not semantic, but a presemantic one, as it does not deal with the relationship between the extension of logic operators and linguistic items, by which they are supposed to be denoted. Gila Sher asserts that logic is indeed dealing only with the extensions of our linguistic expressions. Yet it is difficult to see why such a restriction should be reasonable. Indeed, it almost appears as an inversion of the Kantian restriction that logic should not speak about the relationship between cognition and its objects.

6. Final assessment

Does this all mean that Tarski's approach to logic is incorrect or flawed? We have to be prepared to accept that every attempt to demarcate logic is bound to be only partially successful, since the guiding intuitions are too vague and may always produce new objections against individual proposals. Yet we have seen that there are many objections this specific approach has to face. Not that there are no possibilities to defend it or to amend it. It worth mentioning that Dennis Bonnay is, among others, trying to generalize the notion of invariance and speaks of invariance not just under bijections or isomorphism but also under partial isomorphism and so forth. He thus shows that it is possible to reduce a lot of the problematic interconnections of logics based on invariance criteria with what should be rather extra-logical affairs and especially the dependency on problematic set-theoretical assumptions such as the continuum hypothesis. Even the classical first-order logic can be characterized by means of invariance criteria, namely notions invariant under monadic surjective functions Bonnay (2008). These are for themselves very interesting results which give us the new possibilities of understanding the various logical systems and understand what the difference between the classical logic and its Tarskian amplifications – including the second-order logic – ammounts to.

As I mentioned in the beginning, Tarskian semantics surely is a powerful and handy tool for studying various logical systems. Yet it seems hardly acceptable to see it as revealing the essence of logic (perhaps nothing can achieve such a goal). Yet in the case of the first-order logic we have the happy circumstance that it is complete. Or perhaps we should use a different term, such as axiomatisable. As Etchemendy points out in Etchemendy (1990), the term completeness suggest that the model-theory is something more basic and secure, something the axiomatisation is to be tested against. And I see his suggestion to look at things from the opposite perspective as a very healthy one (this idea is developed in Peregrin (2006) and Peregrin (2014)). This means regarding the axiomatisation as something, which can be seen as being certainly in the realm of logic, at least in the sense that the axioms and inferential rules are of themselves plausibly logically valid. Then the model theoretical system of classical first-order logic gets its foundation by the completeness theorem. Yet this cannot be of itself a sufficient argument in favor of some kind of exclusiveness the classical logic. First of all, the notion of plausibility of the deductive system is problematic, invoking the traditional notion of an axiomatic system as a system of self-evident truths, which is hardly tenable, given the many alternative logics. The second problem is that in the Tarskian semantic we can formulate systems which are stronger than the classical logic and still axiomatisable, such as the system of classical logic plus the *there are uncountably many* quantifier. And here the problems with dependency on the set-theory and epistemological ill-foundedness reemerge.

References

Bonnay, D. (2008). Logicality and invariance. *Bulletin of Symbolic Logic* 14, 29–68.

Brandom, R. (1994). *Making it explicit*. Cambridge: Harvard University Press.

Carnap, R. (1931). Überwindung der metaphysik durch logische analyse der sprache. *Erkenntnis*, Volume 2, Issue 1, pp. 219–241.

Etchemendy, J. (1990). *The Concept of Logical Consequence*. Cambridge, MA: Harvard University Press.

Etchemendy, J. (2008). *Reflections on Consequence*. Douglas Patterson (ed.), New York: Oxford University Press, pp. 263–299.

Frege, G. (1884). *Die Grundlagen der Arithmetik*. Jena, hrsg. von Christian Thiel.

MacFarlane, J. (2000). *What does it mean to say that logic is formal?* Dissertation, University of Pittsburgh.
MacFarlane, J. (2009). *Logical constants*. Available at http://plato.stanford.edu/entries/logical-constants/.
Peregrin, J. (2006). Consequence & inference. Kolman, V. (ed.): *Miscellanea Logica: Truth and Proof*, FF UK, Prague, 1–18.
Peregrin, J. (2014). *Inferentialism: Why Rules Matter.* Palgrave Macmillan.
Quine, W. v. O. (1986). *Philosophy of Logic*. Cambridge: Harvard University Press.
Shapiro, S. (1991). *Foundations without Foundationalism.* Oxford: Oxford University Press.
Sher, G. (1991). *The Bounds of Logic: a Generalized Viewpoint*. Cambridge, MA: MIT Press.
Sher, G. (1996). Did Tarski commit "Tarski's fallacy"? *Journal of Symbolic Logic* 61 (2), 653–686.
Sher, G. (1999). Is logic a theory of the obvious? *European Review of Philosophy* 4.
Tarski, A. (1936). O pojeciu wynikania logicznego. *Przeglad Filozoficzny* 39, 58–68.
Tarski, A. (1986). What are logical notions? *History and philosophy of logic*, 100, 143–154.

2015 ACTA UNIVERSITATIS CAROLINAE PAG. 23–37
PHILOSOPHICA ET HISTORICA 1 / MISCELLANEA LOGICA X

NORMAL DEFAULT RULES AS EPISTEMIC ACTIONS

MICHAL DANČÁK* and MICHAL PELIŠ
Dept. of Logic, Faculty of Arts, Charles University in Prague
* E-mail: mdancak@seznam.cz

ABSTRACT
The goal of this paper is to present a prospective way to 'translate' normal default rules into the framework of action models logic. At the beginning we introduce default logic and normal default logic with their main properties and, separately, action models logic. Then a 'translation' of normal default rules in a slightly modified action models logic is presented.
Keywords: action models, default logic, dynamic epistemic logic

1. Reasoning with default rules

Using the word *reasoning* we mostly mean 'private' act of a subject.[1] Common reasoning that we do all the time is often based on two kinds of information. The first one, *hard information*, is information that we are obliged to accept; we are sure of it, it is our knowledge and trusted data. The other one, *soft information*, is something that 'typically happens', it is very likely that things go that way. Reasoning based exclusively on hard information would be ideally deductive. However, it is not possible in common reasoning. There are many typical situations that are produced by our experience. Some of them are in contradiction, some of them are incomplete. Nonetheless, we have to do conclusions even if there is a lack of hard information. Such conclusions can be out of the scope of classical (deductive) consequence relations.

Let us imagine that we know that Anne is a student of a faculty of arts. A typical student from a faculty of arts does not like mathematics and we could conclude by default that Anne does not like mathematics. The knowledge that Anne attends a faculty of arts together with our prejudice that students from this faculty do not like math can form a (default) 'rule':

> If *AnneStudentOfArtFaculty*, then ¬*AnneLikeMath* under condition that the conclusion ¬*AnneLikeMath* is not in a conflict with our current knowledge.

Later we obtain an information that Anne studies logic. Well, if she studies logic, she might like mathematics. Similarly with this information, we can formulate a 'rule':

[1] Let us call a 'reasoning subject' an *agent*. The word 'private' will be discussed later on.

If *AnneStudentOfLogic*, then *AnneLikeMath* under condition that the conclusion *AnneLikeMath* is not in a conflict with our current knowledge.

More formally, if we put hard information together, we obtain a set (of facts)

$$\Gamma = \{\text{AnneStudentOfArtFaculty}, \text{AnneStudentOfLogic}\}$$

and our reasoning about math's popularity of Anne can follow rules like these:[2]

$$\frac{\text{AnneStudentOfArtFaculty} \ : \ \neg\text{AnneLikeMath}}{\neg\text{AnneLikeMath}}$$

$$\frac{\text{AnneStudentOfLogic} \ : \ \text{AnneLikeMath}}{\text{AnneLikeMath}}$$

Note that once we deduce that Anne does not like mathematics we may not use the second rule because its presupposition is in conflict with our current knowledge, i.e., with the conclusion that Anne does not like math.[3]

In real life, one might introduce a preference relation between these two rules. For example, there is a study program of Logic at the Faculty of Arts in Prague and every student of logic is also a student of this faculty. Therefore the fact that Anne studies logic is more informative here than the fact that she studies at the faculty. One might thus prefer to use the second rule since its assumption is 'more informative'.

The idea of common reasoning formalization is mostly studied in non-monotonic logic – logical systems where monotony can fail. A conclusion of a set of premises needs not be a conclusion anymore if we extend the set of premises.[4] In this paper we will use a formal system called *default logic* with operational semantics; cf. [1], especially. Our aim is to introduce the idea of 'translating' default rules in action model logic. For these purposes we will introduce only the plain version of default logic. We will have a set of premises (hard information) together with a set of default rules that form a default theory.

1.1 Operational semantics

A (general) *default theory* T is a couple (Γ, D) where the set of formulas Γ represents 'hard information', which is accepted as true, and D is a finite set of *default rules* (or *defaults*, for short). A default (rule) can be understood as an accepted way to extend our 'hard information'. It enables us to do conclusions that extend possibilities of classical consequence relation.

A (general) default rule is of the form

$$\frac{\varphi \ : \ \psi_1, \psi_2, \dots, \psi_n}{\chi}$$

2 The rule is in the form of a fraction; hard information is written before the colon together with what is presupposed (behind the colon), a consequent is under the line.

3 If we conclude that Anne does not like mathematics, we cannot consistently assume that she does like mathematics (and vice versa).

4 See, e.g., [1], [2], [3], and [5].

where $\varphi, \psi_1, \ldots, \psi_n$, and χ are formulas of a background logic.[5] The meaning of a general default can be:

> If (a *prerequisite*) φ is 'known' to be true and (*justifications*) $\psi_1, \ldots, \psi_n$ can be consistently presupposed, then (a *consequent*) χ is derivable.[6]

There is an idea presented in [1] on how to work with defaults algorithmically. Let us imagine we form sequences of defaults from the set D without multiple occurrence: Π_1, Π_2, Π_3, etc. The ordering of defaults in $\Pi_j = \langle d_{j_1}, \ldots, d_{j_k} \rangle$, where $\{d_{j_1}, \ldots, d_{j_k}\} \subseteq D$, is an order of their possible applicability. Before we introduce the term *applicability* of a default rule to a deductively closed set of formulas, we define two auxiliary sets, $\mathsf{In}\Pi_j$ and $\mathsf{Out}\Pi_j$, for each sequence Π_j. Both sets must be understood as arising step by step, i.e., default by default, according to an order in Π_j. The definition is by recursion, where Π_j^m denotes the initial segment of $\Pi_j = \langle d_{j_1}, \ldots, d_{j_k} \rangle$ of length m, where $m \leq k$:

$$\begin{array}{lcl}
\mathsf{In}\Pi_j^0 & = & \mathsf{Cn}(\Gamma) \\
\mathsf{In}\Pi_j^{m+1} & = & \mathsf{Cn}\left(\mathsf{In}\Pi_j^m \cup \left\{\chi \mid \frac{\varphi : \psi_1, \psi_2, \ldots, \psi_n}{\chi} = d_{j_{m+1}}\right\}\right) \\
\mathsf{In}\Pi_j & = & \mathsf{In}\Pi_j^k \\
\mathsf{Out}\Pi_j^0 & = & \varnothing \\
\mathsf{Out}\Pi_j^{m+1} & = & \mathsf{Out}\Pi_j^m \cup \left\{\neg\psi_1, \ldots, \neg\psi_n \mid \frac{\varphi : \psi_1, \psi_2, \ldots, \psi_n}{\chi} = d_{j_{m+1}}\right\} \\
\mathsf{Out}\Pi_j & = & \mathsf{Out}\Pi_j^k
\end{array}$$

Step 0 does not apply any default rule, but it prepares all what is (logically) obtainable from the 'hard information' Γ. In step 1, we take the first default in a sequence Π_j and test its applicability (with respect to $\mathsf{In}\Pi_j^0$):[7] φ is included in (the so far obtained) $\mathsf{In}\Pi_j^0$ and no $\psi_1, \ldots, \psi_n$ is in contradiction with (the so far obtained) $\mathsf{In}\Pi_j^0$, i.e., $\neg\psi_l \notin \mathsf{In}\Pi_j^0$, for each $l \in \{1, \ldots, n\}$. If the default is applicable, sets $\mathsf{In}\Pi_j^1$ and $\mathsf{Out}\Pi_j^1$ are created

$$\begin{array}{lcl}
\mathsf{In}\Pi_j^1 & = & \mathsf{Cn}\left(\mathsf{In}\Pi_j^0 \cup \left\{\chi \mid \frac{\varphi : \psi_1, \psi_2, \ldots, \psi_n}{\chi} = d_{j_1}\right\}\right) \\
\mathsf{Out}\Pi_j^1 & = & \left\{\neg\psi_1, \ldots, \neg\psi_n \mid \frac{\varphi : \psi_1, \psi_2, \ldots, \psi_n}{\chi} = d_{j_1}\right\}
\end{array}$$

and we can continue with step 2, and so on.

Definition 1. *A default* $\frac{\varphi : \psi_1, \psi_2, \ldots, \psi_n}{\chi}$ *is applicable to a deductively closed set* Δ *iff* $\varphi \in \Delta$ *and* $\neg\psi_1 \notin \Delta, \ldots, \neg\psi_n \notin \Delta$.

Our introductory example can be formalized as a default theory where $\Gamma = \{\varphi, \psi\}$ and $D = \left\{\frac{\varphi : \neg\chi}{\neg\chi}, \frac{\psi : \chi}{\chi}\right\}$. We can form five sequences of defaults from D:

---

5 In this paper, we use classical propositional language (modalities will be added later on). Consequence relation (resp. operation Cn) is based on classical propositional logic. A set of formulas Δ is deductively closed iff $\Delta = \mathsf{Cn}\Delta$.

6 There are always at least one justification and consequent. It is possible to have no prerequisite. In that case we interpret the empty 'prerequisite place' as *true*, i.e., tautology.

7 For simplicity, let us assume that $d_{j_1} = \frac{\varphi : \psi_1, \psi_2, \ldots, \psi_n}{\chi}$.

$$\Pi_1 = \langle\ \rangle \qquad \Pi_2 = \left\langle \frac{\varphi : \neg\chi}{\neg\chi} \right\rangle \qquad \Pi_3 = \left\langle \frac{\psi : \chi}{\chi} \right\rangle$$

$$\Pi_4 = \left\langle \frac{\varphi : \neg\chi}{\neg\chi}, \frac{\psi : \chi}{\chi} \right\rangle \quad \Pi_5 = \left\langle \frac{\psi : \chi}{\chi}, \frac{\varphi : \neg\chi}{\neg\chi} \right\rangle$$

Let us look at Π_5, for example. The first default is applicable to $\mathsf{Cn}\Gamma$ since $\psi \in \mathsf{Cn}\Gamma$ and $\neg\chi \notin \mathsf{Cn}\Gamma$. But the second default is not applicable now. Formula χ is derivable from $\mathsf{Cn}(\mathsf{Cn}\Gamma \cup \{\chi\})$.

Definition 2 (Process). *A sequence of default rules* Π *is a process of a default theory* T *iff the default* d_m *is applicable to* $\mathsf{In}\Pi^m$, *for every* m *such that* $d_m \in \Pi$.[8]

In the example, Π_1, Π_2, and Π_3 are processes.

Definition 3. *Let* Π *be a process.*

- Π *is successful iff* $\mathsf{In}\Pi \cap \mathsf{Out}\Pi = \varnothing$, *otherwise it is failed.*
- Π *is closed iff every* $d \in D$ *applicable to* $\mathsf{In}\Pi$ *(i.e., in the order) is already in* Π.

Processes Π_1, Π_2, and Π_3, from the example, are successful; however, Π_1 is not closed since there is at least one default in D that is applicable. Generally, it can happen that there is an applicable default in a process whose consequent causes the fail. For example, consider a 'strange' theory $\left(\Gamma = \varnothing, D = \left\{\frac{:p}{\neg p}\right\}\right)$, where p is an atomic formula. It has an applicable rule with the consequent $\neg p$, which is in a conflict with the justification p.

Definition 4 (Extension). *A set of formulas* E *is an extension of* T *iff there is a successful and closed process* Π *such that* $E = \mathsf{In}\Pi$.

Extension is a central notion of default logic. It is a deductively closed set containing conclusions of hard information together with consequents of applied defaults. Moreover, we are sure that there is not any applicable default left. As it is seen in our introductory example, there can be more than one extension. Π_2 and Π_3 are successful and closed processes that form two different extensions. Every extension can be understood as a way to extend deduction over hard information consistently.

The most important properties of (general) default theories are:[9]

(1) Let E_1, E_2 be extensions of T and $E_1 \subseteq E_2$, then $E_1 = E_2$ (*minimality of extensions*).
(2) $T = (\Gamma, D)$ has an inconsistent extension if and only if Γ is inconsistent itself (*consistency preservation*).
(3) Let E be an extension of $T = (\Gamma, D)$, then E is an extension of $T' = (\Gamma \cup \Delta, D)$ for every $\Delta \subseteq E$ (*cautious monotony in declaratives*).

Minimality of extensions (1) says that if there are two different extensions, then they must be incompatible. Consistency preservation (2) guarantees that applications of defaults does not produce inconsistency. As a corollary of this property we obtain: If T has an inconsistent extension E, then E is its only extension.[10] And if $T = (\Gamma, D)$ has two different extensions, then Γ is consistent. The last property (3) is a form of monotonicity;

8 We will omit the index j in Π_j whenever it is not necessary to distinguish various default sequences of one default theory.

9 Proofs are easy for finite sets of defaults. See [1, pp. 42–44].

10 Such extension is just the set of all formulas.

we can add hard information, which is based on an extension, and it does not cause any change in 'conclusions' (extensions).

1.2 Normal default theories

There is a special class of defaults called normal default rules. These rules have the general form of

$$\frac{\varphi : \psi}{\psi}$$

Call a default theory normal if it contains only normal default rules. The introductory example is formalized as a normal default theory. Normal defaults cannot cover the full range of non-monotonic reasoning, but they can formalize much from the common reasoning. Above that, normal default theories have many desirable properties. One of the most important is that normal theories have always at least one extension.[11]

Proposition 1 (Existence of extensions)**.** *If $T = (\Gamma, D)$ is a normal default theory, then T has an extension.*

Proof. We will reason that normal theories have always at least one closed and successful process. The case of inconsistent theories and their extensions is clear. Let us consider consistent ones.

First, every process of a default theory T can be extended to a closed process. This is true for general default theories. If there is a process Π and a default $d \in D$, which is applicable, check whether d is in Π. If not, add it. And so on. (For infinite process, see the proof in [1, pp. 33–34].)

Second, if Π is a process of a normal default theory T, then Π is successful. This follows from the form of normal default rules. The applicability check is processed just on the formula that is in the role of the consequent [1, p. 50]. $\square$

Normal default theories are monotonic in sets of defaults. If we add a new default, then we 'only extend' the original extensions [1, p. 50].

Proposition 2 (Monotony in defaults)**.** *Let $T_1 = (\Gamma, D_1)$ and $T_2 = (\Gamma, D_2)$ be normal default theories such that $D_1 \subseteq D_2$. Then each extension E_1 of T_1 is a subset of some extension E_2 of T_2, i.e., $E_1 \subseteq E_2$.*

In our example with Anne studying logic and faculty of arts, we obtained two different extensions that are incompatible, i.e., they are inconsistent together. This is an inherent property of normal theories [1, p. 52].

Proposition 3 (Orthogonality of extensions)**.** *If a normal default theory T has two different extensions E_1 and E_2, then $E_1 \cup E_2$ is inconsistent.*

All the mentioned properties of normal default theories can be considered as reasonable for common 'private' reasoning of an agent. Here we accept the idea that extensions play the role of conclusions, which can be derived from both hard and soft information

11 We have mentioned the (non-normal) theory $\left(\varnothing, \left\{\frac{:p}{\neg p}\right\}\right)$ that has no extensions. It has two processes. The empty process $\langle\,\rangle$ is not closed. The process $\left\langle\frac{:p}{\neg p}\right\rangle$ is closed, but not successful.

and provide a support for decisions. The first property (*existence of extensions*) coincides with the model of an agent that is obliged to do conclusions and decisions. If there are different conclusions, it means that there are different and incompatible ways of doing decisions (*orthogonality of extensions*). If an agent accepts new soft information, it can extend a decision support without destroying it (*monotony in defaults*).

2. Action models

The system of action models we will be using, was published in [7] as a way to describe and formalize epistemic actions. Action model logic is a variant of dynamic logic. Generally speaking, an action formalizes a transition from one epistemic state to (another) epistemic state. For the modelling of epistemic states we use standard propositional *epistemic logic*.[12]

Epistemic logic is a multimodal system extending the classical propositional logic. The language contains a set of atomic formulas $\mathcal{P}$, a finite set of agents $\mathcal{A}$, and formulas defined by BNF:

$$\psi ::= p \mid \neg\psi \mid \psi \to \psi \mid K_i\psi \mid \hat{K}_i\psi$$

where $i \in \mathcal{A}$ can be interpreted as a name of an agent. We use the well-known S5 semantics. Kripke model is a structure $M = (W, R_i, V)$ where W is a non-empty set of possible worlds, R_i is an accessibility relation of an agent $i \in \mathcal{A}$, and V is a valuation function.[13]

The satisfaction relation $\Vdash$ is defined in a standard way:

- $(M, w) \Vdash p$ iff $w \in V(p)$
- $(M, w) \Vdash \neg\psi$ iff $(M, w) \nVdash \psi$
- $(M, w) \Vdash \psi_1 \to \psi_2$ iff $(M, w) \Vdash \psi_1$ implies $(M, w) \Vdash \psi_2$
- $(M, w) \Vdash K_i\psi$ iff $(M, v) \Vdash \psi$, for each v such that wR_iv

Modality K_i represents knowledge of an agent i and modality $\hat{K}_i$ is understood as a dual to K_i:

$$\hat{K}_i\psi \equiv \neg K_i\neg\psi$$

To obtain an action-model version of epistemic logic, we extend the epistemic language by actions that represent the transformation of one epistemic model to another one. The epistemic language will be extended by new modalities $[\alpha]$ and $\langle\alpha\rangle$ where α is an *action*.[14] The semantics is enriched by the following clauses:

---

12 See, e.g., [7] and [6].

13 For simplicity we will use the single-agent setting throughout the paper, except some comments in subsection 3.1.

14 The form of actions will be discussed immediately. For the moment, the reader can imagine that α is a (computer) program as in dynamic logic [4].

- $(M, w) \Vdash [\alpha]\psi$ iff $(M, w) \xrightarrow{\alpha} (M', w')$ implies $(M', w') \Vdash \psi$, for all (M', w')
- $(M, w) \Vdash \langle\alpha\rangle\psi$ iff $(M, w) \xrightarrow{\alpha} (M', w')$ and $(M', w') \Vdash \psi$, for some (M', w')

Let us note that $\langle\alpha\rangle$ is dual to $[\alpha]$. Now we are obliged to explain how to understand the application of an α-action on an epistemic state (M, w), i.e., what is the meaning of $(M, w) \xrightarrow{\alpha} (M', w')$ in *action model logic*.

An action α causes a change of an epistemic state (M, w) to a state (M', w'). This change is conducted by structures (called *action models*) that are very similar to Kripke structures that we use as models for epistemic logic. An action α can be atomic or composite.[15] Atomic actions and composite actions as well are based on action models. An *action model* is a structure

$$\mathbf{M} = (\mathbf{S}, \mathbf{R}_i, \mathsf{pre})$$

where

- $\mathbf{S}$ is a non-empty set of nodes
- $\mathbf{R}_i$ is a binary relation on $\mathbf{S}$, i.e., $\mathbf{R}_i \subseteq \mathbf{S} \times \mathbf{S}$ for each $i \in \mathcal{A}$
- pre is a function assigning exactly one formula to each node ($\mathsf{pre} : \mathbf{S} \longmapsto \mathsf{Fla}$).

Action models have non-empty domains of action-states (nodes) and an accessibility relation for each agent. These relations have the same constrictions and properties as their epistemic counterparts, i.e., since our underlying logic is S5, the relations are reflexive, transitive, and symmetric. The informal interpretation is the same as in plain S5 system, $\mathsf{s}\mathbf{R}_i\mathsf{t}$ means that an agent i cannot distinguish action-states s and t. Unlike Kripke models however, action models do not contain a binary relation between nodes and formulas (valuation). They instead have a unary function, called *precondition*, that assigns a formula to every node of an action model. This precondition formula has to be satisfied in order for the respective action to happen.

Given an epistemic model, one may apply an action model to it. This 'action execution' is governed by a function of restricted modal product $\otimes$ that takes an epistemic model and an action model and creates a new epistemic model.

Definition 5 (Restricted modal product). *Let $M = (W, R_i, V)$ be an* S5 *epistemic model and $\mathbf{M} = (\mathbf{S}, \mathbf{R}_i, \mathsf{pre})$ an* S5 *action model. A restricted modal product $(M \otimes \mathbf{M})$ is an epistemic model $M' = (W', R'_i, V')$ where*

- $W' = \{(w, \mathsf{s}) \mid w \in W \,\&\, \mathsf{s} \in \mathbf{S} \,\&\, (M, w) \Vdash \mathsf{pre}(\mathsf{s})\}$
- $(w, \mathsf{s})R'_i(w', \mathsf{s}')$ *iff* $(wR_iw' \,\&\, \mathsf{s}\mathbf{R}_i\mathsf{s}')$*,*
  for $w, w' \in W$ *and* $\mathsf{s}, \mathsf{s}' \in \mathbf{S}$
- $(w, \mathsf{s}) \in V'(p)$ *iff* $w \in V(p)$*,*
  for $(w, \mathsf{s}) \in W'$ *and atomic formula* $p \in \mathcal{P}$

Thus, in action model logic, the change from an epistemic state (M, w) to a 'new' epistemic state (M', w') can be conducted by an atomic action $\alpha = (\mathbf{M}, \mathsf{s})$:

$$(M, w) \xrightarrow{(\mathbf{M},\mathsf{s})} (M', w')$$

if and only if

$$w \Vdash \mathsf{pre}(\mathsf{s}) \text{ and } (M', w') = ((M \otimes \mathbf{M}), (w, \mathsf{s}))$$

---

15 E.g., for any actions α_1 and α_2, concatenation of two actions $\alpha = (\alpha_1; \alpha_2)$, finite repetition of an action $\alpha = (\alpha_1)^*$, non-deterministic choice of two actions $\alpha = (\alpha_1 \sqcup \alpha_2)$, and others.

The language of action model logic includes epistemic language and atomic actions that are always of the form $(\mathbf{M}, \mathrm{s})$. Composite actions are usually reducible to atomic ones. More about the properties of this system can be found in [7, chapter 6].

3. Normal default actions

In action model logic we work with the idea that there is a group of agents and these agents change their epistemic states with respect to 'new' information produced by verbal as well as non-verbal actions. There is a slightly modified picture in the introductory example. We played the role of an agent reasoning about 'Anne's liking for math'. We called this reasoning 'private' because it mostly happens inside our head without any connection with other agents. Nonetheless, it need not be completely private, other agents can know both hard and soft information and can follow our steps in reasoning. Even if we do not show publicly the direction of our thoughts, the other agents have to consider all possibilities in their epistemic states.

To incorporate default reasoning inside action model logic we have to understand default rules as actions.[16] Someone (let us call the agent i) can use the default rule

$$\frac{\mathrm{AnneStudentOfArtFaculty} \;:\; \neg\mathrm{AnneLikeMath}}{\neg\mathrm{AnneLikeMath}}$$

in a situation whenever i is not sure whether Anne likes math or not, but there is no information, which is in conflict with the justification that *Anne does not like math*. Simultaneously, the prerequisite *Anne is a student of a faculty of arts* is considered by i as a 'knowledge'. If the rule is applied, then i narrows down the set of possible worlds that are indistinguishable for i. Now, the agent accepts *Anne does not like math* as a (new) 'knowledge'. In fact, if a normal default rule is applied by i, then i changes her epistemic situation that forms a base for a possible application of other defaults. More formally, a normal default rule $\frac{\varphi:\psi}{\psi}$ changes agent's epistemic model such that it separates ψ-worlds from $\neg\psi$-worlds and the agent is ready to work with preferred ψ-worlds from now on. If the agent i goes on with some other default, she only checks the validity of a (new) prerequisite with respect to the possible worlds where ψ is true. Of course, it does not mean that $\neg\psi$-worlds are canceled. They are now distinguishable for i and can be important from the viewpoint of other agents' knowledge.

This idea brings us to a small modification of epistemic models. Every agent will have a set of designated (preferred) possible worlds that are the basis of defaults' applicability.

Definition 6. *A (default) epistemic model M is a structure (W, R_i, V, X_i) where W is a non-empty set of possible worlds, R_i is an accessibility relation of an agent i, V is a valuation function, and $X_i \subseteq W$ is a non-empty set of designated possible worlds for an agent i such that whenever uR_iv, it holds that $u \in X_i$ iff $v \in X_i$.*

[16] Defaults cause changes, which exceed the deductive base of a background formal system that is described by epistemic logic.

The designated set $X_i \neq \emptyset$ marks the states that are 'important' to agent i. In the agent's reasoning, i.e., in application of a default rule, the agent 'ignores' all the states outside of X_i. The designated worlds are not connected via R_i to those that are not designated.

This leads us to a formal solution of the question when a normal default rule $\frac{\varphi : \psi}{\psi}$ is *epistemically applicable* by an agent. An agent has to 'know' the prerequisite φ and consider the justification ψ as unknown, but possible, with respect to the set of the agent's designated worlds.

Definition 7. *A normal default rule* $\frac{\varphi : \psi}{\psi}$ *is* epistemically applicable *by an agent i in an epistemic model M iff for each* $w \in X_i$*:*

- $(M, w) \Vdash K_i\varphi$
- $(M, w) \Vdash \hat{K}_i\psi$
- $(M, w) \Vdash \hat{K}_i\neg\psi$

In other words, formulas $K_i\varphi$, $\hat{K}_i\psi$, and $\hat{K}_i\neg\psi$ are valid in the submodel of M generated by a set X_i.[17]

Our term of (epistemic) *applicability* does not fully correspond to the applicability in default logic. A default rule $\frac{\varphi : \psi}{\psi}$, where the justification ψ is known by an agent, could be applicable in default logic but from the epistemic point of view it does not do any change of agent's epistemic state. Such default rules would be in some sense hollow—empty thinking about things that are already known. However, we want defaults to decide unknown things.[18]

Actions based on normal default rules are of another nature than atomic actions in action models logic. They do not depend on one (action) node and its precondition. The applicability of a default takes over the role of precondition.

An action corresponding to a normal default $\frac{\varphi : \psi}{\psi}$ (used by an agent i) will be understood as a two-node action model

$$\mathbf{D}^i = (\mathbf{S}, \mathbf{R}_i, \mathsf{pre}, \mathbf{X}_i)$$

where

- $\mathbf{S} = \{\mathsf{s}, \mathsf{t}\}$
- $(\mathsf{s}, \mathsf{t}) \notin \mathbf{R}_i$, but $(\mathsf{s}, \mathsf{s}) \in \mathbf{R}_i$ and $(\mathsf{t}, \mathsf{t}) \in \mathbf{R}_i$
- $\mathsf{pre}(\mathsf{s}) = \psi$ and $\mathsf{pre}(\mathsf{t}) = \neg\psi$
- $\mathbf{X}_i = \{\mathsf{s}\}$

The action $\mathbf{D}^i$ has two action-states, which are not connected by the relation $\mathbf{R}_i$ (the agent i can distinguish these two states), and they differ in preconditions with respect to formula ψ. The new aspect is the set $\mathbf{X}_i$, which corresponds to the set of designated worlds in an epistemic model. $\mathbf{X}_i$ contains action states that have the formula ψ as their preconditions. In the simple case of normal defaults, the designated set $\mathbf{X}_i$ contains only one action-state s whose precondition is ψ.[19]

The idea will be complete after we describe how the action $\mathbf{D}^i$ works. We have emphasized that a normal default action is different from atomic actions in action model logic.

17 Compare the notion of 'R-applicability' in [5].

18 Similarly, multiple occurrence of a default is not allowed in default logic.

19 A two-agent version will be mentioned in subsection 3.1.

It is not an update that changes one particular epistemic state (M, w) into a (new) epistemic state (M', w'). Normal default actions operate on whole models and, thus, change epistemic background.

For an epistemic model $M = (W, R_i, V, X_i)$ we define

$$M \xrightarrow{\mathbf{D}^i} M'$$

if and only if

(1) the corresponding default rule is applicable (Definition 7) and
(2) $M' = (M \otimes \mathbf{D}^i)$.

The resulting epistemic model M' will be formed as it is described in Definition 5. The only thing we have to add is how to form the new set of designated worlds X_i'. For each $w \in W$ and $\mathsf{x} \in \mathbf{S}$:

$$(w, \mathsf{x}) \in X_i' \text{ iff } w \in X_i \ \& \ \mathsf{x} \in \mathbf{X}_i$$

The action based on action model $\mathbf{D}^i$ causes the resulting epistemic model M' to contain two parts (submodels) that are disjoint for the accessibility relation R_i. Let us now consider normal defaults $\frac{\varphi : \psi}{\psi}$ and $\frac{\varphi : \neg\psi}{\neg\psi}$, for example. We will write $(\varphi : \psi/\psi)^i$ and $(\varphi : \neg\psi/\neg\psi)^i$ as actions based on these defaults for an agent i. Both of them are applicable (by the agent i) under the same conditions, cf. Definition 7.[20] The corresponding action models are almost the same. The difference lies in sets of designated worlds. The first default action model requires to have designated nodes with the precondition ψ and the other one with $\neg\psi$.

If the action $(\varphi : \psi/\psi)^i$ is applied on an epistemic model $M = (W, R_i, V, X_i)$ (by an agent i), then two separated parts are formed in the new epistemic model $M' = (W', R_i', V', X_i')$. No world from one part is connected by R_i' to any world from the other part. One of the parts consists of ψ-worlds and these worlds are designated (X_i'), the other one consists of $\neg\psi$-worlds. Informally, the agent i did a decision whether ψ or not and preferred ψ as true. If the agent i is in any new epistemic state $w' \in X_i'$, then the formula $K_i\psi$ is true there. Similarly for the formula $K_i\neg\psi$ in states out of X_i'.

In case we know that an agent i is in a particular epistemic state (M, w), then the application of either $(\varphi : \psi/\psi)^i$ or $(\varphi : \neg\psi/\neg\psi)^i$ depends on whether $(M, w) \Vdash \psi$ or $(M, w) \Vdash \neg\psi$. A particular epistemic state provides preferences among default actions based on preconditions.

From the viewpoint of an agent i and her epistemic model $M = (W, R_i, V, X_i)$, a default theory (Γ, D) means that $(M, w) \Vdash K_i\gamma$, for each $w \in X_i$ and each $\gamma \in \Gamma$. At the very beginning, before the use of any default, all states are designated ($X_i = W$). The final epistemic model is the result of 'step by step' applications of default actions given by a successful and closed process. If $\Pi = \langle d_1, \dots, d_k \rangle$ is a successful and closed process, we obtain the final epistemic model (for an agent i) by the concatenation of corresponding actions $(\mathbf{D}_1^i; \mathbf{D}_2^i; \dots; \mathbf{D}_k^i)$:

---

20 With respect to Definition 7 we can make a terminological convention. The meaning of 'a default rule $\frac{\varphi : \psi}{\psi}$ epistemically applicable by an agent i in an epistemic model' is the same as 'a default action $(\varphi : \psi/\psi)^i$ applicable (by i) in an epistemic model'.

$$M \xrightarrow{\mathbf{D}_1^i} \cdots \xrightarrow{\mathbf{D}_k^i} M'$$

The knowledge of the agent i in a submodel (of M') generated by X_i' corresponds to the extension $\mathsf{In}\Pi$.[21]

Let us consider a normal default theory $\left(\varnothing, \left\{\frac{:p}{p}, \frac{:\neg p}{\neg p}, \frac{p:q}{q}\right\}\right)$ of an agent i. There is no hard information and p, q are atomic formulas. We can form two successful and closed processes: $\left\langle \frac{:p}{p}, \frac{p:q}{q} \right\rangle$ and $\left\langle \frac{:\neg p}{\neg p} \right\rangle$. Thus, we have two possible updates of an epistemic model $M = (W, R_i, V, X_i)$. The first one is given by concatenation of two normal default actions $((:\ p/p)^i; (p\ :\ q/q)^i)$ and the second one by the single action $(:\ \neg p/\neg p)^i$. Now all depends on the applicability and agent's decision which one will be done.

If there is a concatenation of normal default actions, then the erasing of an accessibility relation is executed all over the model. For example, whenever i executes $((:\ p/p)^i;$ $(p\ :\ q/q)^i)$, then, after the second action $(p\ :\ q/q)^i$, the relation R_i does not connect q- and $\neg q$-words in both parts formed by the first action $(:\ p/p)^i$. The final set of designated worlds will contain $(p \wedge q)$-worlds.

3.1 Examples

In formalisms and examples throughout the text we used, in fact, a single-agent variant. Our aim was to introduce the basic idea how to use normal defaults as actions. Nonetheless, it will be useful to present some notes concerning a multi-agent variant.

The language of epistemic logic, which we have introduced, is multi-agent friendly. Knowledge operators as well as accessibility relations are indexed by agents' names. Since default rules correspond to 'private' reasoning acts, we use the same indexing by the name of an agent for normal default actions. Nonetheless, we did not introduce group knowledge operators and that is the reason why we are not going to discuss all aspects of the multi-agent setting. We do not solve what is the essence of a default theory, whether 'hard information' is commonly known among agents, whether default rules are shared in a group, and similar questions. For the following examples, let us imagine that agents do their default reasoning individually. Agents do not communicate; however, they may (privately) follow the reasoning of other agents.

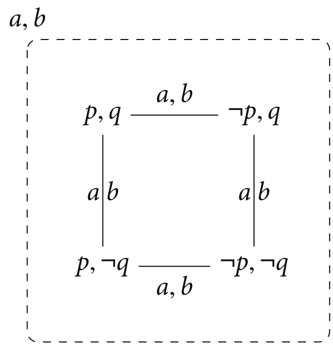
We will show two examples that will present what must be considered even in this simplified epistemic setting. The group of agents will contain just two agents, let us call them Alice (a) and Bob (b). Now, a (default) epistemic model is a structure $M = (W, R_a, R_b, V, X_a, X_b)$ and a normal default action used by the agent a, for example, is a structure $\mathbf{D}^a = (\mathbf{S}, \mathbf{R}_a, \mathbf{R}_b, \mathsf{pre}, \mathbf{X}_a, \mathbf{X}_b)$ where the behavior of this structure depends on the type of reasoning, see Example 2 for an additional commentary. If the reasoning of a does not depend on the activity of b, then a normal default action corresponding to a normal default $\frac{\varphi:\psi}{\psi}$ used by the agent a is the structure $\mathbf{D}^a = (\mathbf{S}, \mathbf{R}_a, \mathbf{R}_b, \mathsf{pre}, \mathbf{X}_a, \mathbf{X}_b)$ where

- $\mathbf{S} = \{\mathsf{s}, \mathsf{t}\}$
- $(\mathsf{s}, \mathsf{t}) \notin \mathbf{R}_a$, but $(\mathsf{s}, \mathsf{s}) \in \mathbf{R}_a$ and $(\mathsf{t}, \mathsf{t}) \in \mathbf{R}_a$

21 Formal details will not be discussed in this introductory paper.

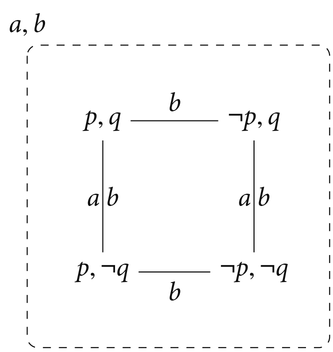
- $(\mathsf{x}, \mathsf{y}) \in \mathbf{R}_b$, for each $\mathsf{x} \in \mathbf{S}$ and $\mathsf{y} \in \mathbf{S}$
- $\mathsf{pre}(\mathsf{s}) = \psi$ and $\mathsf{pre}(\mathsf{t}) = \neg\psi$
- $\mathbf{X}_a = \{\mathsf{s}\}$
- $\mathbf{X}_b = \mathsf{S}$

It means that b's accessibility relation and set of designated worlds will not be changed.
Example 1 In the first example we present the situation where Bob has a default theory $\left(\varnothing, \left\{\frac{:K_a p}{K_a p}\right\}\right)$, i.e., Bob has no hard information, but a default rule on Alice's knowledge of (an atomic fact) p. Consider the following epistemic model:

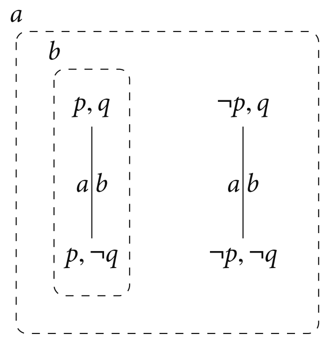


Neither Bob nor Alice know anything about p or q. All four epistemic worlds are indistinguishable for them and their sets of designated worlds are the same (indicated by the dashed line). The default action $(:\ K_a p/K_a p)^b$ is not applicable by Bob in this scenario. He does not admit the epistemic possibility that Alice knows p, see the second condition in Definition 7.

If Alice can distinguish p-worlds and $\neg p$-worlds, as in the following figure, then the action is applicable by Bob.

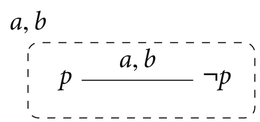


After we apply the default rule we obtain a new epistemic model:

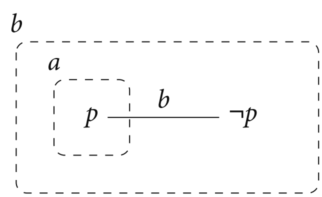


Bob's set of designated worlds has changed and from now on he works with the fact that Alice knows p.

Example 2 The second example presents private and semi-private reasoning and their differences. Recall our example with Anne, a student of logic at Faculty of Arts. Let's consider our two agents, Alice and Bob, who discuss whether Anne likes mathematics or not. They both know the hard information that Anne studies logic at Faculty of Arts. If we label the fact that Anne likes mathematics as p, the situation might look like this:

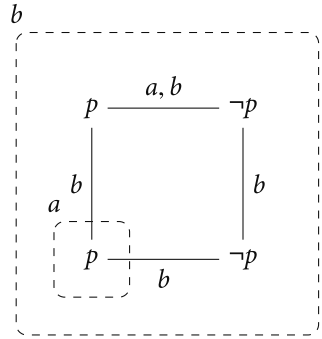


Alice and Bob consider both p and $\neg p$ possible and include both situations in their respective designated sets. After a short discussion Alice decides that she prefers to use the default reasoning that Anne indeed does like mathematics. Bob has access to the same default rules as Alice but has not decided yet.



In this model Alice's designated set contains only the state where p holds. Bob's designated set and accessibility relations were unchanged. However Alice's deduction was in some sense public, or semi-private. Bob knows that Alice knows something new, i.e., the formula $K_b(K_a p \vee K_a \neg p)$ holds in the whole model, resp. in the submodel generated by b's set of designated worlds. We may obtain this model simply by applying the default action $(: p/p)^a$.

Alternatively we may consider that Alice's reasoning is private. In this case Bob's knowledge will in some sense decrease.



Now, Bob has no clue about Alice's knowledge. He doesn't even know whether Alice 'used' a default reasoning or not.

This situation (private reasoning) may be achieved by a slight change in the definition of the default action model. Along the two action states $\{\mathsf{s}, \mathsf{t}\}$ there is a new state u whose precondition is $(p \vee \neg p)$. Bob's accessibility relation will be universal, i.e., any two states are in the relation $\mathbf{R}_b$.[22] The action model structure for a's private use of $(:\ p/p)^a$ is the following $\mathbf{D}^a = (\mathbf{S}, \mathbf{R}_a, \mathbf{R}_b, \mathsf{pre}, \mathbf{X}_a, \mathbf{X}_b)$ where

- $\mathbf{S} = \{\mathsf{s}, \mathsf{t}, \mathsf{u}\}$
- for all $\mathsf{x} \in \mathbf{S}$, just $(\mathsf{x}, \mathsf{x}) \in \mathbf{R}_a$
- $(\mathsf{x}, \mathsf{y}) \in \mathbf{R}_b$, for all $\mathsf{x} \in \mathbf{S}$ and $\mathsf{y} \in \mathbf{S}$
- $\mathsf{pre}(\mathsf{s}) = p$ and $\mathsf{pre}(\mathsf{t}) = \neg p$ and $\mathsf{pre}(\mathsf{u}) = (p \vee \neg p)$
- $\mathbf{X}_a = \{\mathsf{s}\}$
- $\mathbf{X}_b = \mathsf{S}$

4. Conclusion and further research

This paper focuses on showing that normal default reasoning and action models can share a common ground. We investigated the relationship between normal defaults and action models and proposed a way to 'translate' normal default rules into (default) actions in the framework of action models logic. The presented system shows the role of normal defaults in a simple semi-private reasoning or, alternatively, in a completely private way of thought of an agent.

A lot of further research remains in this field. One may obviously investigate general default rules or semi-normal defaults. An interesting generalization stems from the behavior of group epistemic modalities like common knowledge and distributive knowledge that are important for communication among agents. This was mentioned in the previous section.

Another thing to consider is that some rules might be more informative or more 'correct' than other rules. We would want the agents to apply these better defaults before they apply any others. This can be achieved by a preference function. Each agent would have a preference function that would order each of her defaults by preference. This preference

22 Compare the action model and the resulting epistemic model in [7, p. 153, Example 6.13].

function might even be interactive. For example if agent a sees that agent b used default d_1, she might be more inclined to use the same default d_1 instead of d_2. These ideas bring us to a question whether there is a correspondence to belief revision and to a (technical) problem of the combination of default actions with other actions in this framework.

References

[1] Grigoris Antoniou. *Nonmonotonic Reasoning*. The MIT Press, 1997.

[2] Gerhard Brewka, Jürgen Dix, and Kurt Konolige. *Nonmonotonic Reasoning: An Overview*. CSLI, Stanford, California, 1997.

[3] Dov M. Gabbay, C. J. Hogger, and J. A. Robinson, editors. *Handbook of Logic in Artificial Intelligence and Logic Programming*, volume 3. Clarendon Press, Oxford, 1994.

[4] David Harel, Dexter Kozen, and Jerzy Tiuryn. *Dynamic Logic*. The MIT Press, 2000.

[5] Witold Lukaszewicz. *Non-monotonic Reasoning: Formalization of Commonsense Reasoning*. Ellis Horwood, 1990.

[6] John Jules Ch. Meyer and Wiebe van der Hoek. *Epistemic Logic for AI and Computer Science*. Cambridge University Press, 1995.

[7] Hans van Ditmarsch, Wiebe van der Hoek, and Barteld Kooi. *Dynamic Epistemic Logic*. Springer, Dordrecht, The Netherlands, 2008.

2015 ACTA UNIVERSITATIS CAROLINAE PAG. 39–53
PHILOSOPHICA ET HISTORICA 1 / MISCELLANEA LOGICA X

BIRKHOFF'S AESTHETIC MEASURE

VERONIKA DOUCHOVÁ*
Dept. of Logic, Faculty of Arts, Charles University in Prague
* E-mail: veronika.douchova@centrum.cz

ABSTRACT
In this paper, we review and critically evaluate George D. Birkhoff's work concerning formalisation of aesthetics, as it appeared in his book *Aaesthetic Measure* from 1933, and discuss its influence on further research in the field. In the book, Birkhoff defines an *aesthetic measure M* of an art object as the ratio between its order and complexity, or more generally a function f of this ratio: $M = f\left(\frac{O}{C}\right)$, where O stands for order and C for complexity. The specific definitions of O and C depend on the type of the analysed object. Birkhoff applied the formula to multiple classes of objects (e.g. vases, music, or English poetry) and calculated the aesthetic measure for many art objects from these classes.
We give an example of Birkhoff's analysis using polygons, and we further discuss to what extent the ordering of the polygons (or other objects) according to the resulting measures can be used, or interpreted, as an ordering according to a degree of aesthetic preference.
We also include an extensive bibliography, supplemented by a critical discussion of the influence of Birkhoff's work on further research.
Keywords: aesthetic measure, information complexity, theory of information, rational aesthetics, information aesthetics

1. Introduction

In his book *Aesthetic measure* [Bir33], Birkhoff defines an aesthetic measure and applies it to several types of objects with different modes of perception – visual, including 3D objects, and auditory (music, poetry). The measure is defined in relation to the effort which the object demands of the perceiver (complexity), and the pleasing or displeasing features which can be recognised in the object (order).

George David Birkhoff (1884–1944) was a distinguished mathematician, who worked predominantly in the fields of algebra, dynamic systems and number theory. He is best known for his proof of the general form of the Poincare-Birkhoff theorem (1913), for his book *Dynamic Systems* (1927), or for his proof of characterisation of monoids (1935).

Aesthetic Measure was published in 1933, but some of the results were presented earlier on conferences and in papers (starting from 1928). In the introduction to [Bir33], Birkhoff states that he started to be interested in the structural aspects of aesthetic perception while listening to music (mostly classical), almost 30 years previously. He realised that the order, or pattern, of the tones plays an important role in aesthetic perception of music. He later refined these ideas into a theory and applied it to other forms of aesthetic objects (vases, tiles or polygons – visual art, or poetry – auditory art).

Some sources state that Birkhoff based his theory on the work of a Canadian-American artist Jay Hambidge, as he formulated it in his book *Dynamic Symmetry* (1926, Harvard). This information is given for instance in H. J. McWhinnie's *A Review of Research on Aesthetic Measure* [McW68]. Birkhoff himself does not name Hambidge as an inspiration; however, he mentions his name twice in the fourth chapter of the book: *Chapter IV: Vases*. On page 67, he quotes Hambidge's work *Dynamic Symmetry of the Greek Vases*, New Haven and New York, 1920, in connection to Greek vases, and then at the end of chapter *The Appreciable Elements of Order*, where he discusses properties of vases which have impact on the order of the aesthetic measure (and at this place, he disagrees with Hambidge [Bir33, p. 72]).

In our paper, we introduce Birkhoff's aesthetic measure using polygons. The reason is that the majority of contemporary papers focuses on vases, according to Birkhoff's Chapter 4, [Bir33, pp. 67–86]. A modern presentation of the aesthetic measure regarding polygons has been mostly missing from the literature.

2. The aesthetic measure

According to Birkhoff, the first impulse for his study of the aesthetic quality of objects in the context of a semi-mathematical theory was the act of listening to music with reflections on the melody pattern. These reflections were the beginning of work which was

> "… to bring the basic formal side of art within the purview of the simple mathematical formula defining aesthetic measure" [Bir33, p. viii]

According to Birkhoff, aesthetic experience consists of three primary consecutive stages:

> "… (1) a preliminary effort of attention, which is necessary for the act of perception, and which increases in proportion to what we shall call the *complexity (C)* of the object; (2) the feeling of value or *aesthetic measure (M)* which rewards this effort; and finally (3) a realisation that the object is characterised by a certain harmony, symmetry, or *order (O)*, more or less concealed, which seems necessary to aesthetic effect." [Bir33, pp. 3–4]

Thus the aesthetic quality is in relation to the attention which is required to perceive the object in its entirety, and is counterbalanced by the notion of order in the object. The load of attention grows in proportion to the complexity of the object, and therefore Birkhoff denotes this property by *complexity, C,* of the object. *Order, O,* of the object is a counterbalancing quantity, often found in the forms of harmony or symmetry. Birkhoff postulates that the aesthetic measure M is the ratio of these two quantities: to preserve a fixed value of the measure, higher complexity of an object must be counterbalanced by an increased order, and conversely, simpler objects require a smaller value of order for the same effect. More generally, Birkhoff defines the aesthetic measure as a function f of this ratio:

$$M = f\left(\frac{O}{C}\right).$$

Birkhoff did not elaborate on the specific form of f, and in his work he essentially identifies f with the identical function.[1] As we discuss later, the range of M consists of rational numbers, with the usual ordering on rational numbers being implicitly used to order the values of M.[2]

Birkhoff applies his formula in a general setting, irrespective of the mode of perception (visual, auditory) or of the type of object. In his understanding, the formula is universal and transferable between art forms. In his book, he gives examples of objects perceived visually (always plane objects, or objects mapped to a plane), musical objects and melodies, and poetry (which he finds similar to music in his study). In every area, he delimits a class of objects, usually a very narrow class, and formulates for them concrete definitions of order and complexity in order to calculate M. The exact definitions of order and complexity must be chosen carefully to ensure that the resulting M reflects the aesthetic quality of the object; this choice is considered by Birkhoff as the basic problem in aesthetics [sic!].

The first part of aesthetic experience is the initial exertion which is required to perceive an object. As we stated earlier, this exertion is in proportion to the complexity C of the object. Birkhoff maintains that the act of perception leading to an aesthetic feeling necessarily requires a conscious exertion of attention. The amount of this attention corresponds to C. More exactly, the value of C is the sum of all types of exertions multiplied by the number of their occurrences. This leads Birkhoff to defining complexity as the number of units in the object which require a conscious act of attention (e.g. number of tones in a melody or number of edges of a tile). Complexity has an impact on the resulting measure M, which is viewed as a reward for the exertion of attention. This relationship between C and M is counterbalanced by the quantity of order in the object, which can compensate for a higher complexity.

The order O in the object is considered to be a conscious part of the aesthetic process. O is characterised by the pleasant feelings associated with the exertion of attention (corresponding to C). In order to determine the value of order in the object, it is necessary to distinguish two types of associations: formal and connotative.

Formal associations are defined as those associations which are implied by basic properties of objects, such as symmetry, repetition, similarity, contrast, identity, balance, repetitive parts, i.e. properties which invoke pleasant feeling in the act of attention, but also by properties which invoke negative feelings (lack of reward in attention), such as lack of clarity, unpleasant repetition, inessential imperfection, or dissonance in music. The value of O is thus the sum of all formal associations multiplied by the number of occurrences; formal associations occur in this sum either with the positive sign (pleasant

[1] Therefore, in this paper, we use a simplified formula $M = \frac{O}{C}$.

[2] Birkhoff does not specifically discuss this issue, but only mentions some related questions. See the end of Section 2.2 below.

associations) or with the negative sign (unpleasant associations). Indifferent associations have the zero value. Birkhoff specifies concrete associations, and their effect on O, for various types of objects.

Connotative associations do not take part in determining the aesthetic measure. They are defined as such associations that are not formal, i.e. which are not implied by a basic property of the object. Among connotative associations, we can include the usefulness of polygons mentioned by Birkhoff on page 29, where he states that

> "… usefulness corresponds to a connotative factor entirely outside of the scope of the theory." [Bir33, p. 29]

Birkhoff does state that every association becomes an element of the order, disregarding whether it is a formal or a connotative association, but in his definition of the order O for the purposes of M he only allows formal associations.

The notions of order and complexity need to be applied in concreteness to a narrow class of objects to yield a reasonable notion of comparison between these objects. Birkhoff chooses in the book classes of objects such as polygons, ornaments, tiles, or vases. For vases, Birkhoff defines the notions of order and complexity for planar cuts intersecting the axis of a vase, i.e. for planar curves which by rotation draw the contour of the vase (handles and spouts are not considered). In music, based on specific definitions of C and O, Birkhoff calculates M for diatonic scales and chords, harmonies and melodies. From music, Birkhoff finally turns to poetry.

In the next section, we give for illustration of Birkhoff's methods more details for one specific class of objects – polygons.

2.1 Polygons

Birkhoff describes the aesthetic measure of polygons in Chapter II, [Bir33, pp. 16–48]. Polygons are the first class of objects for which he gives details regarding the calculation of M. At the beginning of the chapter, he states that while polygons are often considered merely as geometrical objects, they do have an aesthetic quality which can be used to compare polygons among themselves. As we said earlier, classes of objects for which we calculate M should be defined as narrowly as possible; for this reason Birkhoff limits his attention to plane shapes which can be used as tiles. Moreover, he requires that all polygons should have a similar size, and he disregards colours and materials. These limitations should remove as many connotative associations as possible. Birkhoff also eliminates the role of an observer in order to disregard the role of cultural background or education. This method allows him to evaluate plane shapes more objectively (but he also disregards symmetries which may be recognisable only to an educated or experienced observer).

The formula for calculation of M for polygons is defined as follows:

$$M = \frac{O}{C} = \frac{V + E + R + HV - F}{C},$$

where C is the complexity of the polygon, O is the order defined by the vertical symmetry V, balance E called equilibrium by Birkhoff, rotational symmetry R, horizontal-vertical network or grid HV and non-pleasing or unsatisfactory form F. Parallel edges and the horizontal symmetry are classified as neutral, i.e. they are not included in the computation of M. The properties participating in the order are concerned mostly with movements on the plane and less with properties which we can call numerical. The numerical properties are not completely ignored, though. Sometimes they are captured by other properties: for instance for an equilateral triangle, Birkhoff does not consider the equality of the sides, but they do have a role in the calculation of possible rotations and symmetries.

The complexity C is defined as the number of lines on which there lie all edges of the polygon. Note that this is not the same as the number of edges of the polygon (consider for instance a six-pointed star or the Greek cross).

The vertical symmetry V is considered as a positive property, and therefore is calculated with the positive sign. An object has the vertical symmetry 1 if it is symmetrical along the vertical axis. Otherwise the symmetry has value 0.

An overall balance of the object is denoted E and relates to the visual sense of balance, not necessarily physical balance, and takes values $-1, 0, 1$. The visual balance is evaluated according to the position of the equilibrium. The highest value $E = 1$ is assigned when the equilibrium of the object lies between the vertical lines crossing the extremal points of the polygon, and the distance to these lines is at least $1/6$ of the horizontal width of the object. In particular, if $V = 1$, then $E = 1$. If the equilibrium lies between the lines, but in a smaller distance, the value of E is set to be 0. Otherwise the value is -1.

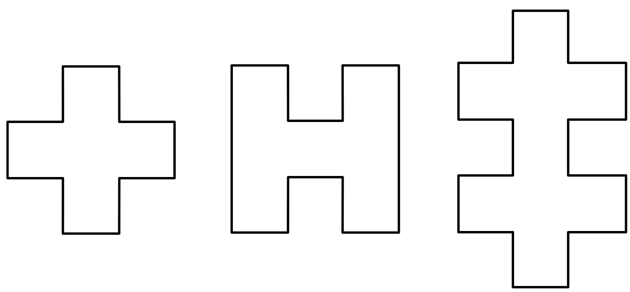


Figure 1: Aesthetic measure from left to right: $M = 0.75$, $M = 0.5$ and $M = 0.4$.[3]

The quantity R captures the rotational symmetries of the polygon. If there exists a rotational symmetry, then q denotes the number of possible rotations before the polygon returns to its original position, i.e. $\frac{360^\circ}{q}$ is the least possible angle of rotation. R is calculated from q as follows:

$$R = \begin{cases} \frac{q}{2} & \text{if (1) there is a rotational symmetry, (2) } q \leq 6 \text{, and (*)} \\ 3 & \text{if (1) there is a rotational symmetry, (2) } q > 6 \text{, and (*),} \\ 1 & \text{for other cases if a rotational symmetry exists and } q \text{ is even,} \\ 0 & \text{otherwise.} \end{cases}$$

[3] This and following figures were created based on the illustrations in [Bir33].

where (*) means that the polygon itself has a vertical symmetry or the convex hull of the given polygon has a vertical symmetry and the concave parts of the given polygon do not adjoin to the vertices of the convex hull.

HV (Horizontal-Vertical grid) is related to the movements of polygons on the plane. HV is considered as positive if the polygon can be moved vertically and/or horizontally to a new position while preserving its relative position in the same horizontal-vertical grid. The value can be $2, 1, 0$. Birkhoff defines that HV takes the value 2 if all edges of the polygon lie in parallel to the vertical-horizontal grid (grid consisting of horizontal and vertical lines). Typical examples are rectangles, Greek cross or a polygon in the form of the letter H. If instead of the vertical-horizontal grid, we consider a grid composed of parallel lines with the same angle with respect to the vertical lines and all edges of the polygon lie on this grid, we set $HV = 1$ (a typical example is a diamond, i.e. an equilateral parallelogram positioned on its vertex). In addition, HV equals 1 if the polygon is situated in a vertical-horizontal grid or in a grid consisting of parallels with the vertical line, but one direction of the edges is diagonal to the gird (or two directions are diagonal to the grid), or finally if some of the edges are not exactly parallel with the lines in the grid. Otherwise, we set $HV = 0$.

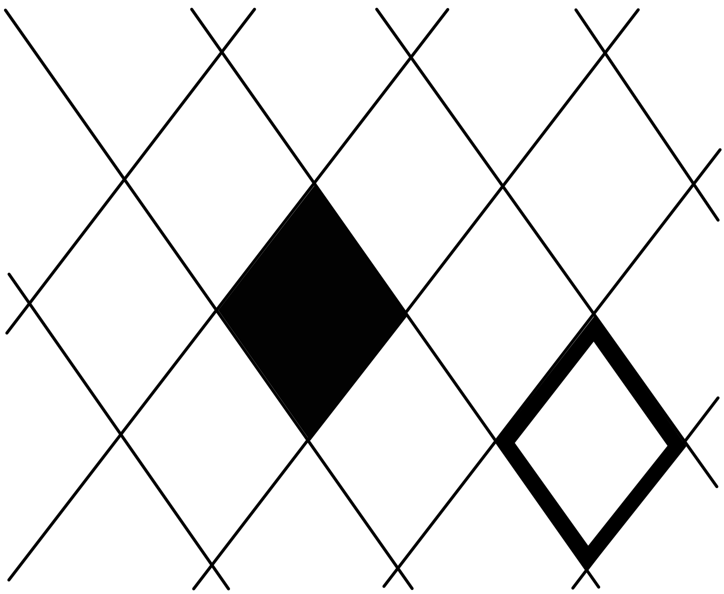


Figure 2: The diamond with the aesthetic measure $M = 1$.

The last quantity participating in O is the non-pleasing (unsatisfactory) form F. Birkhoff refers to it as the *omnium gatherum* of all negative elements of the order. F takes the negative sign, and values $2, 1, 0$. To achieve $F = 0$, the polygon must satisfy the following conditions:

(1) There are no distances between vertices, edges and between a vertex and an edge which are too small. A distance is too small if it is less than $1/10$ of the maximal distance between the vertices of the polygon.
(2) The angles between the edges are not too small (set as less than $20°$).
(3) There are no irregularities. This is defined that no movement of a vertex by a small amount (less than $1/10$ of the distance to the nearest vertex) may impact the values of V, R or HV.
(4) There are no projecting edges.
(5) There is at most one type of a concave angle.
(6) If we view the vertical and horizontal directions as one direction, then the polygon cannot have more than two types of direction.
(7) There is a symmetry which prevents both V and R from taking the value 0.

The reader may notice vague quantifications "too small" or "small amount" occurring in three conditions above. It seems that the exact value of "too small" is not very important for the method – Birkhoff uses the phrase "… for definiteness we shall demand …", [Bir33, p. 41].

If the polygon violates one condition, we set $F = 1$. If more than one condition fails, we set $F = 2$.

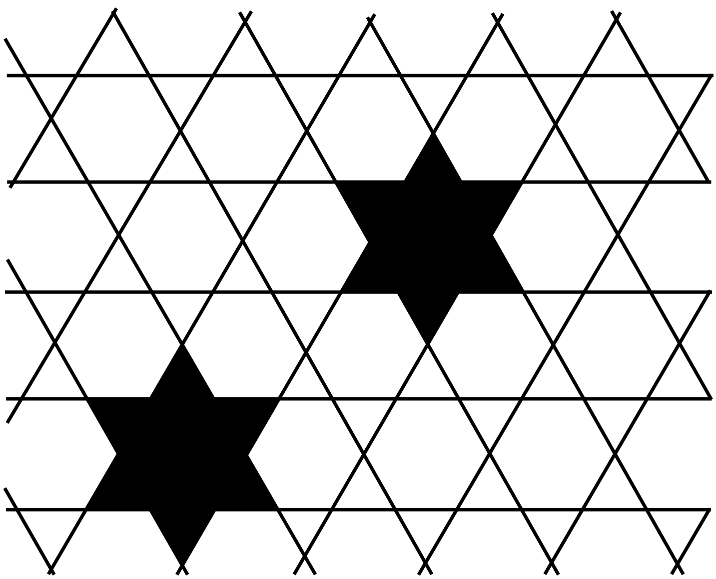


Figure 3: A regular six-pointed star positioned on a vertex, with the grid. $M = 1$.

We illustrate the computation of M with the example of a six-pointed star positioned on a vertex. Complexity C equals to the number of lines on which the edges lie, i.e. $C = 6$. The star is symmetric along the vertical line, therefore $V = 1$ and also $E = 1$. The star can be rotated by $60°$, i.e. $q = 6$ and $R = 3$. The star satisfies the condition that the polygon is situated in a vertical-horizontal grid or in a grid consisting of parallels with the vertical line, but two directions of the edges are diagonal to the grid, therefore $HV = 1$. The form is satisfactory, with $F = 0$. The aesthetic measure of a six-pointed star positioned on a vertex is therefore:

$$M = \frac{1 + 1 + 3 + 1 - 0}{6} = 1.$$

2.2 Comparing objects according to their measure

We will use the specific example of polygons to determine which polygon has the greatest value of the aesthetic measure on the rational line. Birkhoff discusses related questions in the chapter *30. The Mathematical Treatment of Aesthetic Questions* on pages 46–47. A similar discussion, with minor inaccuracies, can be found in [Bar03].

The complexity C can have many values for polygons, but it is always at least 3. $C = 3$ if the edges of the polygon lie on three lines, i.e. if the polygon is a triangle. According to [Bir33], the aesthetic measure of a triangle is an element of the following set: $\{7/6, 2/3, 0, -1/3, -2/3, -1\}$, i.e.

$$M(\text{triangle}) \leq \frac{7}{6},$$

where the greatest value is assigned to the equilateral triangle, with $M = 7/6$.

In general, we have the following values for polygons: $V \leq 1$, $E \leq 1$, $R \leq 3$, $HV \leq 2$, and $F \in \{0, 1, 2\}$, and therefore $O \leq 7$ and

$$M(\text{polygon}) \leq \frac{7}{C}.$$

For the square, we have $C = 4$, $V = 1$, and so $E = 1$, $R = 4/2 = 2$, $HV = 2$, and $F = 0$, hence

$$M(\text{square}) = 1.5$$

We know that $M(\text{triangle}) \leq 7/6$, and $M(\text{square}) = 1.5$, and therefore no triangle (i.e. a polygon with $C = 3$) can have its measure greater than the square. If $C \geq 5$, we get $M \leq 7/5 = 1.4$. The polygon with the greatest measure must therefore be found among four-edged polygons whose edges lie on 4 lines, with two and two lines being parallel to each other. This holds only for the square ($M = 1.5$), a general rectangle ($M = 1.25$), and diamond ($M = 1$). In all other cases, $HV = 0$ and either V or R are equal to 0, hence $O \leq 2$, and $M \leq 0.5$. It follows:

$$M(\text{polygon}) \leq 1.5,$$

where $M = 1.5$ only if the polygon is a square in the upright position.

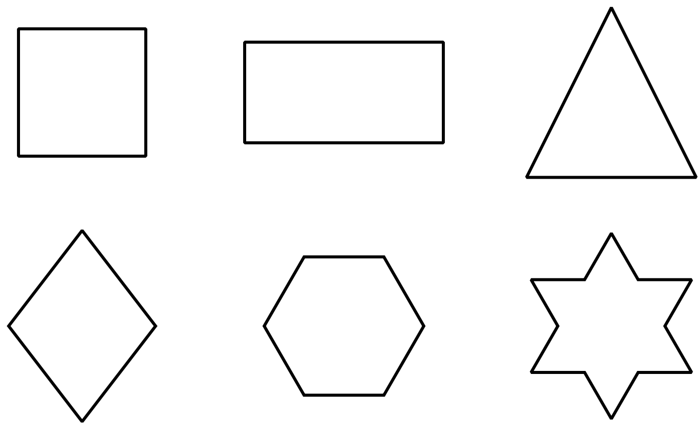


Figure 4: Aesthetic measure from left to right: $M = 1.5$, $M = 1.25$, $M = 7/6$, $M = 1$, $M = 1$, $M = 1$.

We have calculated that the square in the upright position is the polygon with the greatest value of the aesthetic measure. A natural question is whether Birkhoff interprets this result as saying that the square is *the most beautiful polygon*.

Since the measure is a function from objects to the rational numbers, the values of the measure are ordered as rational numbers. Birkhoff does not explicitly say that the underlying order on the rational numbers should be interpreted as the ordering of the values of the measure. However, in some parts of the book he does refer to the ordering of the rational numbers and relates it to the ordering of the measure (we choose a quotation concerning polygons):

> "… the square in horizontal position has the highest rating of all polygonal forms …" [Bir33, p. 25]

However, he also cautions the reader that

> "It follows then as a 'theorem' that the square with horizontal sides with $M = 1.50$ is the best of all possible polygonal forms. Obviously such mathematical treatment upon the basis of the theory becomes a mere game if carried too far." [Bir33, p. 47]

3. The impact of Birkhoff's work

In the time of publication of Birkhoff's book, mathematical community had already been familiar with Birkhoff's search for a simple mathematical formula for the aesthetic measure, see for instance [Bir29], [Bir31]. Birkhoff was also invited to give lectures on this topic, for instance on the International Congress of Mathematicians in Bologna, 1928. Additionally, he was invited to introduce his theory in a four-volume compendium *The*

World of Mathematics, edited by J. Newman (1933, with further editions in 1956 and 2003). The text was published with editor's preface in section XXI, *A Mathematical Theory of Art*, with the title *Mathematics of Aesthetics* [Bir56]. Today, we can find almost 850 citations of Birkhoff's *Aesthetic Measure*, not least because of the interdisciplinary character of various researchers interested in his work. In this paper, we only give references to main sources which refer to Birkhoff's work.

Initially (1930s and 1940s), we mostly find reactions by psychologists. In their papers they published results of empirical studies and compared them with Birkhoff's theory. After an application of the measure, they usually look for a correlation between the empirical results and theoretical predictions; the empirical results were primarily obtained using volunteers who were asked to give their aesthetical preference for the presented objects. Objects used in the studies were of various forms, including polygons. The results were usually negative in the sense that the theory of the aesthetic measure was not corroborated empirically. For more information, see for instance [Dav36], [Eys41], [Wil39], [BCP37] or [Gra55].

A survey of results in psychology aimed at either verifying or refuting Birkhoff's theory can be found in *A Review of Research on Aesthetic Measure* by Harold J. McWhinnie from 1968 [McW68]. McWhinnie states that psychological studies were not interested in the aesthetic judgement concerning an object, but rather in the aesthetic preference of the observer. This seems to be at odds with Birkhoff's prerogative to use only properties of the object itself (formal associations), and disregard subjective preference (connotative associations). Studies based on aesthetic preference run counter to Birkhoff's ideas, and it seems that the negative results of such studies must be re-evaluated carefully. Psychological studies into the ways of quantifying beauty or aesthetics inspired by Birkhoff's aesthetic measure continue into the present, see for instance [BL85] and [PSS13].

Moving away from psychology to theory of information, Birkhoff's work has had an impact in this field as applied in the context of the theory of aesthetics. There are two influential works in this field from 1960s: a book by Abraham Moles *Information Theory and Esthetic Perception*, first published in 1958 and most cited in the English translation from 1966 [Mol66], and Max Bense's work summarised in a four-volume work *Aesthetica* published between 1954 and 1960, and in the book *Aesthetica. Einführung in die neue Aesthetik* from 1965.

Abraham Moles (1920–1992) was a French information theorist who studied the relationship between theory of information and aesthetics, with the focus on the relationship between theory of perception and psychology. Inspired by Shannon's mathematical theory of communication, he developed Birkhoff's ideas into a theory of information and aesthetic perception. He redefined Birkhoff's aesthetic measure from the ratio of order O and complexity C to their multiplication $O \times C$. The original notion of order O in Birkhoff's measure takes in Moles' work the form of low entropy, perceived as redundancy and predictability. High entropy is equated with complexity C, perceived as unpredictability and non-compressibility.

Max Bense (1910–1990) was a German philosopher who is best known today for his work in philosophy of science, aesthetics and semiotics. Building on work of Birkhoff and Shannon, he focused on physical concepts, with the aim of creating rational aesthetics stripped of its subjective component. In addition to the emphasis on strictly scientific methods, Birkhoff and Bense share the definition of the aesthetic measure as the ratio of order and complexity.

Works by Bense and Moles, inspired by Birkhoff, led to the research in the field of Information Aesthetics (sometimes also Informational Aesthetics). Information Aesthetics looks for theoretical foundations of aesthetics, viewed from the point of information and its amount and quality contained in an object. As in Birkhoff's work, the goal is to judge an object by itself, without a subjective component. Research in Information Aesthetics is still in progress, see for instance [Gre05], [McC05], [RFS08], [Gal12].

In general, the quantitative study of aesthetic perceptions (see for instance [AC98]) is often based not only on Shannon's theory of information, but also on Kolmogorov complexity, see for instance [RFS07]. In contrast to the psychological works from 1930s and 1940s, Birkhoff's work is now treated with more flexibility and with less stress on the exact wording of the original text. The emphasis is on the main idea that the aesthetic quality of an object is connected to the notions of order and complexity. The research into the computational methods in aesthetics continues, with many papers referring to Birkhoff's work giving various functions and methods for computation of the aesthetic measure of an object, see for instance [HE10], [Pen98]. Nowadays, an important application of Birkhoff's work lies in the computer-aided design, as in [Cle11].

The field of theoretical aesthetics has had some Czech researchers as well: let us mention T. Staudek [Sta99], [Sta02], R. Kozubík [Koz09], J. Nešetřil [Neš94], [Neš05], [AN01], [BN07] or previous works of the author [DN09], [DN10].

4. A critical discussion of Birkhoff's work

Birkhoff's work in aesthetics has received approximately 850 citations and continues to influence various fields of research. Researchers inspired by Birkhoff often develop certain ideas of his, but seldom provide comprehensive critical analysis of his work. However, on a closer reading one can identify several types of objections appearing in the literature; we summarise these below.

As we stated above, after the publication of Birkhoff's book, psychologists often set up experiments designed to corroborate or refute Birkhoff's thesis. They were mostly interested in the aesthetic preference of volunteers (for instance [Dav36], [BCP37], [Wil39], [Eys41], [Gra55], [McW68], [BL85], and [PSS13]). None of the results verified relevance of the aesthetic measure, and the results did not correlate. Other critical studies, not specifically designed to test Birkhoff's measure, raised the more general objection of the aesthetic preference of volunteers not educated in arts being substantially different from the preference of experts or artists.

Still other papers discussed the appropriateness of the mathematical formula itself (for a summary, see for instance [Gal10]); papers with this focus have appeared more recently, originating from information aesthetics or computational aesthetics. They often raised the objection that the formula $M = f\left(\frac{O}{C}\right)$ seems to measure more an aesthetic efficiency than the level of aesthetic quality, prefers symmetry over beauty, penalises complexity, and views order and complexity as opposing notions. This type of critical analysis has been put forward for the original Birkhoff's measure, its modifications, and theories and models developed by other researchers (such as A. Moles, M. Bense, D. E. Berlyn, and others).

Additionally, researchers often raise doubts regarding the validity of the choice of parameters participating in the application of the formula on the given class of objects, in particular regarding the computation of order, with the underlying problem of distinguishing formal and connotative associations (see for instance [Yor]). The very distinction of formal and connotative associations, strongly defended by Birkhoff, has been questioned.

We can divide critical reactions into three basic groups which address different components of Birkhoff's theory and its subsequent development: *what* quantity is measured by the aesthetic measure (the question of legitimacy), *how* is it measured (the question of method), and *what* is the relevance of the measure for aesthetics (the question of relevance).

Let us start with the first question. Birkhoff gives the following answer to the question what is measured in formal aesthetics:

> "… the basic formal side of art within the perview of the simple mathematical formula …"[Bir33, p. viii]

Thus Birkhoff's intention was to objectify aesthetics, to identify formal rules which would be universally transferable and applicable. After 80 years, Birkhoff's original intention is still present in attempts to model the aesthetic judgment by formal or computational methods (for instance computer-aided design). These attempts need to solve several problems, technically most challenging being the issue of obtaining valid computer-generated data (this relates to the search for suitable and empirically corroborated formulas and algorithms) and of subsequent analysis of the data, not to mention the underlying question of the possibility of real-world use of results obtained by such formal methods.

If we accept the legitimacy of an aesthetic measure, we can move to the second question, i.e. how to define such a measure. Birkhoff refers to older philosophical works to defend his idea of defining the aesthetic measure as the ratio of order and complexity, claiming that these works relate the aesthetic quality of an object to the harmony of the object, to its unity in variety. It seems that the generalised formula $M = f\,(O, C)$, treating M as a function of two variables, for the computation of the aesthetic measure is generally accepted since many researchers do see connection between the aesthetic quality of an object and the two complementary notions: first notion being usually described as order, structure, redundancy of information, repetition, symmetry, fractal pattern, or low entropy; the second notion being equated to complexity, unpredictability, high entropy, and non-compressibility. In addition to the formula, one can also analyse its application, in particular the choice of parameters which take part in the computation. However, the distinction of formal and connotative associations has always been viewed as problematic. Birkhoff, as a mathematician, solved this problem for polygons by defining as formal those associations which are determined by movements of the polygons on the plane, with the effect on the numerical values regarding vertical and rotational symmetries, or on the value of HV which captures the properties of the polygon regarding its position in a grid. One may ask whether such geometrical properties are culturally transferable and whether, and to what extent, they are determined by education and cultural background. A mathematician, for instance, may find symmetry, rotation or other forms of patterns more appealing and interesting, and sees them as providing good *reward for attention of perception* (in Birkhoff's words). However, experiments show that artists, on the other hand, tend to prefer higher entropy and they are less interested in objects with too many regular features. It is therefore unclear to what extent we can define in a strictly scientific way formal associations which are supposed to belong to an object itself. The issue of distinguishing and defining formal and connotative associations has been widely discussed not only in papers in computational aesthetics, but also in psychological papers. Related questions have been studied recently in connection to the information contained in an aesthetic feeling and information processed in an aesthetic judgement.

Finally, we address the last question of relevance of the measure to aesthetics. In the papers we have discussed, these questions appear only indirectly, but from the philosophical point of view they present a major difficulty. It is not only the question of reducing aesthetics or aesthetic feeling to formal properties of an object, but the more general problem whether an aesthetic judgement is composed from individual properties of an object. Birkhoff's measure postulates that an *aesthetic information* can decomposed into components which can be evaluated separately. Research in the field of computational aesthetics proceeds similarly. Hence, these methods discard not only the subject, but also the interaction between subject and object and all other external circumstances. It seems that the resulting aesthetic judgment or aesthetic preference may be a function of the aesthetic measure itself: $A = g\,(M, \dots)$, where A is the final aesthetic judgment. It remains to be seen whether A can be described in more detail and to what extent we can consider M as an input parameter of an aesthetic judgement.

In conclusion, let us emphasise that in evaluating problematic issues in Birkhoff's work we should not forget his initial assumptions – we do not claim that his theory, or its generalisations, should solve the whole question of the *aesthetic judgement*, i.e. we do not search for TOE – theory of everything in the physical sense. The purpose of the research is to gain at least a partial understanding of aesthetics using formal methods, even at the risk of inaccuracies and non-correlation with the subjective notion of aesthetic preference. The wide influence of Birkhoff's work suggests that his methods do provide such an insight.

Acknowledgement

This research has been supported by grant MŠMT ČR ERC-CZ LL 1201.

References

[AC98] Tsion Avital and Gerald C. Cupchik. Perceiving hierarchical structures in nonrepresentational paintings. *Empirical Studies of the Arts*, 16(1): 59–70, 1998.

[AN01] Jan Adamec and Jaroslav Nešetřil. Towards an aesthetic invariant for graph drawing. In *Graph Drawing'01*, volume 2265 of *LNCS*, pages 287–296. Springer-Verlag, 2001.

[Bar03] John D. Barrow. Art and sciences—Les liaisons dangereuses? In J. Casti and A. Karlgvist, editors, *Art and Complexity*, pages 1–20. JAI Press, 2003.

[BCP37] John Gilbert Beebe-Center and Carroll Cornelius Pratt. A test of Birkhoff's aesthetic measure. *The Journal of General Psychology*, 17(2): 339–353, 1937.

[Bir29] George D. Birkhoff. Quelques éléments mathématiques de l'art. *Atti del Congresso Internazionale dei Matematici, Bologna*, I, 1929.

[Bir31] George D. Birkhoff. Une théorie quantitative de l'esthétique. *Bulletin de la Société française de Philosophia*, I, 1931.

[Bir33] George D. Birkhoff. *Aesthetic Measure*. Harvard University Press, 1933.

[Bir56] George D. Birkhoff. Mathematics of aesthetics. In James Newman, editor, *The World of Mathematics*, volume 4, pages 2185–2195. Simon and Schuster, New York, 1956.

[BL85] Frans Boselie and Emanuel Leeuwenberg. Birkhoff revisited: Beauty as a function of effect and means. *The American journal of psychology*, pages 1–39, 1985.

[BN07] Martin Bálek and Jaroslav Nešetřil. Towards mathematical aesthetics. In Michele Emmer, editor, *Mathematics and Culture V*, pages 225–236. Springer-Verlag New York, LLC, 2007.

[Cle11] Paul Charles Cleveland. Colour and dynamic symmetry. In Jennifer DeFore, editor, *2011 Joint Conference of the 41st Annual National PCA-ACA and 32nd Annual SW/Texas PCA-ACA*. SWTX PCA/ACA, 2011.

[Dav36] R. C. Davis. An evaluation and test of Birkhoff's Aesthetic Measure formula. *The Journal of General Psychology*, (15:2): 231–240, 1936.

[DN09] Veronika Douchová and Jaroslav Nešetřil. Fragment of aesthetic processing. In Jana Šafránková and Jiří Pavlů, editors, *WDS'09 Proceedings of Contributed Papers: Part I – Mathematics and Computer Science*, pages 7–10. Matfyzpress, 2009.

[DN10] Veronika Douchová and Jaroslav Nešetřil. S. Zeki and neuroaesthetics – A possible framework to objectify aesthetics. *Miscellanea Logica*, (IX), 2010.

[Eys41] H. J. Eysenck. The empirical determination of an aesthetic formula. *Psychological Review*, pages 83–92, 1941.

[Gal10] Philip Galanter. Complexity, neuroaesthetics, and computational aesthetic evaluation. In *13th International Conference on Generative Art (GA2010)*, 2010.

[Gal12] Philip Galanter. Computational aesthetic evaluation: Past and future. In Jon McCormack and Mark d'Inverno, editors, *Computers and Creativity*, pages 255–293. Springer-Verlag Berlin Heidelberg, 2012.

[Gra55] G. W. Granger. Aesthetic measure applied to color harmony: An experimental test. *The Journal of General Psychology*, 52(2): 205–212, 1955.

[Gre05] Gary Greenfield. On the origins of the term computational aesthetics. In *Proceedings of the First Eurographics conference on Computational Aesthetics in Graphics, Visualization and Imaging*, pages 9–12. Eurographics Association, 2005.

[HE10] E. den Heijer and A. E. Eiben. Comparing aesthetic measures for evolutionary art. In Cecilia Di Chio, Anthony Brabazon, Gianni A. Di Caro, Marc Ebner, Muddassar Farooq, Andreas Fink, Jörn Grahl, Gary Greenfield, Penousal Machado, Michael O'Neill, Ernesto Tarantino, and Neil Urquhart, editors, *Applications of Evolutionary Computing*, volume 6025 of *Lecture Notes in Computer Science*, pages 311–320. Springer-Verlag, Berlin Heidelberg, 2010. EvoApplications 2010: EvoCOMNET, EvoENVIRONMENT, EvoFIN, EvoMUSART, and EvoTRANSLOG, Istanbul, Turkey, April 7–9, 2010, Proceedings, Part II.

[Koz09] Roman Kozubík. *Posuzování strojírenských výrobků z hlediska průmyslového designu*. PhD thesis, VUT Brno, 2009.

[McC05] Jon McCormack. Open problems in evolutionay music and art. In F. Rothlauf, J. Branke, S. Cagnoni, D. W. Corne, R. Drechsler, Y. Jin, P. Machado, E. Marchiori, J. Romero, G. D. Smith, and G. Squillero, editors, *Applications of Evolutionary Computing*, volume 3449 of *Lecture Notes in Computer Science*, pages 428–436, Berlin Heidelberg, 2005. Springer-Verlag. Evoworkshops: EvoBIO, EvoCOMNET, EvoHot, EvoIASP, EvoMUSART, and EvoSTOC.

[McW68] Harold J. McWhinnie. A review of research on aesthetic measure. *Acta Psychologica*, 28: 363–375, 1968.

[Mol66] Abraham A. Moles. *Information Theory and Esthetic Perception*. University of Illinois Press, Urbana and London, 1966.

[Neš94] Jaroslav Nešetřil. An essay on art and mathematics. *From the Logical Point of View*, 2(2): 93, 1994.

[Neš05] Jaroslav Nešetřil. Aesthetics for computers or How to measure a harmony. In Michele Emmer, editor, *Visual Mind II*, pages 35–58. MIT Press, 2005.

[Pen98] Penousal Machado and Amílcar Cardoso. Computing aesthetics. In F. Moreira de Oliveira, editor, *Advances in Artificial Intelligence*, volume 1515 of *Lecture Notes in Computer Science*, pages 219–228, Berlin Heidelberg, 1998. Springer-Verlag.

[PSS13] Stephen E. Palmer, Karen B. Schloss, and Jonathan Sammartino. Visual aesthetics and human preference. *Annual review of psychology*, 64: 77–107, 2013.

[RFS07] Jaume Rigau, Miquel Feixas, and Mateu Sbert. Conceptualizing Birkhoff's aesthetic measure using Shannon entropy and Kolmogorov complexity. In *Proceedings of the Third Eurographics Conference on Computational Aesthetics in Graphics, Visualization and Imaging*, Computational Aesthetics '07, pages 105–112, Aire-la-Ville, Switzerland, 2007. Eurographics Association.

[RFS08] Jaume Rigau, Miquel Feixas, and Mateu Sbert. Informational aesthetics measures. *Computer Graphics and Applications, IEEE*, 28: 24–34, 2008.

[Sta99] Tomáš Staudek. On Birkhoff's aesthetic measure of vases. *Faculty of Information, Masaryk University Report Series (FIMU-RS-99-06)*, 1999.

[Sta02] Tomáš Staudek. *Exact Aesthetics: Object and Scene to Message*. PhD thesis, MU Brno, 2002.

[Wil39] D. J. Wilson. An experimental investigation of Birkhoff's aesthetic measure. *The Journal of Abnormal and Social Psychology*, 34(3): 390–394, 1939.

[Yor] William W. York. Cognitive science, aesthetics, and the development of taste. http://csjarchive.cogsci.rpi.edu/Proceedings/2013/papers/0303/paper0303.pdf.

2015 ACTA UNIVERSITATIS CAROLINAE PAG. 55–72
PHILOSOPHICA ET HISTORICA 1 / MISCELLANEA LOGICA X

LARGE CARDINALS AND THEIR EFFECT ON THE CONTINUUM FUNCTION ON REGULAR CARDINALS

RADEK HONZÍK*
Dept. of Logic, Faculty of Arts, Charles University in Prague
* E-mail: radek.honzik@ff.cuni.cz

ABSTRACT
In this survey paper, we will summarise some of the more and less known results on the generalisation of the Easton theorem in the context of large cardinals. In particular, we will consider inaccessible, Mahlo, weakly compact, Ramsey, measurable, strong, Woodin, and supercompact cardinals. The paper concludes with a result from the opposite end of the spectrum: namely, how to kill all large cardinals in the universe.
Keywords: continuum function, large cardinals

1. Introduction

One of the questions which stood at the birth of set theory as a mathematical discipline concerns the size of real numbers $\mathbb{R}$. Cantor conjectured that there is no subset of the real line whose cardinality is strictly between the size of the set of natural numbers and the size of all real numbers. With the axiom of choice, this is equivalent to saying that the size of $\mathbb{R}$ is the least possible in the aleph hierarchy:

$$\text{The Continuum Hypothesis, CH: } |\mathbb{R}| = 2^{\aleph_0} = \aleph_1.$$

Hilbert included this problem in 1900 as the number one question on his influential list of 23 open problems in mathematics.

It is well known now that CH is independent of the axioms of ZFC.[1] First Gödel showed in 1930s that CH is consistent with ZFC (using the constructible universe L), and then in 1960s Cohen showed that ¬CH is consistent with ZFC (with forcing). Regarding Cohen's result, one naturally inquires how much CH can fail in Cohen's model; it is a witness to the remarkable utility of the method of forcing that virtually the same proof gives the greatest possible variety of results: in principle,

(*) if κ is any cardinal with uncountable cofinality, then $2^{\aleph_0} = \kappa$ is consistent.

There is a small issue how to express (*) properly. We can view (*) as a statement about consistency of a theory, in which case κ should either be a parameter or should be definable in ZFC,[2] or (*) can be taken as a statement about pairs of models of ZFC. It is the latter approach which is more useful and general:

1 If ZFC is consistent, which we will assume throughout the paper.
2 E.g. $\aleph_{\omega+3}$, $\aleph_{\omega_1}$, or the first weakly inaccessible cardinal (if there is one).

Theorem 1.1 (Cohen, Solovay). *Let κ be a cardinal with uncountable cofinality in V, and assume $\kappa^\omega = \kappa$ in V. Then there is a cofinality-preserving extension $V[G]$ of V such that $V[G] \models (2^{\aleph_0} = \kappa)$.*

Easton [9] generalised this result to all regular cardinals. Let us write Card for the class of cardinals and Reg for the regular cardinals. Let F be a function from Reg to Card. Assume further that F satisfies for all κ, λ in Reg:

(i) $\kappa < \lambda \to F(\kappa) \leq F(\lambda)$.

(ii) $\kappa < \mathrm{cf}(F(\kappa))$.

Let us call such an F an *Easton function*. We say that an Easton function F is realised in a model M if $\mathrm{Reg} = \mathrm{Reg}^M$ and for all regular κ in M, $F(\kappa) = 2^\kappa$.

Theorem 1.2 (Easton). *Assume V satisfies* GCH *and let F be an Easton function definable over V. Then there is a definable cofinality-preserving proper-class forcing notion P such that if G is P-generic, then in $V[G]$,*

$$(\forall \kappa \in \mathrm{Reg})(2^\kappa = F(\kappa)),$$

i.e. F is realised in $V[G]$.

There are more general statements of Easton's theorem which remove the restriction of definability of F. Such generalisations usually require additional assumptions above ZFC: one can for instance start with an inaccessible cardinal κ and GCH below κ, and set $M = H(\kappa)$. Then M is a transitive model of ZFC + GCH. An Easton function F for M is now an element of $H(\kappa^+)$, and may not be definable over M. Easton's theorem now generalizes as follows:[3]

Theorem 1.3 (Easton, generalised version). *Let κ be an inaccessible cardinal and denote $M = V_\kappa$, and let F be an Easton function defined on regular cardinals $\alpha < \kappa$. Assume further that* GCH *holds below κ. Then there is a cofinality-preserving forcing notion of size κ such that if G is P-generic over V, then in $M[G]$,*[4]

$$(\forall \alpha \in \mathrm{Reg})(2^\alpha = F(\alpha)),$$

i.e. F is realised in $M[G]$.

Easton's theorem solves the problem of the possible behaviours of the continuum function on regular cardinals in models of ZFC in full generality. Mathematicians briefly conjectured that Easton's theorem could be extended to all cardinals – including the singular cardinals. However, Silver soon proved the following limiting theorem which shows that ZFC controls the continuum function more tightly on singular cardinals:

Theorem 1.4 (Silver). *Let κ be a singular strong limit cardinal of uncountable cofinality. If the set $\{\mu < \kappa \mid 2^\mu = \mu^+\}$ is stationary in κ, then $2^\kappa = \kappa^+$.*

SCH, Singular Cardinal Hypothesis, is a weakening of GCH and says that if κ is a singular strong limit cardinal, then $2^\kappa = \kappa^+$.[5] Silver's theorem claims that the validity of SCH at a

[3] In the rest of the paper, we will not distinguish between these two versions of Easton's theorem.

[4] $M[G]$ is now viewed as a constructible closure of M relative to an additional predicate G.

[5] There are more versions of SCH, some of them formulated for all singular cardinals.

singular strong limit κ is determined by the continuum function on singular strong limit cardinals below κ: in particular, if SCH holds below κ, it must hold at κ.

Surprisingly, similar restrictions hold for regular cardinals which exhibit some combinatorial properties associated to large cardinals (see for instance Lemma 1.17), provided we wish to preserve these properties while realising an Easton function. Acknowledging the importance of large cardinals in current set theory, do we have a satisfactory analogue of Easton's theorem for extensions of ZFC with large cardinals? We will study this question in the following sections, defining all necessary notions as we proceed.

Remark 1.5. Due to lack of space, we completely disregard in this paper other possible, and interesting, generalisations of the Easton theorem: (i) one can for instance study the effect of former large cardinals on the continuum function (e.g. a regular κ with the tree property), (ii) consider other cardinal invariants in addition to 2^κ (see [6]), and finally (iii) consider the continuum function on all cardinals. Regarding (iii), as we mentioned above, there are some analogies between the restrictions valid for singular strong limit cardinals of uncountable cofinality (Silver's theorem) and restrictions valid for e.g. measurable cardinals (Lemma 1.17). However, there are also subtle differences which prevent an easy transfer of the respective results. In particular, in Lemma 1.17, the set A is required to be in a normal measure, not just stationary, as in Silver's theorem.

1.1 Large cardinals

We review some of the more basic large cardinals. The cardinals are listed in the increasing order of strength: inaccessible < Mahlo < weakly compact < Ramsey < measurable < strong < strongly compact, supercompact.[6] Slightly apart, there is the Woodin cardinal which in terms of consistency strength is roughly on the level of a strong cardinal, while it may not be even weakly compact (it is always Mahlo, though).

Proofs of results stated below as facts or mentioned in passing can be found in [14] or [15].

Definition 1.6. *Let κ be a regular uncountable cardinal. We say that κ is inaccessible if $2^\lambda < \kappa$ for every $\lambda < \kappa$ (this property is called being a strong-limit cardinal).*

Note that if GCH holds, then κ is inaccessible if and only if κ is regular and limit cardinal.

A slight strengthening of inaccessibility is Mahloness.

Definition 1.7. *We say that an inaccessible cardinal κ is Mahlo if the set of regular cardinals below κ is stationary.*

Lemma 1.8. *If κ is Mahlo, then the set of inaccessible cardinals is stationary below κ.*

6 < in this case means both the consistency strength and the provable implication: thus for instance a Mahlo cardinal has a strictly larger consistency strength than an inaccessible cardinal, and every Mahlo cardinal is an inaccessible cardinal. It is conjectured that the supercompact and strongly compact cardinals have the same consistency strength; in terms of the implication, a supercompact cardinal is always strongly compact, but not conversely.

Proof. Recall the definition of the function $\beth$: $\beth_0 = \aleph_0$, $\beth_{\alpha+1} = 2^{\beth_\alpha}$, and $\beth_\gamma = \sup\{\beth_\delta \mid \delta < \gamma\}$ for γ limit. By the inaccessibility of κ, the set

$$A = \{\mu < \kappa \mid \beth_\mu = \mu\}$$

is a closed unbounded set of limit cardinals.

We want to show that every closed unbounded set $C \subseteq \kappa$ contains an inaccessible cardinals. By the previous paragraph, $C \cap A$ is a closed unbounded set. By Mahloness, the set of regular cardinals is stationary, and therefore it must meet $C \cap A$. Hence, there is $\mu \in C \cap A$ which is a regular cardinal. By the definition of A, μ is strong-limit and therefore inaccessible. □

As the next large cardinal after Mahlo cardinal, we review the weakly compact cardinal. There are many equivalent definitions of weak-compactness. The one we give first is formulated in terms of trees:

Definition 1.9. *An inaccessible κ is weakly compact if every κ-tree*[7] *has a cofinal branch.*

Note that this definition points to the original motivation for this cardinal: recall that König's theorem (that every ω-tree has a cofinal branch) can be used to prove the compactness theorem for the first-order logic. For a stronger logic which allows infinite quantifications, conjunctions and disjunctions, the similar proof goes through if κ is weakly compact (because the generalisation of König's theorem holds for κ).

An equivalent definition directly postulates a reflection property. We say that a formula φ in the language of set theory with two types of variables is Π^1_1 if it contains at the beginning a block of universal quantifiers over subsets of the target domain (second-order variables), followed by the usual first-order quantification over elements of the target domain (first-order variables). Thus $\forall X \exists x (x \in X)$ is true over a structure $(M, \in)$ if for every $A \subseteq M$ there is some $a \in M$ such that $a \in A$. We write $\varphi(R)$ to indicate that φ contains a free second-order variable R (we call R a parameter).

Fact 1.10. *The following are equivalent:*

(i) *κ is weakly compact.*

(ii) *κ is inaccessible and for every $R \subseteq V_\kappa$ and every Π^1_1 formula $\varphi(R)$,*

$$\text{(1.1)} \quad \textit{If } (V_\kappa, \in, R) \models \varphi(R), \textit{ then} \\ (\exists \alpha < \kappa, \alpha \textit{ inaccessible})(V_\alpha, \in, V_\alpha \cap R) \models \varphi(R \cap V_\alpha).$$

Note that we can also view $(V_\kappa, \in, R)$ as a first-order structure with a predicate R; if κ is Mahlo, then the usual Löwenheim-Skolem theorem implies (ii) of Fact 2.10 for all *first-order formulas* $\varphi(R)$. However, to get (ii) for Π^1_1 formulas, the usual Löwenheim-Skolem theorem no longer suffices because now it should be applied over the first-order structure $(V_{\kappa+1}, V_\kappa, \in, R)$, and there is no guarantee it will yield a substructure of the form $(V_{\alpha+1}, V_\alpha, \in, R \cap V_\alpha)$.

Lemma 1.11. *Suppose κ is weakly-compact and x is a cofinal subset of κ. If $x \cap \alpha \in L$ for every $\alpha < \kappa$, then $x \in L$.*

[7] A tree of height κ whose levels have size $< \kappa$.

Proof. Sketch. Suppose $x \notin L$. Then there is a Π^1_1 sentence φ such that $(V_\kappa, \in, x) \models \varphi(x)$ if and only if x is not in L. φ contains a second-order quantifier which ranges over all subsets of κ which code levels of L of size at most κ and says that in no such level of L, x is constructed.

By weak-compactness, φ is reflected to some $\alpha < \kappa$, which gives $(V_\alpha, \in, x \cap \alpha) \models \varphi(x \cap \alpha)$, which is equivalent to $x \cap \alpha \notin L$, contradicting our initial assumption. □

A weakly compact cardinal has another useful characterisation by means of colourings. If κ is a regular cardinal, then a colouring of two-element subsets of κ by two colours is a function $f : [\kappa]^2 \to 2$. We say that $H \subseteq \kappa$ is homogeneous for f if $f \restriction [H]^2$ has size 1.

Fact 1.12. *The following are equivalent for an inaccessible κ:*
(i) *κ is weakly compact.*
(ii) *Every colouring $f : [\kappa]^2 \to 2$ has a homogeneous set of size κ.*

By considering more complex colourings, we can obtain a stronger large cardinal notion:

Definition 1.13. *Let $\kappa > \omega$ be an inaccessible cardinal. We say that κ is a Ramsey cardinal if every colouring $f : [\kappa]^{<\omega} \to 2$ has a homogeneous set of size κ.*

By definition, every Ramsey cardinal is weakly compact. Moreover, one can show that if there is a Ramsey cardinal, then $V \neq L$. Thus being Ramsey is a substantial strengthening of weak compactness which is compatible with L.

Another cardinal we will mention is the measurable cardinal:

Definition 1.14. *We say that an inaccessible κ is measurable if there is a non-principal*[8] *κ-complete*[9] *ultrafilter U on κ. U is often called a measure.*

Fact 1.15. *The following are equivalent:*
(i) *κ is measurable.*
(ii) *There is an elementary embedding*[10] *$j : V \to M$, where M is a transitive class, $j \restriction \kappa = id$ and $j(\kappa) > \kappa$. (We call κ the critical point of j.)*

If (ii) holds, we can find an embedding $j' : V \to M'$ which in addition satisfies that $\kappa^+ = (\kappa^+)^{M'}$, $H(\kappa^+)^{M'} = H(\kappa^+)$, and M' is closed under κ-sequences in V.

We should say something about proving (i)→(ii) because it features the important concept of an *ultrapower*. Assume that U is a measure on κ. For $f, g : \kappa \to V$ define $f \equiv g \Leftrightarrow \{\xi < \kappa \mid f(\xi) = g(\xi)\} \in U$. For every $f : \kappa \to V$, define

$$[f] = \{g \mid g : \kappa \to V \;\&\; f \equiv g\}.$$

We would like to say that the collection of all $[f]$'s forms a partition of the class of all functions $\kappa \to V$; this is the case, but it presents the problem that this collection is a class

8 For no $\alpha < \kappa$, $\{\alpha\} \in U$.

9 If X_i, $i < \mu < \kappa$ are in U, then $\bigcap_{i<\mu} X_i$ is in U.

10 j is a proper class; thus we should view this definition as taking place in GB set theory, or more technically – but preferably – as a statement expressible in ZFC because the relevant part of j which we need, $j \restriction H(\kappa^+)$, is a set.

of classes, making it an illegal object in set theory. We will therefore identify $[f]$ with the sets in $[f]$ of minimal rank. Using this identification, denote

$$\mathrm{Ult}(V,U) = \{[f] \mid f : \kappa \to V\}.$$

Define the interpretation of $\in$ on elements of $\mathrm{Ult}(V,U)$: $[f] \in [g] \Leftrightarrow \{\xi < \kappa \mid f(\xi) \in g(\xi)\} \in U$.

Theorem 1.16 (Łos)**.** *For every φ and $f_1, \dots, f_n$:*

$$\mathrm{Ult}(V,U) \models \varphi[[f_1], \dots, [f_n]] \Leftrightarrow \{\xi < \kappa \mid \varphi(f_1(\xi), \dots, f_n(\xi))\} \in U. \tag{1.2}$$

By ω_1-completeness of the measure U, the relation $\in$ on $\mathrm{Ult}(V,U)$ is well-founded, and one can therefore collapse the structure $(\mathrm{Ult}(V,U), \in)$, obtaining a transitive proper class model. The proof (i)→(ii) is finished by taking for j the composition of the canonical ultrapower embedding $j' : V \to \mathrm{Ult}(V,U)$ defined by

$$j'(x) = [c_x],$$

where $c_x : \kappa \to \{x\}$, and of the collapsing isomorphism π:

$$j = \pi \circ j'.$$

We say that U is *normal* if

$$[id] = \kappa. \tag{1.3}$$

One can show that if κ is measurable, there always exists a normal measure. Property (1.3) is useful for computing information about ultrapowers; see Lemma 1.17 for an application.

Lemma 1.17. *Assume κ is measurable and let U be a normal measure. If $A = \{\alpha < \kappa \mid 2^\alpha = \alpha^+\}$ is in U, then $2^\kappa = \kappa^+$.*

Proof. Let $\mathrm{Ult}(V,U)$ be the transitive collapse of the ultrapower, as discussed above after Fact 1.15. By Łos theorem, $A \in U$ implies

$$\mathrm{Ult}(V,U) \models 2^{[id]} = [id]^+$$

which is by normality the same as

$$\mathrm{Ult}(V,U) \models 2^\kappa = \kappa^+.$$

As stated in Fact 1.15, $\kappa^+ = (\kappa^+)^{\mathrm{Ult}(V,U)}$, and $H(\kappa^+) = (H(\kappa^+))^{\mathrm{Ult}(V,U)}$. This implies $\mathcal{P}(\kappa) = (\mathcal{P}(\kappa))^{\mathrm{Ult}(V,U)}$. Therefore any bijection $g \in \mathrm{Ult}(V,U)$ between $(\kappa^+)^{\mathrm{Ult}(V,U)}$ and $\mathcal{P}(\kappa)^{\mathrm{Ult}(V,U)}$ is a bijection between κ^+ and $\mathcal{P}(\kappa)$ in V, proving $2^\kappa = \kappa^+$. □

A useful set which belongs to any normal measure is

$$I = \{\alpha < \kappa \mid \alpha \text{ is inaccessible}\}.$$

I is stationary and co-stationary, i.e. $(\kappa \setminus I)$ is also stationary. I is in every normal measure because $\kappa = [id]$ is inaccessible in $\mathrm{Ult}(V,U)$; by Łos theorem this implies that I is in U. By a similar argument one can show that if C is club in κ, then $C \in U$: in the ultrapower, $\kappa \in j(C)$, which by Łos theorem is equal to $C \in U$. Note that Lemma 1.17 depends on ultrafilter U in the following sense. Denote

$$A = \{\alpha < \kappa \mid 2^\alpha = \alpha^+\}.$$

To argue that $2^\kappa = \kappa^+$ it suffices to find *at least one* normal measure U which contains A. As we discussed, if A is club or a set of inaccessibles, then all normal measures contain A. However, if A is just stationary, then it is not the case in general that there is some normal measure U which contains A. In fact, it is consistent that A is stationary and $2^\kappa > \kappa^+$ (see Lemma 2.14).

By strengthening the properties of the elementary embedding in the definition of a measurable cardinal, we get the notion of a strong cardinal. For more motivation and properties of strong cardinals, see Section 2.3.

Definition 1.18. *We say that an inaccessible cardinal κ is $H(\lambda)$-strong, $\kappa < \lambda$ regular, if there is an elementary embedding $j : V \to M$ with critical point κ, $j(\kappa) > \lambda$, $H(\lambda) \subseteq M$, and M is closed under κ-sequences in V.*

We say that κ is strong if it is $H(\lambda)$-strong for every regular $\lambda > \kappa$.

By definition, being measurable is the same as being $H(\kappa^+)$-strong.

By strengthening the closure properties of the target model M in the definition of a strong cardinal, we obtain an even stronger notion of a supercompact cardinal (see Definition 1.21). However, we first define the notion of a strongly compact cardinal, using a generalisation of the ultrafilter definition of a measurable cardinal. In preparation for the definition, let us define the following: Let $\kappa \le \lambda$ be cardinals, κ regular, and set

$$P_\kappa\lambda = \{x \subseteq \lambda \mid |x| < \kappa\}.$$

For $x \in P_\kappa\lambda$, define

$$\hat{x} = \{y \in P_\kappa\lambda \mid x \subseteq y\}.$$

Finally, define

$$F(\kappa, \lambda) = \{X \subseteq P_\kappa\lambda \mid (\exists x \in P_\kappa\lambda)\ \hat{x} \subseteq X\}.$$

We call $F(\kappa, \lambda)$ *a fine filter on $P_\kappa\lambda$.*

Lemma 1.19. *$F = F(\kappa, \lambda)$ is a κ-complete filter.*

Proof. Follows because for $\{x_i \mid i < \mu < \kappa\} \subseteq P_\kappa\lambda$,

$$\bigcap_{i<\mu} \hat{x}_i = \widehat{\bigcup_{i<\mu} x_i}.$$

□

Definition 1.20. *Assume $\kappa \le \lambda$ are cardinals, κ inaccessible. We call κ λ-strongly compact if the fine filter $F(\kappa, \lambda)$ can be extended into a κ-complete ultrafilter on $P_\kappa\lambda$. We call κ strongly compact if it is λ-strongly compact for all $\lambda \ge \kappa$.*

Strongly compact cardinals are much stronger than measurable cardinals (regarding consistency strength); however, by a result of Magidor from 70s the first measurable cardinal can be strongly compact.

By demanding that there is a κ-complete ultrafilter extending $F(\kappa, \lambda)$ which is also *normal* (we will not define this notion, see [14], p. 374), we get the notion of a *supercompact* cardinal. A characterisation of supercompactness by means of elementary embeddings is very convenient:

Definition 1.21. *Let κ be an inaccessible cardinal, and let $\lambda \geq \kappa$ be a cardinal. We say that κ is λ-supercompact if there is an elementary embedding $j : V \to M$ with critical point κ such that $j(\kappa) > \lambda$ and ${}^{\lambda}M \subseteq M$. A cardinal κ is supercompact if it is λ-supercompact for every $\lambda \geq \kappa$.*

Finally, we define a large cardinal notion due to Woodin which he used in the analysis of the Axiom of Determinacy.

Definition 1.22. *Let $\delta > \omega$ be an inaccessible cardinal. We say that δ is a Woodin cardinal if for every function $f : \delta \to \delta$ there is a $\kappa < \delta$ with $f''\kappa \subseteq \kappa$ and there is $j : V \to M$ with critical point κ such that $V_{j(f)(\kappa)} \subseteq M$.*

A Woodin cardinal is always Mahlo, but may not be weakly compact. Its consistency strength is quite high (by definition, there are many cardinals on the level of a $H(\mu)$-strong cardinal, for some μ, below a Woodin cardinal).

2. The continuum function with large cardinals

Assume κ is a large cardinal in V which satisfies GCH and F is an Easton function. Can we find a generic extension of V which realises F and preserves the largeness of a fixed large cardinal κ? Clearly, a necessary condition on F is that it should keep κ strong limit. We can formulate this property globally for the class of large cardinals we wish to preserve. Let Γ be a class of regular cardinals. We say that *F respects* Γ if

$$(\forall \kappa \in \Gamma)(\forall \mu \in \mathrm{Reg} \cap \kappa)(F(\mu) < \kappa). \tag{2.4}$$

In anticipation of the generalisation of Easton's theorem to large cardinals, we can tentatively formulate two distinguishing criteria, (R^-), (R) and (L^-), (L), which help to characterise large cardinals according to their sensitivity to the manipulation with the continuum function:

- (R^-) Cardinals without obvious reflection properties relevant to the continuum function[11] such as inaccessible, Mahlo, weakly compact, and Woodin or Ramsey cardinals.
- (R) Cardinals with reflection properties relevant to the continuum function, such as measurable cardinals.

A typical effect of reflection of measurable cardinals regarding the continuum function is captured in Lemma 1.17 above.

Remark 2.1. The notion of reflection is often used in a broad sense (for instance, Fact 2.10 provides a notion of reflection for Π^1_1-formulas). In (R^-) and (R), we use it in a very restricted sense: κ has a reflection property (relevant to the continuum function) if 2^κ depends on the values of 2^α, for $\alpha < \mu$.

A different classification is based on what is called *fresh subsets*:

[11] Of course, only after we generalise Easton's theorem to these cardinals we know for certain that they have no "hidden" reflection properties.

Definition 2.2. *Let $M \subseteq N$ be two transitive models of set theory with the same ordinals. Let κ be a cardinal in N. We say that $x \subseteq \kappa$ is fresh if $x \in N \setminus M$ and for all $\alpha < \kappa$, $x \cap \alpha \in M$.*

For instance Cohen forcing[12] at a regular cardinal κ adds a fresh subset κ.

(L^-) Cardinals which are not obviously influenced by fresh subsets such as inaccessible and Mahlo cardinals.

(L) Cardinals which may be destroyed by adding fresh subsets such as weakly compact cardinals, or measurable cardinals.

Lemma 1.11 identifies this restriction for weakly compact cardinals.

As we will see, the first distinction (R^-) and (R) is relevant for the possible patterns of the continuum function which can be realised, while the second distinction (L^-) and (L) is relevant for the appropriate method of forcing.

The following forcing, defined in Easton [9], will be refereed to as the *product-style Easton forcing*, and denoted it P_F^{product}.

Definition 2.3. *Let F be an Easton function. For all regular cardinals α, define Q_α to be the Cohen forcing* $\mathrm{Add}(\alpha, F(\alpha))$. *Define*

$$P_F^{\text{product}} = \prod_{\alpha \in \text{Reg}}^{\text{Easton}} Q_\alpha,$$

where the upper index indicates that the forcing has the "Easton support": for every inaccessible α and any condition $p \in P_F^{\text{product}}$, $\mathrm{dom}(p) \cap \alpha$ is bounded in α.

Note in particular that if there are no inaccessible cardinals, then the forcing is just a full-support product of Cohen forcings. It is relatively straightforward to compute that if GCH holds in the ground model, then P_F^{product} preserves all cofinalities and forces $2^\alpha = F(\alpha)$, for all regular α.

As we indicated above in the paragraph after the definition of a fresh subset, a product-style forcing will not be good enough for preservation of large cardinals with reflection as in Lemma 1.11. In anticipation of a solution to this problem, we define a variant of Easton forcing which appeared already in [17]. For this definition, let us first define some notions. If F is an Easton function, let C_F be the closed unbounded class of limit cardinals which are the closure points of F: i.e.

$$C_F = \{\alpha \mid \alpha \text{ limit cardinal } \ \& \ (\forall \beta \in \alpha \cap \text{Reg})(F(\beta) < \alpha)\}.$$

Notice that if F respects Γ, see (2.4), then Γ is included in C_F.

Definition 2.4. *Let F be an Easton function. By reverse Easton forcing we mean the forcing P_F defined as follows. For every pair (α, β) of successive elements of C_F, let us write*

$$Q_{\alpha,\beta} = \prod_{\gamma \in [\alpha,\beta) \cap \text{Reg}}^{\text{Easton}} \mathrm{Add}(\gamma, F(\gamma)).$$

12 If α is a limit ordinal and $\beta > 0$ is an ordinal, we define the Cohen forcing at α for adding β-many subsets of α, $\mathrm{Add}(\alpha, \beta)$, as the collection of all functions from $\alpha \times \beta$ to 2 with domain of size $< |\alpha|$. Ordering is by reverse inclusion. Of course, $\mathrm{Add}(\alpha, \beta)$ is equivalent to $\mathrm{Add}(|\alpha|, |\beta|)$, but the more general notation is often useful.

P_F is the iteration $(\langle P_\alpha \mid \alpha \in \mathrm{Ord}\rangle, \langle \dot{Q}_\alpha \mid \alpha \in \mathrm{Ord}\rangle)$ with Easton support such that $\dot{Q}_\alpha$ is the canonical name for the trivial forcing whenever α is not in C_F. If α is in C_F, let $\dot{Q}_\alpha$ be a name for the forcing $Q_{\alpha,\beta}$, where β is the successor of α in C_F.

2.1 Inaccessible and Mahlo cardinals

Let F be an Easton function respecting inaccessible cardinals, i.e. respecting $\Gamma = \{\alpha \mid \alpha$ is inaccessible$\}$ according to (2.4). To generalise Easton's theorem to F, it suffices to check that the forcing P_F^{product} preserves cofinalities of all $\kappa \in \Gamma$. As we indicated after Definition 2.3, cofinalities are preserved for all cardinals if V satisfies GCH, which yields the following theorem:

Theorem 2.5. *Let V satisfy* GCH *and let F be an Easton function respecting inaccessible cardinals. Let A_0 be the class of all inaccessible cardinals. Then in any generic extension $V[G]$ by P_F^{product}, the set of inaccessible cardinals coincides with A_0.*

One can formulate a version of the theorem for Mahlo cardinals.

Theorem 2.6. *Let V satisfy* GCH *and let F be an Easton function respecting Mahlo cardinals. Let A_0 be the class of all Mahlo cardinals. Then in any generic extension $V[G]$ by P_F^{product}, the set of Mahlo cardinals coincides with A_0.*

Proof. Let G be P_F^{product}-generic and let κ be a Mahlo cardinal in V. Since the set of inaccessible cardinals I is stationary in κ in V, $C_F \cap I$ is also stationary. It follows by Theorem 2.5 that all inaccessible $\alpha \in C_F \cap I$, and also κ, remain inaccessible in $V[G]$. To finish the argument, it suffices to check that $C_F \cap I$ is still stationary in $V[G]$. Factor P_F^{product} into $P_0 \times P_1$ such that P_1 is κ-closed and P_0 is κ-cc:[13] P_1 preserves stationary subsets of κ because it is κ-closed; as P_1 forces that P_0 is κ-cc, P_0 preserves stationary subsets over V^{P_1}. Thus $P = P_0 \times P_1$ preserves stationary subsets of κ, and in particular stationarity of $C_F \cap I$. □

Actually, the reverse Easton iteration P_F achieves the same result here. The point is that for every Mahlo κ, one can show that $(P_F)_\kappa$, the restriction of P_F to κ, is κ-cc, and the tail iteration is forced to be κ-closed.

Remark 2.7. We have argued that the relevant forcings do not kill inaccessible or Mahlo cardinals. To get the results above, we also need to argue that the forcings do not create new large cardinals. However, notice that P_F^{product} and P_F cannot create new inaccessible cardinals because these forcings preserve cofinalities, and therefore a non-inaccessible cardinal α in the ground model must remain non-inaccessible in the extension. Similarly, a non-stationary set of inaccessible cardinals cannot become stationary, and thus new Mahlo cardinals cannot be created.

---

13 P_0 is defined as P_F^{product}, but with the domain of the functions in the product limited to $\kappa \cap \mathrm{Reg}$; similarly, P_1 has the domain limited to $\mathrm{Reg} \setminus \kappa$.

It is easy to find an example where the product-style Easton forcing P_F^{product} destroys weak-compactness of some cardinal κ, over some well-chosen ground model such as L.

Lemma 2.8. *Assume that κ is weakly compact and let F be an Easton function. Then over L, P_F^{product} kills weak-compactness of κ.*

Proof. P_F^{product} factors at κ to $P_0 \times P_1 \times P_2$, where P_0 is P_F^{product} restricted to regular cardinals $< \kappa$, P_1 is the forcing $\mathrm{Add}(\kappa, F(\kappa))$, and P_2 is the restriction to regular cardinals $> \kappa$. We argue that P_1 kills the weak-compactness of κ, and neither P_0, nor P_2 can resurrect it.

The fact that P_1 kills weak-compactness of κ follows from Lemma 1.11 (because it adds many fresh subsets of κ over L). It follows that after forcing with P_1, there exists a κ-tree without a cofinal branch. Since P_2 cannot add a branch to a κ-tree because it is κ^+-distributive over V^{P_1}, κ is not weakly compact in $V^{P_1 \times P_2}$.

Finally notice that P_0 is κ-Knaster in $V^{P_1 \times P_2}$ by the usual Δ-lemma argument (and the fact that κ is Mahlo here). Using the fact that a κ-Knaster forcing cannot add a branch to a κ-tree (see [1]), we conclude that in V^{P_F} there exist a κ-tree without a cofinal branch, contradicting weak-compactness of κ. □

In order to formulate Theorem 2.6 for weakly compact cardinals, we need to introduce a very universal technique for verification of preservation of large cardinals. This technique uses the characterisation of many large cardinals by means of suitable elementary embeddings between transitive sets or classes. In order to show that a certain large cardinal κ remains large in a generic extension, it suffices to check that the original embedding from V "lifts" to an embedding in the generic extension (this is in general easier than to verify that there exists *an* elementary embedding in the extension). The following lemma of Silver is the key ingredient:

Lemma 2.9 (Silver)**.** *Assume M and N are transitive models of* ZFC*, $P \in M$ is a forcing notion, and $j : M \to N$ is an elementary embedding. Let G be P-generic over M, and let H be $j(P)$-generic over N. Then the following are equivalent:*

(i) $(\forall p \in G)(j(p) \in H)$.

(ii) There exists an elementary embedding $j^+ : M[G] \to N[H]$ such that $j^+(G) = H$ and $j^+ \restriction M = j$.

We say that j^+ is a *lifting of j*. If j has some nice property (like being an extender embedding), the lifting j^+ will often have it as well. More details about these concepts can be found in [5].

This is a useful characterisation of weakly compact cardinals (proof can be found in [5]):

Fact 2.10. *Let κ be an inaccessible cardinal. The following are equivalent.*

(i) κ is weakly compact.

(ii) For every transitive set M with $|M| = \kappa$, $\kappa \in M$, and ${}^{<\kappa}M \subseteq M$, there is an elementary embedding $j : M \to N$ where N is transitive, $|N| = \kappa$, ${}^{<\kappa}N \subseteq N$, and the critical point of j is κ.

Now, using the characterisation of weak-compactness by elementary embeddings, one can show:

Theorem 2.11. *Let V satisfy* GCH *and let F be an Easton function respecting weakly compact cardinals. Let A_0 be the class of all weakly compact cardinals. Then in any generic extension $V[G]$ by P_F, the set of weakly compact cardinals coincides with A_0.*

Proof. The proof has two parts: Part 1 proves that all weakly compact cardinals in V remain weakly compact in $V[G]$. In Part 2, which corresponds to Remark 2.7 above, we argue that the forcing does not create new weakly compact cardinals.

Part 1.

The proof is given in [3]; we will only briefly identify the main points, assuming some familiarity with lifting arguments. The proof is similar to an argument in [5], section 16 – when one uses the forcing P_F – with one extra twist to be resolved: assuming κ is weakly compact, in [5], one forces below κ with a reverse Easton forcing which at every inaccessible $\alpha < \kappa$ forces with $\mathrm{Add}(\alpha, 1)$. At κ, one can force with $\mathrm{Add}(\kappa, \mu)$ for an arbitrary μ because any κ-tree which supposedly does not have a cofinal branch is captured by a subforcing of $\mathrm{Add}(\kappa, \mu)$ which is isomorphic to $\mathrm{Add}(\kappa, 1)$; thus the preparation below κ matches the forcing at κ, making it possible to use a standard lifting argument with a master condition. In Theorem 2.11, the preparation below κ is determined by F so it may not be possible to force just with $\mathrm{Add}(\alpha, 1)$ at every inaccessible $\alpha < \kappa$; in particular if $j : M \to N$ is an embedding ensured by Fact 2.10, we need to force with $\mathrm{Add}(\kappa, j(F)(\kappa))$ on the N-side; this introduces a mismatch between the forcings at κ between M and N: $\mathrm{Add}(\kappa, 1)$ vs. $\mathrm{Add}(\kappa, j(F)(\kappa))$. In order to lift to $j(\mathrm{Add}(\kappa, 1))$, one therefore needs to make sure to have on the N-side available the generic filter g for $\mathrm{Add}(\kappa, 1)$. In [3], the solution is to include g on the first coordinate of the generic filter for $\mathrm{Add}(\kappa, j(F)(\kappa))$. The rest of the argument is standard.

Part 2.

The situation of a weakly compact cardinal is a bit more complicated than in the analogous Remark 2.7. By Kunen's construction [16], it is possible to turn a weakly compact cardinal κ into a Mahlo non-weakly compact cardinal by forcing a κ-Souslin tree, and resurrect its weak-compactness by forcing with the Souslin tree added earlier. However, it is easy to check that this kind of anomaly will not occur with our forcing.

Let κ be a Mahlo non-weakly compact cardinal in V which is a closure point of F; it follows there is a κ-tree T in V which has no cofinal branch in V. Denote $R = (P_F)_\kappa$, and $\dot{Q} = \mathrm{Add}(\kappa, F(\kappa))$; it suffices to check that $R * \dot{Q}$ cannot add a branch through T. R cannot add a cofinal branch because it is κ-Knaster. Over V^R, $\dot{Q}$ cannot add a branch to T because it is κ-closed (if p in $\dot{Q}$ forced that $\dot{B}$ is a cofinal branch through T, then one could find a decreasing sequence of conditions $\langle p_i \mid i < \kappa\rangle$, $p_0 = p$ and a $\leq_T$-increasing sequence $\langle b_i \mid i < \kappa\rangle$ such that $p_i \Vdash b_i \in T$; the sequence $\langle b_i \mid i < \kappa\rangle$ would be a cofinal branch in T in V^R). □

Thus for inaccessible, Mahlo and weakly compact cardinals, there are no restrictions on the Easton functions F which can be realised, except that these cardinals must be closure points of F. In particular, the reflection property identified in Lemma 1.11 did have an effect on the technique (P_F over P_F^{product}), but not on the result. In the next section, we learn that the case of measurable cardinals is far more complicated.

It follows from Lemma 1.17 that to preserve measurable cardinals, we must expect that the full generalisation along the lines of Theorems 2.6 and 2.11 cannot be achieved. There are two easy properties to notice regarding restrictions on the continuum function by measurable cardinals:

(a) There is an obvious asymmetry in the sense that Lemma 1.17 prohibits 2^κ "jumping up" with respect to the values 2^α for $\alpha < \kappa$, while "jumping down" is perfectly possible. See Lemma 2.12.

(b) The restrictions which a measurable cardinal κ puts on the continuum function also depend on the normal measures which exist on κ (and not only on the fact that κ is measurable). See Lemma 2.14.

We first deal with (a).

Lemma 2.12. *Assume that κ is measurable and $2^\kappa > \kappa^+$. Let P be the collapsing forcing $\mathrm{Col}(\kappa^+, 2^\kappa)$ which collapses 2^κ to κ^+ by functions of size at most κ. Then in V^P, κ is still measurable and $2^\kappa = \kappa^+$.*

Proof. By κ^+-closure of P, every measure on κ in V remains a measure in V^P because P does not add new subsets of κ to measure (nor new κ-sequences of such sets). Notice that the result did not assume that $\{\alpha < \kappa \mid 2^\alpha = \alpha^+\}$ is big in the sense of some measure on κ. $\square$

We will deal with (b) after we define the notion of an $H(\lambda)$-strong cardinal.

Apart from the easy observations (a) and (b), we in addition have:

(c) The consistency strength of a measurable cardinal κ with $2^\kappa > \kappa^+$ is $o(\kappa) = \kappa^{++}$, see [12]. Thus to play with the continuum function and preserve measurability of cardinals, one typically needs to assume that these cardinals are larger than measurable in the ground model.

In view of (c), we now define a suitable strengthening of measurability.

Definition 2.13. *We say that an inaccessible cardinal κ is $H(\lambda)$-strong, $\kappa < \lambda$ regular, if:*

(i) There is an elementary embedding $j : V \to M$ with critical point κ, $j(\kappa) > \lambda$, such that

(ii) $H(\lambda) \subseteq M$, and M is closed under κ-sequences in V.

We say that κ is strong if it is $H(\lambda)$-strong for every regular $\lambda > \kappa$.

We note that with GCH, κ being $H(\kappa^{++})$-strong is equivalent to having Mitchell order of $\kappa^{++}+1$, a slight strengthening of the assumption identified by [12] as optimal for obtaining the failure of GCH at a measurable cardinal.

As promised, we now deal with the property (b).

Lemma 2.14. *Assume κ is $H(\kappa^{++})$-strong and that* GCH *holds in the universe. Denote $I = \{\alpha < \kappa \mid \alpha \text{ is inaccessible}\}$.*

Then there exist a stationary subset X of I, distinct normal measures U, W on κ, and a forcing notion P such that:

(i) $X \in W$ and $(I \setminus X) \in U$,

(ii) Assume G is P-generic. In $V[G]$, κ is measurable, $2^\kappa = \kappa^{++}$, $2^\alpha = \alpha^+$ for all $\alpha \in X$, and $2^\alpha = \alpha^{++}$ for all $\alpha \in (I \setminus X)$.

In particular, W cannot be extended into a normal measure in V^P.

Proof. Let U, W be two distinct normal measures on κ in V. We know that I is in both U and W; therefore for some $A \subseteq I$, $A \in U$ and $B = (I \setminus A) \in W$ (if U and W agreed on all subsets of I, they would agree on all subsets of κ).

Let $j : V \to M$ be an elementary embedding witnessing $H(\kappa^{++})$-strength of κ. Without loss of generality assume that $\kappa \in j(A)$ (and $\kappa \notin j(B)$). We define P so that $B = X$ is as desired.

Let Q be the standard reverse Easton iteration which at every $\alpha \in (I \setminus X)$ forces with $\mathrm{Add}(\alpha, \alpha^{++})$. By an argument involving "surgery", see [5], one can show that there is a forcing $\dot{R}$ such that denoting $P = Q * \dot{R}$, in V^P all cofinalities are preserved, $2^\kappa = \kappa^{++}$, and κ is measurable. Moreover, in V^P, X and $(I \setminus X)$ are stationary subsets of inaccessible cardinals such that $2^\alpha = \alpha^{++}$ for $\alpha \in (I \setminus X)$, and $2^\alpha = \alpha^+$ for $\alpha \in X$.

It follows that U extends to a normal measure in V^P, while by Lemma 1.17, W (and any other normal measure containing X) cannot extend into a normal measure in V^P. □

This lemma should be understood as follows: while W prohibits certain values of the continuum function in V because $X \in W$ (e.g. implies $2^\kappa = \kappa^+$), this restriction is not persistent to larger models: in V^P, $2^\kappa = \kappa^{++}$ is possible even though X is still a stationary subset composed of inaccessible cardinals. This scenario is made possible by the assumption that there is at least one embedding j in V for which the set $I \setminus X$ is big – using this j we can kill all normal measures which contain X, while ensuring that some normal measures still exist in V^P.

These consideration lead to the following theorem (see [10]):

Theorem 2.15. *Let F be an Easton function respecting every κ which is $H(F(\kappa))$-strong, and assume* GCH *holds in the universe. There is a cofinality-preserving iteration P which realises F such that whenever G is P-generic over V, we have:*

Whenever in V, κ is $H(F(\kappa))$-strong and there is $j : V \to M$ witnessing $H(F(\kappa))$-strength of κ such that

$$j(F)(\kappa) \geq F(\kappa), \tag{2.5}$$

then κ remains measurable in $V[G]$.

The proof is beyond the scope of this paper, but let us at least comment on the method of proof. As we mentioned in Lemma 2.14, the manipulation of 2^κ with κ measurable using the Cohen forcing and Woodin's "surgery argument" requires us to use an extra forcing denoted $\dot{R}$ in the proof of Lemma 2.14. It seems quite hard to incorporate this extra forcing at every relevant stage into a global result along the lines of Theorem 2.15. Instead, to prove Theorem 2.15 we use the generalised product-style α-Sacks forcing $\mathrm{Sacks}(\alpha, \beta)$, for an inaccessible α and an ordinal $\beta > 0$ (see [10] for details): P is a reverse Easton iteration defined similarly as in Definition 2.4 with $\mathrm{Add}(\gamma, F(\gamma))$ replaced by $\mathrm{Sacks}(\gamma, F(\gamma))$ whenever γ is an inaccessible closure point of F.[14] The use of Sacks forcing has the advantage that to lift an embedding, no extra forcing $\dot{R}$ is required.

[14] Since one mixes the α-Sacks forcing with the α^+-Cohen forcing (and other Cohen forcings – but only the stage α^+ requires an argument), one needs to argue that they work well together: in particular, one can show (see [10]) that $\mathrm{Sacks}(\alpha, F(\alpha))$ forces that $\mathrm{Add}(\alpha^+, F(\alpha^+))$ is still α^+-distributive. In fact, this is true for any α^+-closed forcing in place of $\mathrm{Add}(\alpha^+, F(\alpha^+))$.

The property (2.5) is essential for lifting the embedding at κ, and captures the degree of reflection which F needs to satisfy for preservation of measurability of κ. The proof is relatively straightforward when $F(\kappa)$ is regular, but is more involved when $F(\kappa)$ is a singular cardinal (the most difficult case is when $F(\kappa)$ has cofinality $> \kappa^+$ in V and is singular in V, but is regular in M, where $j : V \to M$ is an embedding witnessing (2.5)).

Note that the apparent lack of uniformity in the statement of the theorem (the condition (2.5)) is unavoidable as illustrated in Lemma 2.14. Also note that the use of $H(F(\kappa))$-strong cardinals is almost optimal, as mentioned above in the discussion of property (c).

Remark 2.16. We have not checked whether every measurable cardinal κ in $V[G]$ is measurable also in V, obtaining an analogue of Theorems 2.6 and 2.11, but we consider it likely.

2.4 Ramsey, Woodin and supercompact cardinals

We shall more briefly review results for some other large cardinals, most notably Ramsey, Woodin and supercompact.

A Ramsey cardinal, see Definition 1.13, is large enough to imply $V \neq L$, but it may not be measurable (and its consistency strength is less than measurability). In the classification following (2.4), Ramsey cardinals are in (R^-) and (L). We will see below in Theorem 2.17 that indeed Ramsey cardinals have no reflection properties relevant for the continuum function.

In terms of consistency, Woodin cardinals (see Definition 1.22) are much stronger than measurable cardinals, being in principle inaccessible limits of $H(\lambda)$-strong cardinals introduced above (for certain λ's). However, a Woodin cardinal may not even be weakly compact (while it is a Mahlo cardinal). Its classification is still (R^-) and (L), as will be apparent from Theorem 2.18.

The following theorem appears in [3] as Theorem 4.5:

Theorem 2.17. *Let V satisfy* GCH *and let F be an Easton function respecting Ramsey cardinals. Let A_0 be the class of all Ramsey cardinals. Then in any generic extension $V[G]$ by P_F, F is realised and the set of Ramsey cardinals contains A_0.*

We should note that the proof of Theorem 2.17 utilizes a characterisation of Ramseyness by means of elementary embeddings, to apply an appropriately tailored lifting argument.

The following theorem appears in [2] as Theorem 1:

Theorem 2.18. *Let V satisfy* GCH *and let F be an Easton function respecting Woodin cardinals. Let A_0 be the class of all Woodin cardinals. Then in any generic extension $V[G]$ by a certain cofinality preserving forcing P, F is realised and the set of Woodin cardinals contains A_0.*

The forcing P in the statement of the theorem contains the α-Sacks forcing at the critical stages (regular closure points α of F), similarly as we discussed below Theorem 2.15. The key lemma for the preservation of Woodiness is Lemma 14 in [2].

We now turn to supercompact cardinals. The first generalisation of the Easton theorem for large cardinals actually appeared for the supercompact cardinals, see [17]. Since supercompact cardinals have reflection properties, it is not possible to realise every F and preserve supercompact cardinals; Menas identified a property of F which is sufficient for preservation of supercompact cardinals:

Definition 2.19. *An Easton function F is said to be locally definable if the following condition holds:*

There is a sentence ψ and a formula $\varphi(x, y)$ with two free variables such that ψ is true in V and for all cardinals γ, if $H(\gamma) \vDash \psi$, then $F[\gamma] \subseteq \gamma$ and

$$\forall \alpha, \beta \in \gamma (F(\alpha) = \beta \Leftrightarrow H(\gamma) \vDash \varphi(\alpha, \beta)). \tag{2.6}$$

The following is a theorem in section 18 of [17]:

Theorem 2.20. *Let V satisfy* GCH *and let F be a locally definable Easton function respecting supercompact cardinals. Let A_0 be the class of all supercompact cardinals. Then in any generic extension $V[G]$ by the forcing P_F of Definition 2.4, F is realised and the set of supercompact cardinals contains A_0.*

The theorem is proved using a "master condition" argument[15] applied to the forcing, which makes it possible to use Cohen forcing at closure points of F; compare with the discussion below Theorem 2.15. Theorem 2.20 was generalised also for the strong cardinals (see Definition 2.13); see [10, Theorem 3.17].

Theorem 2.21. *Let V satisfy* GCH *and let F be a locally definable Easton function respecting strong cardinals. Let A_0 be the class of all strong cardinals. Then in any generic extension $V[G]$ by a certain cofinality-preserving forcing P, F is realised and the set of strong cardinals contains A_0.*

The forcing P contains the α-Sacks forcing at regular closure points α of F.

Let us conclude this section by remarking that there are results similar to these theorems which are formulated for a λ-supercompact cardinal κ which is also $H(\nu)$-strong for some $\lambda < \nu$; see [11, 4].

2.5 Open questions

Considering the variety of large cardinal concepts, it is no surprise that many of them have not been studied from the point of their compatibility with patterns of the continuum function. For instance the following cardinals have not been studied:[16]

- While strong compactness is close to supercompactness in the consistency strength, the dropping of normality of the witnessing ultrafilter makes it less well-behaved. In particular, a characterisation by means of an elementary embedding only gives the following (compare with Definition 1.21):

---

15 Roughly, in order to lift an embedding j between transitive classes M and N, the pointwise image of a P-generic filter g, $j''g$, is an element of N, generating a suitable $j(P)$-generic filter h over N containing $j''g$. $j''g$ is called a master condition. In crucial situations, $j''g$ is usually too big to be in N; a typical case where $j''g$ is in N is when j is a supercompact embedding. There is no master condition for arguments starting with $H(F(\kappa))$-strong cardinals.

16 To our knowledge, no complete generalisation of the Easton theorem has been formulated yet.

Definition 2.22. *Let κ be an inaccessible cardinal and $\lambda > \kappa$ a cardinal. κ is λ-strongly compact if there is an elementary embedding $j : V \to M$ with critical point κ such that $j(\kappa) > \lambda$ and for any $X \subseteq M$ with $|X| \leq \lambda$ there is $Y \in M$, $Y \supseteq X$, such that $M \models |Y| < j(\kappa)$.*

These weaker properties of the embedding suggest a different lifting method – instead of lifting an embedding, one can lift directly the ultrafilter (as in [13], albeit in a different context).

- We say that κ is a *Shelah cardinal* if for every $f : \kappa \to \kappa$ there is $j : V \to M$ with critical point κ such that $V_{j(f)(\kappa)} \subseteq M$. Very little has been published about this cardinal with respect to the continuum function.
- Rank-to-rank embeddings (the hypotheses I3–I0). A partial result appeared in [8].

There are many other cardinals which can be studied, so our list is far from complete.

3. In the converse direction

In the whole paper, we studied the question of preserving large cardinals while manipulating the continuum function. As a curiosity, we show in this section that by manipulating the continuum function, it is possible to wipe out all large cardinals.

Theorem 3.1. *Let $M = V_\kappa$, where κ is an inaccessible cardinal. Suppose $I = \{\alpha < \kappa \mid \alpha \text{ is inaccessible}\}$ is a non-stationary subset of κ. Then there is a forcing P of size κ, definable in $H(\kappa^+)$ such that in $M[G]$, there are no inaccessible cardinals, for any P-generic filter G over V.*

Proof. Let C be a club disjoint from I, and let $\langle c_i \mid i < \kappa\rangle$ be the increasing enumeration of C. Define P to be a product of Cohen forcings with Easton support as follows: define $Q_i = \mathrm{Add}(c_i^+, c_{i+1})$ for $0 \leq i < \kappa$, and $Q_{-1} = \mathrm{Add}(\aleph_0, c_0)$. Set

$$P = \prod_{-1<i<\kappa}^{\text{Easton}} Q_i,$$

where the superscript "Easton" denotes the Easton support.

Let G be a P-generic filter over V. By definition of P, if $\mu < \kappa$ is a limit cardinal closed under the continuum function in $V[G]$, then $\mu \in C$. Since $C \cap I = \emptyset$, it implies that in $V[G]$ there are no inaccessible cardinals below κ.

Finally, since κ is still inaccessible in $V[G]$, $M[G]$ is a transitive model of set theory without inaccessible cardinals as desired. □

Note that if M satisfies GCH, then the forcing P preserves cofinalities.

To destroy all inaccessible cardinals in M, it suffices to find a forcing which forces a club disjoint from inaccessible cardinals. The idea comes from [7].

Theorem 3.2. *Let M be as above. There is a forcing P of size κ which does not change $V_\kappa = M$ such that in $V[G]$ there is a club $C \subseteq \kappa$ disjoint from I, the set of inaccessibles below κ.*

Proof. Let conditions be functions from ordinals $\alpha < \kappa$ to 2 such that if $\beta < \kappa$ is inaccessible, then $\{\gamma \in \mathrm{dom}(p) \cap \beta \mid p(\gamma) = 1\}$ is bounded in β.

The forcing is κ-distributive because it is κ-strategically closed. So V_κ is not changed, and consequently all cardinals $\leq \kappa$ are preserved. Moreover since κ is inaccessible, P has size κ, so all cardinals are preserved.

Clearly, if G is P-generic over V, then

$$A = \lim\{\alpha < \kappa \mid (\exists p \in G)(p(\alpha) = 1)\}$$

is a club disjoint from I. □

Remark 3.3. Note that the same proof can be rephrased as turning a Mahlo cardinal into a non-Mahlo inaccessible cardinal.

Acknowledgments

The author was supported by GAČR project I 1921-N25.

References

[1] J. E. Baumgartner. Iterated forcing. In A. R. D. Mathias, editor, *Surveys in Set Theory*, pages 1–59. Cambridge University Press, 2010. London Math. Soc. Lecture Note Ser. 87.

[2] Brent Cody. Easton's theorem in the presence of Woodin cardinals. *Archive for Mathematical Logic*, 52(5–6): 569–591, 2013.

[3] Brent Cody and Victoria Gitman. Easton's theorem for Ramsey and strongly Ramsey cardinals. *Annals of Pure and Applied Logic*, 166: 934–952, 2015.

[4] Brent Cody and Menachem Magidor. On supercompactness and the continuum function. *Annals of Pure and Applied Logic*, 165(2): 620–630, 2014.

[5] James Cummings. Iterated forcing and elementary embeddings. In Matthew Foreman and Akihiro Kanamori, editors, *Handbook of Set Theory*, volume 2. Springer, 2010.

[6] James Cummings and Saharon Shelah. Cardinal invariants above the continuum. *Annals of Pure and Applied Logic*, 75: 251–268, 1995.

[7] René David. Some applications of Jensen's coding theorem. *Annals of Mathematical Logic*, 22: 177–196, 1982.

[8] Vincenzo Dimonte and Sy-David Friedman. Rank-into-rank hypotheses and the failure of GCH. *Archive for Mathematical Logic*, 53: 351–366, 2014.

[9] William B. Easton. Powers of regular cardinals. *Annals of Mathematical Logic*, 1: 139–178, 1970.

[10] Sy-David Friedman and Radek Honzik. Easton's theorem and large cardinals. *Annals of Pure and Applied Logic*, 154(3): 191–208, 2008.

[11] Sy-David Friedman and Radek Honzik. Supercompactness and failures of GCH. *Fundamenta Mathematicae*, 219(1): 15–36, 2012.

[12] Moti Gitik. On measurable cardinals violating the continuum hypothesis. *Annals of Pure and Applied Logic*, 63: 227–240, 1993.

[13] Moti Gitik and Saharon Shelah. On certain indestructibility of strong cardinals and a question of Hajnal. *Archive for Mathematical Logic*, 28: 35–42, 1989.

[14] Tomáš Jech. *Set Theory*. Springer, 2003.

[15] Akihiro Kanamori. *The Higher Infinite*. Springer, 2003.

[16] Kenneth Kunen. Saturated ideals. *The Journal of Symbolic Logic*, 43(1), 1978.

[17] Telis K. Menas. Consistency results concerning supercompactness. *Transactions of the American Mathematical Society*, 223: 61–91, 1976.

2015 ACTA UNIVERSITATIS CAROLINAE PAG. 73–91
PHILOSOPHICA ET HISTORICA 1 / MISCELLANEA LOGICA X

A SURVEY OF SPECIAL ARONSZAJN TREES

RADEK HONZÍK and ŠÁRKA STEJSKALOVÁ*
Dept. of Logic, Faculty of Arts, Charles University in Prague
* E-mail: sarka@logici.cz

ABSTRACT
The paper reviews special Aronszajn trees, both at ω_1 and κ^+ for an uncountable regular κ. It provides a comprehensive classification of the trees and discusses the existence of these trees under different set-theoretical assumptions. The paper provides details and proofs for many folklore results which circulate (often without a proper proof) in the literature.
Keywords: special Aronszajn trees

1. Introduction

A tree, which is now called Aronszajn, was first constructed by Nachman Aronszajn and the construction can be found in [Kur35]. The constructed tree was actually a special ω_1-Aronszajn tree. The definition of special Aronszajn tree has several equivalent variants and in the literature can be found many generalizations of the definition of a special Aronszajn tree. In this paper we focus on the question what are the relationships between them and provide a basic classification.

1.1 Preliminaries

In this section, we provide a review of basic definitions and facts relating to trees.

Definition 1.1. We say that $(T, <)$ is a *tree* if $(T, <)$ is a partial order such that for each $t \in T$, the set $\{s \in T | s < t\}$ is wellordered by $<$.

Definition 1.2. We say that $S \subseteq T$ is a *subtree* of $(T, <)$ in the induced ordering $<$ if $\forall s \in S\, \forall t \in T (t < s \rightarrow t \in S)$.

Definition 1.3. Let T be a tree
(i) If $t \in T$, then $ht(t, T) = ot(\{s \in T | s < t\})$ is *height of* t in T;
(ii) For each ordinal α, we define the *α-th level* of T as $T_\alpha = \{t \in T | ht(t) = \alpha\}$;
(iii) The *height* of T, $\mathrm{ht}(T)$, is the least α such that $T_\alpha = \emptyset$;
(iv) $T \restriction \alpha = \bigcup_{\beta<\alpha} T_\beta$ is a subtree of T of height α.

Definition 1.4. For a regular $\kappa \geq \omega$, T is called a *κ-tree* if T has height κ, and $|T_\alpha| < \kappa$ for each $\alpha < \kappa$.

Many κ-tree is isomorphic to a subtree of the full tree $({}^{<\kappa}\kappa, \subset)$. More precisely, this is the case whenever the κ-tree is normal. See the definition below.

Definition 1.5. A *normal κ-tree* is a tree T such that:

(i) $ht(T) = \kappa$;
(ii) $|T_\alpha| < \kappa$, for every $\alpha < \kappa$;
(iii) $|T_0| = 1$;
(iv) If $ht(s, T) = ht(t, T)$ is a limit ordinal, then $s = t$ if and only if $\{r \in T | r < s\} = \{r \in T | r < t\}$.

Note that the conditions (i) and (ii) ensure that a normal κ-tree is a κ-tree.

Fact 1.6. *Let κ be a regular cardinal. Then every normal κ-tree is isomorphic to a subtree T' of the full tree $({}^{<\kappa}\kappa, \subset)$.*

If we consider a successor cardinal κ^+ in the previous fact, then the levels of the κ^+-tree have size $\leq\kappa$. Hence we can strengthen the formulation of the previous fact for successor cardinals as follows: Every normal κ^+-tree is isomorphic to a subtree T' of the full tree $({}^{<\kappa^+}\kappa, \subset)$.

Definition 1.7. Let T be a tree. We say that B is a *branch* if it is a maximal chain in T.

Definition 1.8. Let κ be a regular cardinal. We say that a κ-tree T is a *κ-Aronszajn tree* if it has no branch of size κ. We denote the class of all Aronszajn trees at κ as $\mathbb{A}(\kappa)$.

By König's Lemma, no ω-Aronszajn trees exist. On the other hand, by result of Aronszajn, there exists an ω_1-Aronszajn tree. Moreover, if we assume GCH, then there exists a κ^+-Aronszajn tree for each regular cardinal κ, by a result of Specker [Spe49].

There are two common strengthenings of the notion of an Aronszajn tree. The first one leads to the notion of a special Aronszajn tree, to which we dedicate the next section. The second leads to the notion of a Suslin tree.

Definition 1.9. Let κ be a regular cardinal. We say that a κ-Aronszajn tree is Suslin, if it has no antichain[1] of size κ. We denote the class of all Suslin trees at κ as $\mathbb{S}(\kappa)$.

The notion of an ω_1-Suslin tree first appeared in connection with the Suslin problem of the characterization of the real line. Actually, in [Kur35] Kurepa showed that the original Suslin hypothesis (SH) can be formulated as the claim that there are no Suslin trees. For more details about Suslin hypothesis see [Jec03].

2. Special Aronszajn trees at ω_1

2.1 Classification

In this section, we classify different types of special Aronszajn trees at ω_1. Most of the notions are standard, but dispersed through diverse papers, so we think it is useful to provide a unified treatment here.

[1] $A \subset T$ is an antichain if for every $t, s \in A$, if $t \neq s$, then there is no $u \in T$ such that $u \geq t$ and $u \geq s$.

Definition 2.1. We say that an ω_1-Aronszajn tree T is *special* if T is a union of countable many antichains. We denote the class of all special Aronszajn trees at ω_1 as $\mathbb{A}^{sp}(\omega_1)$.

Definition 2.2. Let κ be a regular cardinal, T be a κ-Aronszajn tree and $\mathbb{P} = \langle P, <_{\mathbb{P}} \rangle$ be a partially ordered set. We say that T is $\mathbb{P}$-*embeddable* if there is a function $f : T \to P$ such that $s <_T t \to f(s) <_{\mathbb{P}} f(t)$. We denote the class of all $\mathbb{P}$-embeddable trees at κ as $\mathbb{T}(\mathbb{P})(\kappa)$.

Note that the previous definition can be generalized for arbitrary partially ordered set.

Definition 2.3. Let κ be a regular cardinal, $\mathbb{R} = \langle R, <_{\mathbb{R}} \rangle$ and $\mathbb{P} = \langle P, <_{\mathbb{P}} \rangle$ be a partially ordered sets. We say that $\mathbb{R}$ is $\mathbb{P}$-*embeddable* if there is a function $f : R \to P$ such that $s <_{\mathbb{R}} t \to f(s) <_{\mathbb{P}} f(t)$.

Fact 2.4. *The following are equivalent for an ω_1-Aronszajn tree T:*
(i) T is special;
(ii) There is $f : T \to \omega$ such that if s, t are comparable in T, then $f(s) \neq f(t)$;
(iii) T is $\mathbb{Q}$-embeddable, i.e. $T \in \mathbb{T}(\mathbb{Q})(\omega_1)$.

When we work with $\mathbb{Q}$-embeddable Aronszajn trees it is natural to consider also $\mathbb{R}$-embeddable Aronszajn trees and ask what is the connection between them. The following fact tells us how to characterise $\mathbb{R}$-embeddable Aronszajn trees using $\mathbb{Q}$-embeddable Aronszajn trees. It was first proved in [Bau70].

Fact 2.5. *Let T be an ω_1-tree. T is $\mathbb{R}$-embeddable if and only if $T^* = \bigcup_{\alpha<\omega_1} T_{\alpha+1}$ is $\mathbb{Q}$-embeddable.*

Now, we introduce the concept of an M-special Aronszajn tree.

Definition 2.6. We say that an ω_1-Aronszajn tree T is M-*special* if T is isomorphic to the subtree of $\{s \in {}^{<\omega_1}\omega | s \text{ is 1-1}\}$. We denote the class of all M-special ω_1-Aronszajn trees as $\mathbb{A}^{\text{M-}sp}(\omega_1)$.

We use the notation M-special to distinguish special Aronszajn trees defined by Mitchell in [Mit72] from now more used Definition 2.1. Note that Mitchell's definition includes just normal trees in contrast to Definition 2.1. In this sense the notion of a special tree is more general than M-special. However, if we consider just normal trees, then every special normal tree can be represented by an M-special tree. The converse may not hold in general, see Lemma 2.16.

Lemma 2.7. *If T is a normal special ω_1-Aronszajn tree, then T is* M-*special.*

Proof. Fix for every $\alpha < \omega_1$ a 1-1 function $g_\alpha : T_\alpha \to \omega$, and write $T = \bigcup_{n<\omega} A_n$, where A_n is an antichain for each $n < \omega$.

We define by induction on $\alpha < \omega_1$ a tree T' and an isomorphism $i : T \to T'$, where T' is a subtree of $\{s \in {}^{<\omega_1}(\omega \times \omega) | s \text{ is 1-1}\}$. The isomorphism i will be a union of partial isomorphisms $i_\alpha : T \restriction \alpha \to T' \restriction \alpha$.

Set $T'_0 = \{\emptyset\}$ and $i_1(r) = \emptyset$, where r is the root of T. As we assume that T is normal, i_1 is an isomorphism between $T \restriction 1$ and $T' \restriction 1$.

Suppose that we have constructed $i_\beta : T \restriction \beta \to T' \restriction \beta$ for each $\beta < \alpha$. First, if α is limit, set $i_\alpha = \bigcup_{\beta<\alpha} i_\beta$ and $T' \restriction \alpha = \bigcup_{\beta<\alpha} T' \restriction \beta$.

If $\alpha = \gamma + 1$ and γ is a successor, then we define i_α by extending i_γ setting for each $s \in T_\gamma$:

$$i_\alpha(s) = i_\gamma(t) \cup \{\langle \gamma, \langle g_\gamma(s), n\rangle\rangle\}, \tag{2.1}$$

where the node t is the immediate predecessor of s and $s \in A_n$. Let $T' \restriction \alpha = T' \restriction \gamma \cup T'_\gamma$, where $T'_\gamma = \{i_\alpha(s) | s \in T_\gamma\}$. It is clear that each function in T'_γ is 1-1 since each two comparable nodes must be in different antichains.

If $\alpha = \gamma + 1$ and γ is limit, then we define i_α by extending i_γ setting for each $s \in T_\gamma$:

$$i_\alpha(s) = \bigcup\{i_\gamma(t) | t < s\}. \tag{2.2}$$

By (iv) of Definition 1.5, i_α is 1-1 and clearly it is also an isomorphism. Let $T' \restriction \alpha = T' \restriction \gamma \cup T'_\gamma$, where $T'_\gamma = \{i_\alpha(s) | s \in T_\gamma\}$. Again it is obvious that each function in T'_γ is 1-1 since it is a union of 1-1 functions with gradually increasing domains.

At the end, set $T' = \bigcup_{\alpha<\omega_1} T' \restriction \alpha$ and $i = \bigcup_{\alpha<\omega_1} i_\alpha$. It is easy to see that the tree T' is isomorphic to a subtree of $\{s \in {}^{<\omega_1}\omega | s \text{ is 1-1}\}$ by any bijection between $\omega \times \omega$ and ω. Hence T is M-special. $\square$

Note that at limit steps we use just the assumption that the tree is normal. Hence we can generalize this lemma to $\mathbb{R}$-embeddable trees. The proof of the implication from right to left can be found in [Dev72].

Lemma 2.8. *Let T be an ω_1-Aronszajn tree. T is normal $\mathbb{R}$-embeddable if and only if T is* M-*special.*

Proof. ($\Rightarrow$) Let T be a normal $\mathbb{R}$-embeddable. Then $T^* = \bigcup_{\alpha<\omega_1} T_{\alpha+1}$ is $\mathbb{Q}$-embeddable and so $T^* = \bigcup_{n<\omega} A_n$ where A_n is an antichain for each n. The rest of the proof is the same as the proof of Lemma 2.7 since we used the antichains only in the successor step.

($\Leftarrow$) Let T be M-special. We define $f : T \to \mathbb{R}$ by setting

$$f(t) = \sum_{i=0}^{\infty} \frac{\mathcal{X}_{\mathrm{Rng}(t)}(i)}{10^i},$$

where $\mathcal{X}_X$ is the characteristic function of a set $X \subset \omega$. Since every node of T is a 1-1 function from some ordinal $\alpha < \omega_1$ to ω, if $s < t$ then $\mathrm{Rng}(s) \subset \mathrm{Rng}(t)$ and so there is $n < \omega$ such that $0 = \mathcal{X}_{\mathrm{Rng}(s)}(n) < \mathcal{X}_{\mathrm{Rng}(t)}(n) = 1$ and $\mathcal{X}_{\mathrm{Rng}(s)} \restriction n = \mathcal{X}_{\mathrm{Rng}(t)} \restriction n$. Hence $f(s) < f(t)$. $\square$

By Fact 2.5, if the tree T is $\mathbb{R}$-embeddable then $T \restriction S$ for $S = \{\alpha + 1 | \alpha < \omega_1\}$ is $\mathbb{Q}$-embeddable. So it is natural to introduce the concept of S-special for arbitrary unbounded subset of $S \subseteq \omega_1$. The following definition is from [She98].

Definition 2.9. Let S be an unbounded subset of ω_1. We say that an ω_1-tree T is *S-special* if $T \restriction S$ is $\mathbb{Q}$-embeddable, where

$$T \restriction S = \{t \in T | ht(t, T) \in S\},$$

with the induced ordering. We say that an ω_1-tree T is *$\mathcal{S}$-special* if there is S, an unbounded subset of ω_1, such that T is S-special. We denote the class of all $\mathcal{S}$-special ω_1-Aronszajn trees as $\mathbb{A}^{\mathcal{S}\text{-}sp}(\omega_1)$.

The following fact from [DJ74] says that if we only consider S-special trees for closed unbounded subsets S of ω_1, we get nothing new.

Fact 2.10. *Let C be a closed unbounded subset of ω_1. If T is a C-special ω_1-Aronszajn tree, then T is special.*

The following Fact 2.11, which can be found in [She98], says that if all Aronszajn trees are S-special for some given unbounded subset of ω_1, then all of them are in fact special. As an easy corollary, we have by Fact 2.5 that if every ω_1-Aronszajn tree is $\mathbb{R}$-embeddable, then every ω_1-Aronszajn tree is $\mathbb{Q}$-embeddable.

Fact 2.11. *Let S be an unbounded subset of ω_1. If every ω_1-Aronszajn tree is S-special then every ω_1-Aronszajn tree is special. In particular, if every ω_1-Aronszajn tree is $\mathbb{R}$-embeddable, then every ω_1-Aronszajn tree is $\mathbb{Q}$-embeddable.*

Note that $\mathcal{S}$-special Aronszajn trees, including special, $\mathbb{R}$-embeddable, and M-special Aronszajn trees, are not Suslin in the following strong sense: every uncountable subset of such tree contains an uncountable antichain. This motivates the following definition.

Definition 2.12. We say that an ω_1-tree T is *non-Suslin* if every uncountable subset U of T contains an uncountable antichain. We denote the class of all non-Suslin Aronszajn trees at ω_1 as $\mathbb{A}^{NS}(\omega_1)$.

The name of non-Suslin trees is inspired by the fact that every non-Suslin tree is not Suslin. On the other hand, every tree that is not non-Suslin has a Suslin subtree, as follows from the next fact that can be found in [Han81].

Fact 2.13. *Let T be an ω_1-Aronszajn tree. If T is not non-Suslin, then T has a subtree which is Suslin.*

Lemma 2.14. *Let T be an ω_1-Aronszajn tree. If T is $\mathcal{S}$-special, then T is non-Suslin.*

Proof. Assume for contradiction that T is an $\mathcal{S}$-special ω_1-Aronszajn tree which is not non-Suslin. By the previous fact T has a subtree T' which is Suslin. Since T is $\mathcal{S}$-special, T' is $\mathcal{S}$-special, too. Hence there is an unbounded subset S of ω_1 such that $T' \upharpoonright S = \bigcup_{n<\omega} A_n$, where A_n is an antichain for each n. By pigeon-hole principle, for some $n < \omega$ the size of A_n must be greater than ω. This contradicts the fact that T' is Suslin. $\square$

To sum up, for general trees we obtain:

$$\mathbb{A}^{sp}(\omega_1) = \mathbb{T}(\mathbb{Q})(\omega_1) \subseteq \mathbb{T}(\mathbb{R})(\omega_1) \subseteq \mathbb{A}^{\mathcal{S}\text{-}sp}(\omega_1) \subseteq \mathbb{A}^{NS}(\omega_1). \tag{2.3}$$

If we consider only normal trees, we get:

$$\mathbb{A}^{sp}(\omega_1) = \mathbb{T}(\mathbb{Q})(\omega_1) \subseteq \mathbb{T}(\mathbb{R})(\omega_1) = \mathbb{A}^{\text{M-}sp}(\omega_1) \subseteq \mathbb{A}^{\mathcal{S}\text{-}sp}(\omega_1) \subseteq \mathbb{A}^{NS}(\omega_1). \tag{2.4}$$

In the next section, for each of these inclusions, we examine if there is a model in which it is proper.

The existence of special Aronszajn trees at ω_1 can be proved in ZFC and by Baumgartner's theorem published in [BMR70] it is consistent with ZFC that every Aronszajn tree at ω_1 is special, so $\mathbb{A}^{sp}(\omega_1) = \mathbb{T}(\mathbb{R})(\omega_1) = \mathbb{A}^{\mathcal{S}\text{-}sp}(\omega_1) = \mathbb{A}^{NS}(\omega_1)$ is consistent with ZFC. On the other hand, consistently, each inclusion can be proper.

The following Fact 2.15 was first published in [Bau70]. It says that it is consistent that there is an Aronszajn tree which is M-special but not special. As a corollary we obtain that the first inclusion in (2.3) can be consistently proper.

Fact 2.15. *Assume $\Diamond$. Then there is a non-special Aronszajn tree which is a subtree of $\{s \in {}^{<\omega_1}\omega | s \text{ is } 1\text{-}1\}$. In particular, there is an $\mathbb{R}$-embeddable ω_1-Aronszajn tree which is not special.*

Proof. This has been proved by Baumgartner (see [Dev72]). We have extended his proof to obtain a more general result, see Theorem 3.27. □

The following lemma is a consequence of Fact 2.15 and it shows us that the second inclusion in (2.3) can be consistently proper.

Lemma 2.16. *Assume $\Diamond$. Then there is an ω_1-Aronszajn tree, which is $\mathcal{S}$-special and not $\mathbb{R}$-embeddable.*

Proof. By Fact 2.15, assuming $\Diamond$, there is an ω_1-Aronszajn tree which is $\mathbb{R}$-embeddable, but not $\mathbb{Q}$-embeddable. Let $\alpha < \omega_1$ be a limit ordinal and let $t \in T_\alpha$. For the chain $C = \{s \in T | s < t\}$ we add a new node t_C such that $t_C < t$ and $t_C > s$ for all $s \in C$. Consider the tree T' which is created by adding such t_C for every limit node t. Note that $\bigcup_{\alpha<\omega_1} T'_{\alpha+1} = T \setminus T_0$. Now, T' is not $\mathbb{R}$-embeddable since $\bigcup_{\alpha<\omega_1} T'_{\alpha+1}$ is not $\mathbb{Q}$-embeddable. But T' is S-special for $S = \{\alpha + 2 | \alpha < \omega_1\}$ since $T' \restriction S = \bigcup_{\alpha<\omega_1} T_{\alpha+1} \setminus T_1$. □

The claim that the last inclusion in (2.3) can be consistently proper is a consequence of the theorem published in [Sch14], which says that if ZFC is consistent, so is ZFC + SH[2] + there is an Aronszajn tree T at ω_1 which is not $\mathcal{S}$-special. If SH holds, then by Fact 2.13 every Aronszajn tree is non-Suslin. Therefore T is non-Suslin and it witnesses that ZFC + $\mathbb{A}^{\mathcal{S}\text{-}sp}(\omega_1) \neq \mathbb{A}^{NS}(\omega_1)$ is consistent.

3. Special Aronszajn trees at larger κ

3.1 Generalisations of $\mathbb{Q}$

In this section we consider some common generalisations of $\mathbb{Q}$ at higher cardinals. The following definitions of $\mathbb{Q}_\kappa$ and $\mathbb{Q}^*_\kappa$ are taken from [Tod84]. In addition, we introduce our definition of a generalisation of the real line for higher cardinals because we want to generalize the concept of an $\mathbb{R}$-embeddable tree (see Definition 2.2).

2 Suslin Hypothesis

Definition 3.1. Let κ be a regular cardinal. Then

(3.1) $\quad \mathbb{Q}^*_\kappa = (\{f \in {}^\omega\kappa \mid \{n < \omega \mid f(n) \neq 0\} \text{ is finite}\} \setminus \{\bar{0}\}, <_{\mathrm{lex}});$

(3.2) $\quad \mathbb{Q}_\kappa = (\{f \in {}^\kappa 2 \mid |\{\alpha < \kappa \mid f(\alpha) \neq 0\}| < \kappa\} \setminus \{\bar{0}\}, <_{\mathrm{lex}});$

(3.3) $\quad \mathbb{R}_\kappa = (\{f \in {}^\kappa 2 \mid (\neg\exists \alpha < \kappa)[f(\alpha) = 0 \text{ and } (\forall \beta > \alpha)(f(\beta) = 1)]\} \setminus \{\bar{0}, \bar{1}\}, <_{\mathrm{lex}});$

where $<_{lex}$ is the lexicographical ordering, $\bar{0}$ ($\bar{1}$) denotes the sequence of zeros (ones) of length ω in (3.1) and of length κ in (3.2) and (3.3).

Note that in the definition of $\mathbb{R}_\kappa$, we allow all 1's on a tail, but restrict this configuration by demanding that in this case there is no greatest α with $f(\alpha) = 0$.[3]

Remark 3.2. Note that $\mathbb{Q}_\omega \cong \mathbb{Q} \cong \mathbb{Q}^*_\omega$. On the other hand, for $\kappa > \omega$, $\mathbb{Q}_\kappa \not\cong \mathbb{Q}^*_\kappa$, even if $|\mathbb{Q}_\kappa| = \kappa$. This holds, because $\mathbb{Q}^*_\kappa$ does not contain any decreasing sequence of uncountable length. However, in $\mathbb{Q}_\kappa$ there are decreasing sequences of length κ.

In this paper we work mainly with $\mathbb{Q}_\kappa$ because it has some nice properties: in particular, one can generalize Kurepa's Theorem for $\mathbb{Q}_\kappa$ and prove Lemma 3.4 which is very useful and plays the key role in proving Lemma 3.13. On the other hand, the main advantage of $\mathbb{Q}^*_\kappa$ is that it always has size κ. When we work with $\mathbb{Q}_\kappa$, we need to assume that $\kappa^{<\kappa} = \kappa$ to control its size.

The following easy lemma tells us that $\mathbb{Q}_\kappa$ has the properties which we want from a generalisation of $\mathbb{Q}$, with the exception that it does not have to have size κ. The proof is left as an exercise.

Lemma 3.3. *The ordering $\mathbb{Q}_\kappa$ is linear, dense, without endpoints and $|\mathbb{Q}_\kappa| = \kappa^{<\kappa}$.*

There is an asymmetry in $\mathbb{Q}_\kappa$ between decreasing and increasing sequences:

Lemma 3.4. *Assume $\kappa > \omega$ is regular.*

(i) Let $A = \langle f_\alpha \mid \alpha < \lambda \rangle$ be a strictly decreasing sequence in $\mathbb{Q}_\kappa$, where λ is a limit ordinal such that $\omega \leq \lambda < \kappa$. Then A does not have an infimum in $\mathbb{Q}_\kappa$.

(ii) Let $B = \langle g_\alpha \mid \alpha < \lambda \rangle$ be a strictly increasing sequence in $\mathbb{Q}_\kappa$ where λ is a limit ordinal such that $\omega \leq \lambda < \kappa$. Then B has a supremum in $\mathbb{Q}_\kappa$.

Proof. Ad (i). Let $A = \langle f_\alpha \mid \alpha < \lambda \rangle$ be given. Assume for contradiction that there is the infimum $f \in \mathbb{Q}_\kappa$ of A. Since $f \in \mathbb{Q}_\kappa$, there is $\beta_0 < \kappa$ such that for each $\beta \geq \beta_0$ $f(\beta) = 0$. Since $\lambda < \kappa$ and κ is regular, there is $\gamma_0 < \kappa$ such that for each $\gamma \geq \gamma_0$ and for each $\alpha < \lambda$ $f_\alpha(\gamma) = 0$. Let $\delta = \max\{\beta_0, \gamma_0\}$. We define $f^* = f \restriction \delta \cup \{\langle \delta, 1 \rangle\} \cup \{\langle \beta, 0 \rangle \mid \beta > \delta\}$. Clearly, $f^* > f$. Since $f < f_\alpha$ for every $\alpha < \lambda$ and since $\delta = \max\{\beta_0, \gamma_0\}$, $f^* < f_\alpha$ for every $\alpha < \lambda$. This is a contradiction because we assume that f is the infimum of A.

Ad (ii). Let $B = \langle g_\alpha \mid \alpha < \lambda \rangle$ be given. We define supremum g by induction on $\beta < \kappa$.

For $\beta = 0$. Set

$$g(0) = \begin{cases} 1 & \text{if } \exists \alpha < \lambda (g_\alpha(0) = 1); \\ 0 & \text{otherwise.} \end{cases}$$

[3] If $f \in {}^\kappa 2$ does not satisfy (3.3) and α is the greatest position with 0, then we can define $g \in \mathbb{R}_\kappa$ which is the immediate successor of f in the lexicographical order: define g exactly as f below α, and set $g(\beta) = 1$ for all $\beta \geq \alpha$. To prohibit this situation (which violates density of the ordering), we choose to disallow such f's in (3.3). If there is no greatest α where $f(\alpha) = 0$, this problem does not arise.

Assume that $g \restriction \beta$ is defined, then we define $g(\beta)$ as follows:

$$g(\beta) = \begin{cases} 1 & \text{if } \exists\alpha < \lambda \text{ such that } g_\alpha(\beta) = 1 \text{ and } g_\alpha \restriction (\beta+1) > g \restriction \beta \cup \{\langle \beta, 0\rangle\}; \\ 0 & \text{otherwise.} \end{cases}$$

First note that g is in $\mathbb{Q}_\kappa$ since κ is regular and $\lambda < \kappa$.

Now, we show that g is the supremum of B. It is obvious that $g_\alpha < g$ for every $\alpha < \lambda$. Hence it is enough to show that g is the least upper bound of B. Let $h < g$ be given. Then there is $\beta_0 < \kappa$ such that $h \restriction \beta_0 = g \restriction \beta_0$ and $0 = h(\beta_0) < g(\beta_0) = 1$. By definition of g there is α such that $g_\alpha \restriction (\beta_0 + 1) > g \restriction \beta_0 \cup \langle \beta_0, 0\rangle$. As $h \restriction \beta_0 = g \restriction \beta_0$ and $h(\beta_0) = 0$, $g \restriction \beta_0 \cup \langle \beta_0, 0\rangle = h \restriction (\beta_0+1)$ and so $g_\alpha \restriction (\beta_0+1) > h \restriction (\beta_0+1)$. Therefore $g_\alpha > h$. □

Note that it was important in (i) of the previous lemma that λ is a limit ordinal $< \kappa$. One can easily find decreasing sequences in $\mathbb{Q}_\kappa$ of length κ which do have the infimum.[4]

Now, we present the generalisation of Kurepa's Theorem for $\mathbb{Q}_\kappa$:

Theorem 3.5. ***(Generalised Kurepa's Theorem)*** *Assume $\kappa^{<\kappa} = \kappa$. Let $(E, <)$ be a partially ordered set. Then the following are equivalent:*

(i) E is embeddable in $\mathbb{Q}_\kappa$;

(ii) E is the union of at most κ-many antichains.

Proof. (i) $\Rightarrow$ (ii) Let f be the embedding. Let $\{q_\alpha | \alpha < \kappa\}$ be an enumeration of $\mathbb{Q}_\kappa$. We define $A_\alpha = f^{-1}(q_\alpha)$ for each $q_\alpha \in \mathrm{Rng}(f)$. Obviously, each A_α is an antichain since f is an embedding.

(ii) $\Rightarrow$ (i) We assume that $\bigcup_{\alpha<\kappa} A_\alpha = E$, where each A_α is an antichain. Moreover, without loss of generality, we may assume that for each $\beta, \alpha < \kappa$, $A_\alpha \cap A_\beta = \emptyset$. Let $f : E \to \kappa$ be a function such that $A_\alpha = f^{-1}(\alpha)$. For $x \in E$ define $g(x)$ so that $g(x)(\alpha) = 1$ if and only if $\alpha \leq f(x)$ and $\{y \in E | y \leq x\} \cap A_\alpha \neq \emptyset$.

Notice that $g(x)$ is in $\mathbb{Q}_\kappa$ because $g(x)(\alpha) = 1$ implies that $\alpha \leq f(x)$, where $f(x) \in \kappa$.

Now, we check that g is an embedding. Assume that $x < y$ are in E and $x \in A_\alpha, y \in A_\beta$ for some $\beta \neq \alpha$. We distinguish two cases.

First suppose that $\alpha < \beta$. Then $g(x)(\alpha) = 1$ and also $g(y)(\alpha) = 1$ since $x < y$ and $x \in A_\alpha$. And for all $\gamma < \alpha$ if $g(x)(\gamma) = 1$ then $g(y)(\gamma) = 1$ and so $g(x) \restriction \alpha \leq_{lex} g(y) \restriction \alpha$. If $g(x) \restriction \alpha <_{lex} g(y) \restriction \alpha$, then $g(x) < g(y)$ and we are finished. If $g(x) \restriction \alpha = g(y) \restriction \alpha$, then we can continue as follows: for all $\gamma > \alpha$ it holds that $g(x)(\gamma) = 0$ since $\gamma > f(x)$. Hence $g(x)(\beta) = 0$ and $g(y)(\beta) = 1$; therefore $g(x) < g(y)$.

Next suppose that $\beta < \alpha$. Again for all $\gamma < \beta$, if $g(x)(\gamma) = 1$ then $g(y)(\gamma) = 1$ and so $g(x) \restriction \beta \leq_{lex} g(y) \restriction \beta$. Now, we show that $g(x)(\beta) = 0$ and $g(y)(\beta) = 1$. Assume for contradiction that $g(x)(\beta) = 1$. Then by definition of the function g, we know there exists $z \in A_\beta$ and $z \leq x$. Hence $z < y$ and this is a contradiction since there are two comparable elements in A_β. By the definition of g, $g(y)(\beta) = 1$ and so $g(x) < g(y)$. □

Remark 3.6. Note that the assumption $\kappa^{<\kappa} = \kappa$ is necessary just in the proof of (i) $\Rightarrow$ (ii).

[4] Compare with $\mathbb{Q}$: some infinite decreasing sequences have an infimum; since there is no limit ordinal below ω, the analogue of (i) of the previous Lemma does not appear in $\mathbb{Q}$.

Remark 3.7. We cannot prove Kurepa's Theorem for $\mathbb{Q}^*_\kappa$, for $\kappa > \omega$ a regular cardinal, since it does not contain strictly decreasing sequence of uncountable length. Consider the ordinal κ with reverse ordering $<^*$, i.e. $\alpha <^* \beta$ that $\alpha <^* \beta$ if and only if $\beta \in \alpha$ for α, $\beta \in \kappa$. Then κ is a union of κ-many antichains and cannot be embedded to $\mathbb{Q}^*_\kappa$.

Partials orders from Theorem 3.5 have another useful characterisation. The proof of the following lemma is easy and it is left as an exercise for the reader.

Lemma 3.8. *Let κ be regular and let $(E, <)$ be a partially ordered set. Then the following are equivalent:*
(i) E is the union of at most κ-many antichains;
(ii) there is $f : E \to \kappa$ such that if s, t are comparable in E, then $f(s) \neq f(t)$.

Now, we focus on the partial order $\mathbb{R}_\kappa$. We show that it has similar properties as $\mathbb{R}$.

Lemma 3.9. *The partial order $\mathbb{R}_\kappa$ is*
(i) linear, without endpoints;
(ii) $\mathbb{Q}_\kappa$ is dense in $\mathbb{R}_\kappa$;
(iii) Dedekind complete.

Proof. It is easy to verify that $\mathbb{R}_\kappa$ satisfies (i).

Ad (ii). Let $f <_{\mathbb{R}_\kappa} g$ in $\mathbb{R}_\kappa$ be given. Let α_0 be the least ordinal such that $0 = f(\alpha_0) < g(\alpha_0) = 1$. By definition of $\mathbb{R}_\kappa$, there is the least $\beta_0 > \alpha_0$ such that $f(\beta_0) = 0$. Let $h = f \restriction \beta_0 \cup \{\langle \beta_0, 1\rangle\} \cup \{\langle \gamma, 0\rangle \mid \gamma > \beta_0\}$. It is easy to see that $h \in \mathbb{Q}_\kappa$ and $f <_{\mathbb{R}_\kappa} h <_{\mathbb{R}_\kappa} g$.

Ad (iii). It is enough to show that every increasing sequence with upper bound has the supremum. First note that each increasing sequence in $\mathbb{R}_\kappa$ has cardinality at most $\kappa^{<\kappa}$ since $\mathbb{Q}_\kappa$ is dense in $\mathbb{R}_\kappa$ as we proved in the previous paragraph. Let $A = \langle f_\alpha \in \mathbb{R}_\kappa | \alpha < \lambda\rangle$ for some ordinal $\lambda \leq \kappa^{<\kappa}$ be given and let $f \in \mathbb{R}_\kappa$ be the upper bound of A. Let F_C be a choice function from $\mathcal{P}(\mathbb{Q}_\kappa)$ to $\mathbb{Q}_\kappa$. We define the sequence $A_{\mathbb{Q}}$ in $\mathbb{Q}_\kappa$ as follows:

$$A_{\mathbb{Q}} = \langle g_\alpha \in \mathbb{Q}_\kappa | g_\alpha = F_C(\{q \in \mathbb{Q}_\kappa | f_\alpha < q < f_{\alpha+1}\}) \text{ and } \alpha < \lambda\rangle . \tag{3.4}$$

We show that $A_{\mathbb{Q}_\kappa}$ has the supremum g in $\mathbb{R}$ and that g is also the supremum of A in $\mathbb{R}_\kappa$. We define a function $g^* : \kappa \to 2$ by induction on $\beta < \kappa$.

For $\beta = 0$. Set

$$g^*(0) = \begin{cases} 1 & \text{if } \exists \alpha < \lambda (g_\alpha(0) = 1); \\ 0 & \text{otherwise.} \end{cases}$$

Let $g^* \restriction \beta$ be defined, then we define $g^*(\beta)$ as follows:

$$g^*(\beta) = \begin{cases} 1 & \text{if } \exists \alpha < \lambda \text{ such that } g_\alpha(\beta) = 1 \text{ and } g_\alpha \restriction (\beta + 1) > g^* \restriction \beta \cup \{\langle \beta, 0\rangle\}; \\ 0 & \text{otherwise.} \end{cases}$$

Note that g^* may not be in $\mathbb{R}_\kappa$, but it holds that $g^* \neq \{\langle \alpha, 1\rangle \mid \alpha < \kappa\}$ since the sequence has an upper bound in $\mathbb{R}_\kappa$.

Now, we need to show that g^* is the supremum of $A_{\mathbb{Q}}$ in $(2^\kappa, <_{\text{lex}})$. However, the proof of this is the same as the proof of Lemma 3.4 (ii). Note that in the Lemma 3.4 (ii) we used the assumption that the sequence has length less than κ just for showing that the supremum is in $\mathbb{Q}_\kappa$.

As we mentioned earlier, g^* may not be in $\mathbb{R}_\kappa$, but note that $g^* \neq \{\langle \alpha, 1\rangle \mid \alpha < \kappa\}$. If g^* is not in $\mathbb{R}_\kappa$, there is $\beta_0 < \kappa$ such that $g^*(\beta_0) = 0$ and $g^*(\beta) = 1$ for every $\beta > \beta_0$. Let $\overline{g} = g^* \restriction \beta_0 \cup \{\langle \beta_0, 1\rangle\} \cup \{\langle \beta, 0\rangle \mid \beta > \beta_0\}$. Clearly $\overline{g} \in \mathbb{R}_\kappa$ and there is no function between g^* and $\overline{g}$ in 2^κ. Now we define $g \in \mathbb{R}_\kappa$ by

$$g = \begin{cases} g^* & \text{if } g^* \in \mathbb{R}_\kappa; \\ \overline{g} & \text{otherwise.} \end{cases}$$

It is obvious that $g \in \mathbb{R}_\kappa$ and since g^* is the supremum of $A_{\mathbb{Q}_\kappa}$ in 2^κ, g is the supremum of $A_{\mathbb{Q}_\kappa}$ in $\mathbb{R}_\kappa$.

To finish the proof of the theorem, it suffices to show that g is also the supremum of A. The function g is clearly the upper bound of A. Now, we show that g is the least upper bound. Let $h < g$. Since g is the supremum of $A_{\mathbb{Q}_\kappa}$, there is $q \in A_{\mathbb{Q}_\kappa}$, such that $h < q$. But $q < r$ for some $r \in A$ by the definition of $A_{\mathbb{Q}_\kappa}$. Hence $h < r$. □

3.2 Classification

In the previous section we have built the foundations for the investigation of special κ^+-Aronszajn trees for any regular κ. We introduced the concept of special, $\mathbb{R}$-embeddable, M-special and $\mathcal{S}$-special ω_1-Aronszajn trees. Now, we generalize these concepts to higher Aronszajn trees, which are in the center of our interest. When we talk about an Aronszajn tree in this section, we mean a κ^+-Aronszajn tree for some regular cardinal $\kappa > \omega$.

Definition 3.10. Let κ be a cardinal. We say that κ^+-Aronszajn tree T is *special* if T is a union of κ-many antichains. We denote the class of all special Aronszajn trees at κ^+ as $\mathbb{A}^{sp}(\kappa^+)$.

As in the previous section, the concept of a special Aronszajn tree has more equivalent definitions. However, we need to be careful when we talk about $\mathbb{Q}_\kappa$-embeddability, since this partial order in general does not have to have size κ.

Lemma 3.11. *Let κ be regular. The following are equivalent for a κ^+-Aronszajn tree T:*
(i) T is special;
(ii) There is $f : T \to \kappa$ such that if s, t are comparable in T, then $f(s) \neq f(t)$.

Proof. This is a direct consequence of Lemma 3.8. □

Lemma 3.12. *Assume $\kappa^{<\kappa} = \kappa$. Then κ^+-Aronszajn tree T is special if and only if T is $\mathbb{Q}_\kappa$-embeddable.*

Proof. It follows from Theorem 3.5. □

Again as in the previous section, we can characterise $\mathbb{R}_\kappa$-embeddable Aronszajn trees using $\mathbb{Q}_\kappa$-embeddable Aronszajn trees. This is our generalisation of Fact 2.5.

Theorem 3.13. *Assume $\kappa^{<\kappa} = \kappa$. Let T be an κ^+-tree. T is $\mathbb{R}_\kappa$-embeddable if and only if $T^* = \bigcup_{\alpha<\kappa^+} T_{\alpha+1}$ is $\mathbb{Q}_\kappa$-embeddable.*

Proof. ($\Rightarrow$) Let T be $\mathbb{R}_\kappa$-embeddable and $T^* = \bigcup_{\alpha<\kappa^+} T_{\alpha+1}$. Let f be the embedding, $t \in T^*$ and let $s \in T$ be the immediate predecessor of t. We define $f' : T^* \to \mathbb{Q}_\kappa$ as follows: $f'(t) = q$ where $q \in \mathbb{Q}_\kappa$ such that $f(s) < q < f(t)$.

($\Leftarrow$) Let $T^* = \bigcup_{\alpha<\kappa^+} T_{\alpha+1}$ be $\mathbb{Q}_\kappa$-embeddable and let f be the embedding.

We first define a function $g : \mathbb{Q}_\kappa \to \mathbb{Q}_\kappa \times \mathbb{Q}_\kappa$ which will "replace" every $q \in \mathbb{Q}_\kappa$ with an open interval $(g(q)^1, g(q)^2)$,[5] while preserving the ordering. More precisely, we will define g by induction on κ and ensure it satisfies the following for all $q < q'$ in $\mathbb{Q}_\kappa$:

$$g(q)^2 < g(q')^1. \tag{3.5}$$

Enumerate $\mathbb{Q}_\kappa$ as $\{q_\beta | \beta < \kappa\}$. We will construct by induction on $\alpha < \kappa$ embeddings $g_\alpha : \{q_\beta | \beta < \alpha\} \to \mathbb{Q}_\kappa \times \mathbb{Q}_\kappa$ which will be used to define the final function g.

As we will see below, at the successor step, we define $g_{\alpha+1}$ as an extension of g_α to q_α. Suppose $g_{\alpha+1}(q_\alpha) = \langle q, q' \rangle$ for some $q < q'$ in $\mathbb{Q}_\kappa$. In addition to choosing q, q', fix also two elements $a(q_\alpha) < b(q_\alpha)$ in the interval (q, q') and two sequences as follows: a strictly increasing sequence of elements in (q, q') of length κ converging to $a(q_\alpha)$ and a strictly decreasing sequence of elements in (q, q') of length κ converging to $b(q_\alpha)$. We denote these sequences $\langle a(q_\alpha)_i | i < \kappa \rangle$ and $\langle b(q_\alpha)_i | i < \kappa \rangle$, respectively.

Now we provide an inductive definition of the functions g_α, $\alpha < \kappa$:

Set $g_0 = \emptyset$.

Let α be a limit ordinal. Define

$$g_\alpha = \{\langle q_\beta, \langle a(q_\beta)_{\alpha+1}, b(q_\beta)_{\alpha+1} \rangle \rangle | \beta < \alpha\}.$$

The idea behind this definition is to take the intervals defined in the previous stages of the construction and "shrink" them to get more space. The shrinking of the intervals makes sure that the construction can continue on the successor steps.

At $\alpha + 1$, define $g_{\alpha+1}$ by

$$g_{\alpha+1} = g_\alpha \cup \{\langle q_\alpha, \langle q, q' \rangle \rangle\},$$

for some suitable interval (q, q'), i.e. for all $s < s'$ in the domain of $g_{\alpha+1}$, we should have $g_{\alpha+1}(s)^2 < g_{\alpha+1}(s')^1$.[6]

When all functions g_α, $\alpha < \kappa$, have been constructed, set

$$g = \{\langle q_\alpha, \langle a(q_\alpha), b(q_\alpha) \rangle \rangle \, | \alpha < \kappa\}.$$

By the construction, it follows that g is as required.

Now we can finish the proof of the theorem. Define a function $i : \mathbb{Q}_\kappa \to \mathbb{Q}_\kappa$ by $i(q) = r$, where r is some element of the open interval $(g(q)^1, g(q)^2)$. We define an embedding $f' : T \to \mathbb{R}_\kappa$ as follows:

$$f'(t) = \begin{cases} i(f(t)) & \text{if } t \in T_{\alpha+1} \text{ for } \alpha < \kappa^+; \\ \sup\{i(f(s)) | s < t \text{ and } s \in T_{\beta+1} \text{ and } \beta < \alpha\} & \text{otherwise.} \end{cases}$$

---

5 $g(q)^1$ denotes the left coordinate and $g(q)^2$ the right coordinate of the pair $g(g)$.

6 When defining $g_{\alpha+1}$, we need to ensure that we can map q_α into an interval which is disjoint from the intervals $g_\alpha(\beta)$, $\beta < \alpha$, while respecting the ordering. Without the shrinking at the limit stages of the construction, the intervals might converge in a way which prevents the definition of $g_{\alpha+1}(q_\alpha)$.

Now we need to check that the function f' is the embedding of T to $\mathbb{R}$. If $s < t$ and s, $t \in T^*$, then it is easy to see that $f'(s) < f'(t)$ because i is order-preserving. If $t \in T_\alpha$ for α limit, then $f'(s) < f'(t)$ since $f'(t)$ is the supremum. The only interesting case is if $s \in T_\alpha$ for α limit and $t \in T_{\alpha+1}$. Then we need to show

$$f'(t) = i(f(t)) > \sup\{i(f(r)) | r < s \text{ and } r \in T_{\beta+1} \text{ and } \beta < \alpha\} = f'(s). \tag{3.6}$$

This follows from the construction of g. For every $r < s$ it holds that $i(f(r)) < q < i(f(t))$ where $q = g(f(t))^1$. Hence

$$f'(s) = \sup\{i(f(r)) | r < s \text{ and } r \in T_{\beta+1} \text{ and } \beta < \alpha\} \leq q < i(f(t)) = f'(t). \tag{3.7}$$

□

Definition 3.14. Let κ be a cardinal. We say that κ^+-Aronszajn tree T is M-*special* if T is isomorphic to a subtree of $\{s \in {}^{<\kappa^+}\kappa | s \text{ is 1-1}\}$

The following lemma is a generalisation of Lemma 2.7, hence we left the proof as an exercise.

Lemma 3.15. *Let κ be a regular cardinal. If T is a normal special κ^+-Aronszajn tree then T is* M-*special.*

As in the case for ω_1, at the limit step we use just the assumption that the tree is normal. Hence we can generalize this lemma to the following lemma. Note that for this we do not need the assumption $\kappa^{<\kappa} = \kappa$ since we use that the tree $\bigcup_{\alpha<\kappa^+} T_{\alpha+1}$ is special instead of $\mathbb{Q}_\kappa$-embeddable. We explicitly state this lemma here so it is clear that M-special trees are exactly those trees that are normal and whose successor levels form a special tree, as was the case at ω_1.

Lemma 3.16. *Let κ be a regular cardinal. Let T be a normal κ^+-Aronszajn tree. Then $T^* = \bigcup_{\alpha<\kappa^+} T_{\alpha+1}$ is special if and only if T is* M-*special.*

Proof. ($\Rightarrow$) Let $T^* = \bigcup_{\alpha<\kappa^+} T_{\alpha+1}$ be special. Then $T^* = \bigcup_{\xi<\kappa} A_\xi$ where A_ξ is an antichain for each $\xi < \kappa$. The rest of the proof is the same as the proof of Lemma 3.15.

($\Leftarrow$) Let T be an M-special tree. Then T is isomorphic to a subtree T' of $\{s \in {}^{<\kappa^+}\kappa | s \text{ is 1-1}\}$ via i. We define $f : T^* \to \kappa$ by setting $f(t) = i(t)(\alpha)$ for $ht(t,T) = \alpha + 1$. Let $s < t \in T^*$. Then $ht(s,T) = \beta + 1 < \alpha + 1 = ht(t,T)$. Since $i(s) \subset i(t)$, $i(s)(\beta) = i(t)(\beta)$. As $i(t)$ is 1-1, $i(t)(\beta) \neq i(t)(\alpha)$. Therefore $f(s) \neq f(t)$.

□

On the other hand, generalisation of Lemma 2.8 requires the additional assumption that $\kappa^{<\kappa} = \kappa$ since we need to use Generalised Kurepa's Theorem.

Lemma 3.17. *Assume $\kappa^{<\kappa} = \kappa$. Let T be a κ^+-Aronszajn tree. T is a normal $\mathbb{R}_\kappa$-embeddable tree if and only if T is* M-*special.*

Proof. It follows by Theorem 3.13 and Lemmas 3.12 and 3.16. □

Unlike special ω_1-Aronszajn trees, it is not provable in ZFC that special κ-Aronszajn trees exist for $\kappa > \omega_1$. Hence we are also interested in the question how the existence of one kind of special Aronszajn trees influences the existence of other kinds of special

Aronszajn trees. The following lemma claims that if there are no M-special Aronszajn trees then there are no special Aronszajn trees at all.

Lemma 3.18. *Let κ be a regular cardinal. If there exists a special κ^+-Aronszajn tree, then there exists an* M*-special Aronszajn tree.*

Proof. Let T be a special κ^+-Aronszajn tree. We first add one root r such that $r < t$ for each $t \in T_0$. Now we wish to guarantee the condition (iv) of Definition 1.5. Let $\alpha < \kappa^+$ be a limit ordinal and let C be a cofinal branch in $T \restriction \alpha$ such that there exists node $t \in T$ greater than all nodes $c \in C$. Then we add one extra node t_C to the limit level α such that $t_C > c$ for all $c \in C$ and $t_C < t$ for all $t > C$, where $t > C$ means $t > c$ for all $c \in C$.

Since for every chain we add one extra node to the limit level, this new tree satisfies (iv). Denote this tree T'. This tree is normal and $T = \bigcup_{\alpha<\kappa^+} T'_{\alpha+1}$. By Lemma 3.16 the tree T' is M-special. □

As in previous section it makes sense to introduce the concept of $\mathcal{S}$-special Aronszajn trees.

Definition 3.19. Let κ be a regular cardinal and S be an unbounded subset of κ^+. We say that the κ^+-tree T is S*-special* if $T \restriction S$ is special, where $T \restriction S = \{t \in T | ht(t,T) \in S\}$ with the induced ordering. We say that a κ^+-tree T is $\mathcal{S}$*-special* if there is S, an unbounded subset of κ^+, such that T is S-special. We denote the class of all $\mathcal{S}$-special κ^+-Aronszajn trees as $\mathbb{A}^{\mathcal{S}\text{-}sp}(\kappa^+)$.

The proofs of the following lemmas are direct generalisations of proofs of Facts 2.10 and 2.11.

Lemma 3.20. *Let C be a closed unbounded subset of κ^+, where κ is a regular cardinal. If T is a C-special κ^+-Aronszajn tree, then T is special.*

Proof. Let T be a C-special κ^+-Aronszajn tree. Then $T \restriction C = \bigcup_{\nu<\kappa} A_\nu$, where each A_ν is an antichain. Let $\{a^\nu_\alpha | \alpha < \kappa^+\}$ be an enumeration of A_ν for each $\nu < \kappa$. Let $\{c_\alpha | \alpha < \kappa^+\}$ be the monotone enumeration of C. For $\alpha < \kappa^+$ and for $x \in T_{c_\alpha}$, we define $S_x = \{y \in T \restriction c_{\alpha+1} | x <_T y\}$. Since each S_x has size less than κ^+, let $\{s_\mu(x) | \mu < \kappa\}$ be an enumeration of S_x. Set

$$A_{\nu,\mu} = \{s_\mu(a^\nu_\alpha) | \alpha < \kappa^+\}. \tag{3.8}$$

Clearly, $A_{\nu,\mu}$ is an antichain of T for each $\nu, \mu < \kappa$. Since C is closed unbounded, $T = \bigcup_{\nu<\kappa} A_\nu \cup \bigcup_{\nu,\mu<\kappa} A_{\nu,\mu}$. Hence T is special. □

Lemma 3.21. *Let κ be a regular cardinal and S be an unbounded subset of κ^+. If every κ^+-Aronszajn tree is S-special then every κ^+-Aronszajn tree is special.*

Proof. Let $S = \{\alpha_\mu | \mu < \kappa^+\}$ be an unbounded subset of κ^+ and T be a S-special κ^+ Aronszajn tree. We define a new tree

$$T' = \{\langle t, \beta\rangle \,| t \in T \text{ and } \beta < \alpha_{ht(t,T)} \text{ and } \forall s < t(\alpha_{ht(s,T)} < \beta)\}. \tag{3.9}$$

The tree T' is ordered by $<_{T'}$ as follows: $\langle t, \beta\rangle <_{T'} \langle s, \gamma\rangle$ if and only if $t < s$ or ($t = s$ and $\beta < \gamma$). It is obvious that T satisfies our definition of Aronszajn tree. Hence T' is

$\mathcal{S}$-special, i.e. $T' \restriction S$ is special. Since T is isomorphic to $T' \restriction S = \{\langle t, \alpha_{ht(t,T)}\rangle | t \in T\}$, T is special. $\square$

Again, note that $\mathcal{S}$-special κ^+-Aronszajn trees are not Suslin in a strong sense. This means that every subset of size κ^+ of such tree contains an antichain of size κ^+. Hence we can generalize Definition 2.12 and Lemma 2.14.

Definition 3.22. Let κ be a regular cardinal and T be a κ^+-Aronszajn tree. We say that T is *non-Suslin* if every subset U of T, which has size κ^+, contains an antichain of size κ^+. We denote the class of all non-Suslin Aronszajn trees at κ^+ as $\mathbb{A}^{NS}(\kappa^+)$.

The proof of following lemma is a direct generalisation of proof of Fact 2.13.

Lemma 3.23. *Let κ be a regular cardinal and T be a κ^+-Aronszajn tree. If T is not non-Suslin, then T has a subtree which is Suslin.*

Proof. Let T be a κ^+-Aronszajn tree, which is not non-Suslin. Then there is a subset X of T such that $|X| = \kappa^+$ and X does not contain antichain of size κ^+. We define $T' = \{s \in T | \exists t \in X(s < t)\}$. It is easy to verify that T' is Suslin. $\square$

The proof of following lemma is a direct generalisation of proof of Lemma 2.14.

Lemma 3.24. *Let κ be a regular cardinal and T be a κ^+-Aronszajn tree. If T is $\mathcal{S}$-special, then T is non-Suslin.*

The following theorem is only the summary of what we have showed about the relative existence of different kinds of special Aronszajn trees. It tells us that the weak tree property[7] at κ^+ is equivalent to the claim that there are no M-special κ^+-Aronszajn trees and also to the claim that there are no $\mathcal{S}$-special κ^+-Aronszajn trees.

Theorem 3.25. *Let κ be a regular. The following are equivalent*

(i) $\mathbb{A}^{sp}(\kappa^+) = \emptyset$;
(ii) $\mathbb{A}^{M}(\kappa^+) = \emptyset$;
(iii) $\mathbb{A}^{\mathcal{S}-sp}(\kappa^+) = \emptyset$.

Proof. Ad (i) $\Leftrightarrow$ (ii). The claim from left to right follows from Lemma 3.16 and the converse follows from Lemma 3.18.

Ad (i) $\Leftrightarrow$ (iii). This follows from the definition of $\mathcal{S}$-special κ^+-Aronszajn tree. $\square$

To sum up:

$$\mathbb{A}^{sp}(\kappa^+) \subseteq \mathbb{A}^{\mathcal{S}-sp}(\kappa^+) \subseteq \mathbb{A}^{NS}(\kappa^+) \text{ and } \mathbb{A}^{M}(\kappa^+) \subseteq \mathbb{A}^{\mathcal{S}-sp}(\kappa^+). \tag{3.10}$$

If moreover we only consider normal trees and assume that $\kappa^{<\kappa} = \kappa$, we get:

$$\mathbb{A}^{sp}(\kappa^+) = \mathbb{T}(\mathbb{Q}_\kappa)(\kappa^+) \subseteq \mathbb{T}(\mathbb{R}_\kappa)(\kappa^+) = \mathbb{A}^{\text{M-}sp}(\kappa^+) \subseteq \mathbb{A}^{\mathcal{S}\text{-}sp}(\kappa^+) \subseteq \mathbb{A}^{NS}(\kappa^+). \tag{3.11}$$

7 We say that a cardinal κ^+ has the weak tree property, if there are no special κ^+-Aronszajn trees.

We are interested in special Aronszajn trees at successors of regular cardinals. While the existence of a special ω_1-Aronszajn tree can be proved in ZFC, at higher cardinals we need some additional assumption, for example $\kappa^{<\kappa} = \kappa$ or weak square principle. The first one was used in construction by Specker in [Spe49] and the second one in the construction by Jensen in [Jen72]. On the other hand, it is possible to find a model with no special κ^+-Aronszajn tree where $\kappa > \omega$ is regular, but this requires much stronger assumption. Throughout this section we assume that κ is a regular cardinal and $\kappa > \omega$.

Definition 3.26. $E_\kappa^{\kappa^+} = \{\alpha < \kappa^+ | cf(\alpha) = \kappa\}$

This theorem is our generalisation of Fact 2.15. As a corollary we obtain that the first inclusion in (3.10) can be consistently proper.

Theorem 3.27. *Assume $\kappa^{<\kappa} = \kappa$ and $\diamondsuit_{\kappa^+}(E_\kappa^{\kappa^+})$. Then there is an* M*-special κ^+-Aronszajn tree, which is not special.*

Proof. By $\diamondsuit_{\kappa^+}(E_\kappa^{\kappa^+})$ there is a sequence $\langle f_\alpha | \alpha \in E_\kappa^{\kappa^+}\rangle$ such that f_α is a function from α to α and for any function $f : \kappa^+ \to \kappa^+$ the set $\{\alpha \in E_\kappa^{\kappa^+} | f_\alpha = f \restriction \alpha\}$ is stationary in κ^+. We fix this sequence for the rest of the proof.

We construct the tree T and the function $\pi : T \to \kappa^+$, which will code the tree in κ^+, by induction on $\alpha < \kappa^+$. For each $\alpha < \kappa^+$ we require the following conditions:

(T1) If $s \in T \restriction \alpha$ then $|\kappa \setminus \mathrm{Rng}(s)| = \kappa$.

(T2) If $s \in T \restriction \alpha$ and $x \in [\kappa \setminus \mathrm{Rng}(s)]^{<\kappa}$ then there is $s' \supseteq s$ on each higher level of $T \restriction \alpha$ such that $\mathrm{Rng}(s') \cap x = \emptyset$.

($\pi 0$) π_α is a 1-1 map from $T \restriction \alpha$ to κ^+ such that $s \subseteq t \to \pi_\alpha(s) < \pi_\alpha(t)$ and for $\beta < \alpha$, $\pi_\beta \subseteq \pi_\alpha$.

Let $T_0 = \{\emptyset\}$ and π_1 is an arbitrary function from $T \restriction 1 = T_0$ to κ^+. It is clear that T_0 satisfies both conditions and π_1 satisfies ($\pi 0$).

Let $\alpha = \beta + 1$. Suppose $T \restriction (\beta + 1)$ and $\pi_{\beta+1}$ are defined and they satisfy the conditions mentioned above. We want to construct level T_α. For each $s \in T_\beta$ we add all one-point extensions $s \cup \{\langle \alpha, \nu\rangle\}$ of s such that $\nu \in \kappa \setminus \mathrm{Rng}(s)$. This is possible by (T1), which guarantees the existence of κ-many such extensions. Since we add all such extension of s, for each $x \in [\kappa \setminus \mathrm{Rng}(s)]^{<\kappa}$ we can always find $t \in T_\alpha$ such that $s \subseteq t$ and $x \cap \mathrm{Rng}(s) = \emptyset$; therefore $T \restriction (\alpha + 1)$ satisfies (T2). As $T \restriction (\beta + 1)$ satisfies (T1), $T \restriction (\alpha + 1)$ satisfies (T1), too. To obtain $\pi_{\alpha+1}$, we extend π_α arbitrarily such that it satisfies the condition ($\pi 0$).

Let α be limit. For each $\beta < \alpha$, suppose $T \restriction \beta$ and π_β are defined and they satisfy the conditions mentioned above. We need to distinguish two cases. First, if α has cofinality less than κ then we add all possible sequences. We can do that since $\kappa^{<\kappa} = \kappa$.

Second, if α has cofinality κ then let $T'_\alpha = \bigcup_{\beta<\alpha} T_\beta$ and $\pi^*_\alpha = \bigcup_{\beta<\alpha} \pi_\beta$. We construct for each $s \in T'_\alpha$ and $x \in [\kappa \setminus \mathrm{Rng}(s)]^{<\kappa}$ node s_x above s of height α such that $x \cap \mathrm{Rng}(s) = \emptyset$. Let us fix for the rest of the proof a bijection g from κ to $\mathbb{Q}_\kappa$. Again, we need to distinguish two cases. First, if $f_\alpha \circ g$ embeds $\pi^{*\prime\prime}_\alpha T'_\alpha$ to $\mathbb{Q}_\kappa$ and $\mathrm{Dom}(f_\alpha) = \pi^{*\prime\prime}_\alpha T'_\alpha$, then set

$$X_\alpha = \{(s, x) | s \in T'_\alpha \;\&\; x \in [\kappa]^{<\kappa} \;\&\; \mathrm{Rng}(s) \cap x = \emptyset\}. \tag{3.12}$$

For (s,x), (t,y) in X_α, we define $(s,x) \leq_\alpha (t,y)$ if and only if $s \subseteq t$ and $x \subseteq y$. For each $q \in \mathbb{Q}_\kappa$, set

$$\begin{aligned} \Delta^\alpha_q = \big\{(s,x) \in X_\alpha \,|\, & g(f_\alpha(\pi^*_\alpha(s))) \geq_{\mathbb{Q}} q \text{ or} \\ & (\forall (t,y) \in X_\alpha)((t,y) \geq_\alpha (s,x) \to g(f_\alpha(\pi^*_\alpha(t))) <_{\mathbb{Q}_\kappa} q)\big\}. \end{aligned} \tag{3.13}$$

It is easy to see that Δ^α_q is cofinal in X_α.

Let $s \in T'_\alpha$ and $x \in [\kappa \setminus \mathrm{Rng}(s)]^{<\kappa}$ be given. First we fix an increasing sequence $\langle \alpha_\gamma | \gamma < \kappa\rangle$ with limit α and $\alpha_0 = \mathit{length}(s)$. By induction we construct an increasing sequence $\langle (s_\gamma, x_\gamma)\Delta^\alpha_{g(\gamma)} | \beta < \kappa\rangle$ with $\mathit{length}(s_\gamma) \geq \alpha_\gamma$ for all $\gamma < \kappa$.

Let $s'_0 = s$ and $x'_0 = x$. By definition of X_α, (s'_0, x'_0) is in X_α and as $\Delta^\alpha_{g(0)}$ is cofinal in X_α, we can find $(s_0, x_0) \geq_\alpha (s'_0, x'_0)$ in $\Delta^\alpha_{g(0)}$.

If $\gamma < \kappa$ is a successor ordinal $\gamma = \beta + 1$ we can proceed as follows. Assume (s_β, x_β) is defined. By (T1) there is $v_\beta \in \kappa \setminus (\mathrm{Rng}(s_\beta) \cup x_\beta)$. Let $x'_{\beta+1} = x_\beta \cup \{v_\beta\}$. By (T2) we can find $s'_{\beta+1} \in T'_\alpha$ such that $s'_{\beta+1} \supseteq s_\beta$, $\mathit{length}(s'_{\beta+1}) \geq \alpha_{\beta+1}$ and $\mathrm{Rng}(s'_{\beta+1}) \cap x'_\beta = \emptyset$. By definition of X_α, (s'_β, x'_β) is in X_α and as $\Delta^\alpha_{g(\beta)}$ is cofinal in X_α, we can find $(s_{\beta+1}, x_{\beta+1}) \geq_\alpha (s'_{\beta+1}, x'_{\beta+1})$ in $\Delta^\alpha_{g(\beta+1)}$.

Let $\gamma < \kappa$ be limit. Since $\gamma < \kappa$ we can take $s''_\gamma = \bigcup_{\beta<\gamma} s_\beta$ and $x'_\gamma = \bigcup_{\beta<\gamma} x_\beta$. As κ is regular, $|x'_\gamma| < \kappa$. Note that $\mathrm{Rng}(s''_\gamma) \cap x'_\gamma = \emptyset$, but $\mathit{length}(s''_\gamma)$ does not have to be greater or equal to α_γ. However, by (T2) there exists $s'_\gamma \supseteq s''_\gamma$ such that $\mathrm{Rng}(s'_\gamma) \cap x'_\gamma = \emptyset$ and $\mathit{length}(s'_\gamma) \geq \alpha_\gamma$. By definition of X_α, (s'_γ, x'_γ) is in X_α and as $\Delta^\alpha_{g(\gamma)}$ is cofinal in X_α, we can find $(s_\gamma, x_\gamma) \geq_\alpha (s'_\gamma, x'_\gamma)$ in $\Delta^\alpha_{g(\gamma)}$.

In the other case, if $f_\alpha \circ g$ does not embed $\pi^{*\prime\prime}_\alpha T'_\alpha$ to $\mathbb{Q}_\kappa$, then we proceed similar as before. Let $s \in T'_\alpha$, $x \in [\kappa \setminus \mathrm{Rng}(s)]^{<\kappa}$ and $\langle \alpha_\gamma | \gamma < \kappa\rangle$ be cofinal in α with $\alpha_0 = \mathit{length}(s)$. By induction we construct an increasing sequence $\langle (s_\gamma, x_\gamma) | \beta < \kappa\rangle$ with $\mathit{length}(s_\gamma) \geq \alpha_\gamma$ for all $\gamma < \kappa$.

Let $s_0 = s$ and $x_0 = x$.

If $\gamma < \kappa$ is a successor ordinal $\gamma = \beta + 1$ we can proceed as follows. Assume (s_β, x_β) is defined. By (T1) there is $v_\beta \in \kappa \setminus (\mathrm{Rng}(s) \cup x_\beta)$. Let $x_{\beta+1} = x_\beta \cup \{v_\beta\}$. By (T2) we can find $s_{\beta+1} \in T'_\alpha$ such that $s_{\beta+1} \supseteq s_\beta$, $\mathit{length}(s_{\beta+1}) \geq \alpha_{\beta+1}$ and $\mathrm{Rng}(s_{\beta+1}) \cap x_{\beta+1} = \emptyset$.

Let $\gamma < \kappa$ be limit. Since the size of γ is less than κ, we can take $s'_\gamma = \bigcup_{\beta<\gamma} s_\beta$ and $x_\gamma = \bigcup_{\beta<\gamma} x_\beta$. As κ is regular, $|x_\gamma| < \kappa$. Note that $\mathrm{Rng}(s'_\gamma) \cap x_\gamma = \emptyset$, but $\mathit{length}(s'_\gamma)$ does not have to be greater or equal to α_γ. However by (T2) there exist $s_\gamma \supseteq s'_\gamma$ such that $\mathrm{Rng}(s_\gamma) \cap x_\gamma = \emptyset$ and $\mathit{length}(s_\gamma) \geq \alpha_\gamma$.

Let $s_x = \bigcup_{\gamma<\kappa} s_\gamma$. We define the level $T_\alpha = \{s_x | s \in T'_\alpha \text{ and } x \in [\kappa \setminus \mathrm{Rng}(s)]^{<\kappa}\}$. It is easy to verify that $T \upharpoonright (\alpha+1) = T'_\alpha \cup T_\alpha$ satisfies the condition (T1) and (T2). Again, we can extend π^*_α to $\pi_{\alpha+1}$ on $T \upharpoonright (\alpha+1)$ arbitrarily such that it satisfies the condition (π0).

Finally, set $T = \bigcup_{\alpha<\kappa^+} T_\alpha$ and $\pi = \bigcup_{\alpha<\kappa^+} \pi_\alpha$. Then $\pi : T \to \kappa^+$ is a function such that $s \subseteq t \to \pi(s) < \pi(t)$.

For a contradiction assume that T is special. As we assume $\kappa^{<\kappa} = \kappa$, by Lemma 3.12 T is special if and only if T is $\mathbb{Q}_\kappa$-embeddable. Therefore there is a function $f : \kappa^+ \to \kappa$ such that $f \circ g$ embeds $\pi'' T$ in $\mathbb{Q}$. Let

$$\begin{aligned} C = \{\alpha < \kappa^+ \mid\ & \alpha \text{ is a limit ordinal and } \pi''(T \restriction \alpha) = \pi^{*}_{\alpha}{}'' T'_\alpha \text{ and} \\ & f \circ g \restriction \alpha \text{ embeds } \pi^{*}_{\alpha}{}'' T'_\alpha \text{ in } \mathbb{Q}_\kappa \text{ and} \\ & (\forall s \in T'_\alpha)(\forall x \in [\kappa \setminus \mathrm{Rng}(s)]^{<\kappa})(\forall q >_{\mathbb{Q}} g(f(\pi(s))) \\ & ((\exists t \in T)(t \supseteq s \,\&\, \mathrm{Rng}(t) \cap x = \varnothing \,\&\, g(f(\pi(t))) \geq_{\mathbb{Q}} q) \\ & \to (\exists t' \in T'_\alpha)(t' \supseteq s \,\&\, \mathrm{Rng}(t') \cap x = \varnothing \,\&\, g(f(\pi(t'))) \geq_{\mathbb{Q}} q)\Big\}. \end{aligned} \tag{3.14}$$

It is easy to verify that C is a closed unbounded subset of κ^+. As we assume $\diamondsuit_\kappa(E^{\kappa^+}_\kappa)$, the set $\{\alpha \in E^{\kappa^+}_\kappa | f \restriction \alpha = f_\alpha\}$ is stationary, so there is $\alpha \in C$ such that $f \restriction \alpha = f_\alpha$ and α has cofinality κ. Let $t \in T_\alpha$ and let $q = g(f(\pi(t)))$. By the construction of T, there is $(s, x) \in \Delta^\alpha_q$ such that $\mathrm{Rng}(s) \cap x = \varnothing$ and $s \subset t$. Since $f \circ g$, and π are order-preserving, $g(f(\pi(s))) <_{\mathbb{Q}} g(f(\pi(t))) = q$.

Since $g(f(\pi(s))) <_{\mathbb{Q}} q$ and $g(f(\pi(t))) \geq_{\mathbb{Q}} q$, by the definition of C there exists $t' \in T'_\alpha$ such that $t' \supseteq s$, $\mathrm{Rng}(t') \cap x = \varnothing$ and $g(f(\pi(t'))) \geq_{\mathbb{Q}} q$. Note that (s, x), (t', x) are in X_α and $(s, x) \leq_\alpha (t', x)$. Since (s, x) is in Δ^α_q and $f \restriction \alpha = f_\alpha$, by (3.13) it must hold that $g(f_\alpha(\pi(s))) \geq_{\mathbb{Q}} q$. But $f_\alpha = f \restriction \alpha$ and so $g(f(\pi(s))) \geq_{\mathbb{Q}} q$. This contradicts our earlier inequality $g(f(\pi(s))) <_{\mathbb{Q}} q$. □

Corollary 3.28. *Assume $\kappa^{<\kappa} = \kappa$ and $\diamondsuit_{\kappa^+}(E^{\kappa^+}_\kappa)$. Then there is an $\mathbb{R}_\kappa$-embeddable κ^+-Aronszajn tree, which is not special.*

Proof. By Lemma 3.17, every M-special κ^+-Aronszajn tree is $\mathbb{R}_\kappa$-embeddable. □

Corollary 3.29. *Assume $\kappa^{<\kappa} = \kappa$ and $\diamondsuit_{\kappa^+}(E^{\kappa^+}_\kappa)$. Then there is an $\mathcal{S}$-special κ^+-Aronszajn tree, which is not special.*

Proof. By Lemma 3.16, every M-special κ^+-Aronszajn tree is $\mathcal{S}$-special for $S = \{\alpha + 1 | \alpha < \kappa^+\}$. □

The next lemma is a straightforward generalisation of Lemma 2.16 and tells us that the last inclusion in (3.10) can be consistently proper.

Lemma 3.30. *Assume $\kappa^{<\kappa} = \kappa$ and $\diamondsuit_{\kappa^+}(E^{\kappa^+}_\kappa)$. Then there is a κ^+-Aronszajn tree, which is S-special for some S unbounded subset of κ^+ and it is not* M*-special and by our assumption it is not $\mathbb{R}_\kappa$-embeddable.*

Proof. The proof is the same as in Lemma 2.16. □

To show that the second inclusion in (3.10) can be consistently proper, i.e. that $\mathbb{A}^{\mathcal{S}\text{-}sp} \neq \mathbb{A}^{NS}$, we need to introduce the notion of an ω-ascent path, which is due to Laver.

Definition 3.31. Let κ be a regular cardinal. We say that a κ^+-Aronszajn tree T has the property of the *ω-ascent path* if there is a sequence $\langle x^\alpha | \alpha < \kappa^+\rangle$ such that

(i) for each $\alpha < \kappa^+$, x^α is a function from ω to T_α;
(ii) if $\alpha, \beta < \kappa$ with $\alpha < \beta$ then $\exists n \in \omega\, \forall m \geq n\ x^\alpha_m < x^\beta_m$.

If the tree T has a cofinal branch, then this branch is a 1-ascent path and it is obvious that T is not special. But Aronszajn trees do not have cofinal branches. Thus an ω-ascent path is a pseudo-branch with width ω which prevents the tree from being special.

The following fact is due to Shelah ([SS88]), building on work of Laver and Todorčević.

Fact 3.32. *Let $\kappa > \omega$ be a regular cardinal. Let T be a κ^+-Aronszajn tree with the property of an ω-ascent path. Then T is not special.*

Remark 3.33. No such argument can exist for ω_1-trees since it is important for the proof that there is a regular cardinal between ω and κ^+. This is the difference between the specialization forcing for ω_1 and for higher cardinals. In the case of higher cardinals, if T has an ω-ascent path, then any specialization forcing must collapse cardinals. On the other hand, as was pointed out by a referee, Baumgartner showed that an ω_1-tree has a cofinal branch if and only if it contains an ascent path of finite width. In particular, the nonexistence of paths of finite width implies that the corresponding specialization forcing has the ccc.

Corollary 3.34. *Let κ be a regular cardinal. Let T be a κ^+-Aronszajn tree with the property of an ω-ascent path. Then T is not $\mathcal{S}$-special.*

Proof. Let $S \subseteq \kappa^+$ be an unbounded subset of κ^+ and $\langle x^\alpha | \alpha < \kappa^+ \rangle$ be an ω-ascent path. Then $\langle x^\alpha | \alpha < \kappa^+ \rangle \restriction S$ is ω-ascent path for $T \restriction S$ and by the previous theorem $T \restriction S$ is not special. $\square$

The construction of the following tree can be found in [SS88].[8]

Fact 3.35. *Let κ be a regular cardinal. Assume $\square_\kappa$. Then there is a non-Suslin κ^+-Aronszajn tree with ω-ascent path.*

Hence we can conclude that the second inclusion in (3.10) can be consistently proper.

Corollary 3.36. *Let κ be a regular cardinal. Assume $\square_\kappa$. Then there is a non-Suslin κ^+-Aronszajn tree T such that T is not $\mathcal{S}$-special.*

Proof. It follows from Corollary 3.34 and Fact 3.35. $\square$

References

[Bau70] James Baumgartner. Decomposition and embeddings of trees. *Notices of the American Mathematical Society*, 17, 1970.

[BMR70] James Baumgartner, Jerome Malitz, and William Reinhardt. Embedding Trees in the Rationals. *Proceedings of the National Academy of Sciences*, 67(4): 1748–1753, 1970.

[Dev72] Keith Devlin. Note on a theorem of J. Baumgartner. *Fundamenta Mathematicae*, 76(3): 255–260, 1972.

[DJ74] Keith Devlin and Håvard Johnsbråten. *The Souslin Problem*. Lecture Notes in Mathematics. Springer-Verlag, Berlin, third edition, 1974. The third millennium edition, revised and expanded.

[Han81] Masazumi Hanazawa. Various Kinds of Aronszajn Tree with No Subtree of a Different Kind. *Lecture Notes in Mathematics*, 891: 1–21, 1981.

8 It is worth noting that Todorčević also constructed Aronszajn trees with ascent paths from weaker assumptions, see [Tod89].

[Jec03] Thomas Jech. *Set Theory*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 2003.

[Jen72] Ronald Jensen. The Fine Structure of The Constructible Hierarchy. *Annals of Mathematical Logic*, 4: 229–308, 1972.

[Kur35] Djuro R. Kurepa. Ensembles ordonnés et ramifiés. *Publications mathématiques de l'Université de Belgrade*, 4: 1–138, 1935.

[Mit72] William Mitchell. Aronszajn Trees and Independence of the Transfer Property. *Annals of Mathematical Logic*, 5: 21–46, 1972.

[Sch14] Chaz Schlindwein. How Special is your Aronszajn Tree? http://arxiv.org/pdf/math/0312445v1.pdf, 2014.

[She98] Saharon Shelah. *Proper and Improper Forcing*. Perspectives in Mathematical Logic. Springer-Verlag, Berlin, second edition, 1998.

[Spe49] Ernst Specker. Sur un problème de Sikorski. *Colloquium Mathematicum*, 2: 9–12, 1949.

[SS88] Saharon Shelah and Lee Stanley. Weakly Compact Cardinals and Nonspecial Aronszajn Trees. *Proceedings of the American Mathematical Society*, 104(3): 887–897, 1988.

[Tod84] Stevo Todorčević. Trees and linearly ordered sets. In Kenneth Kunen and Jerry Vaughan, editors, *Handbook of Set-Theoretic Topology*. Elsevier Science Publisher B.V., 1984.

[Tod89] Stevo Todorčević. Special square sequences. *Proc. Amer. Math. Soc.*, 105(1): 199–205, 1989.

2015 ACTA UNIVERSITATIS CAROLINAE PAG. 93–101
PHILOSOPHICA ET HISTORICA 1 / MISCELLANEA LOGICA X

DYNAMICAL BRANCHING

PETR ŠVARNÝ*
Dept. of Logic, Faculty of Arts, Charles University in Prague
* E-mail: svarnypetr@gmail.com

ABSTRACT
We investigate the Branching space-times logic in a dynamic setting. We present a new logic, called Dynamic branching logic and show some of its basic properties.
Keywords: branching time, branching space-times, dynamic logic

1. Introduction

Branching temporal logics, those originating from Branching space-times (BST), are attempting to model time. Although branching should capture the changes of possibilities, the point-events that constitute these branches are given in advance and in a tenseless manner. Common view of time and changes, however, is more connected to dynamic evolution of options, possibilities, and the future. For this reason, we find it suitable to try a synthesis of dynamic and branching logic.[1] We will shortly discuss the philosophical background of our work, present the original BST approach, and then present a dynamical branching time model.

2. Philosophical background

How does our work approach the questions of eternalism or presentism, determinism or indeterminism?

Let us define the terms clearly:

> *Eternalism* is committed to the tenseless coexistence of all events, while *presentism* is committed to the thesis that existence is confined just to the present events, while future and past events do not exist.
>
> (Dorato, 2012)

From these definition stems also our motivation. We hold the view in this paper that a present event can give rise to the following event by a dynamic process and this would be a presentist model, as opposed to the static eternalist model.

---

[1] The idea to approach branching as a dynamic process was suggested to the author by Ondrej Majer at a session of the 'Prague dynamic group' (O. Majer, M. Peliš and others).

The idea of determinism is more complex than it might seem at first sight and we can differentiate between approaches based on what exactly we mean by determinism (Müller & Placek, 2015). Nevertheless, we use the following definition of determinism: "given the state of the world now, there is only one possible future outcome". Thus, indeterminism means that given the present state we have multiple possible future outcomes.

Our theory would like to be a presentist indeterministic theory. Therefore it would allow multiple possible outcomes from present events without committing to the existence of any future or past events. Now, as we do not wish to ignore (entirely) advances made in the field of physics, we need to address the question of 'present'. The arrival of Special Theory of Relativity (STR) marked also the end of a simple idea of present or simultaneity[2] (Dieks, 1988). It showed how closely time and space are interwoven. We do not want to ignore this and therefore get inspiration from the Machian view of Barbour (Barbour, 2000) that is consistent with STR. While BST uses the notion of space-time fully, we will use a space and time approach. This dichotomy could be viewed as a version of the endurantism and perdurantism discussion[3] or it can be considered simply as a different formulations of eternalism and presentism (Dorato, 2012). Instead of taking entities as entities with four dimensions, three spatial and one temporal, as seen in STR or BST, we consider entities to be only three dimensional. Therefore our basic building blocks are not space-time events, but 3D configurations of 3D entities. We could, in order to accommodate the results of STR, discuss the role of observers in these configurations. However, at this point, we focus merely on the introduction of some ideas how such presentist approach could start out and do not venture deeper into this topic. An observer-related[4] dynamics was discussed already for BT. The so called BT+I+AC[5] structures are introduced in (Belnap, Perloff, & Xu, 2001). Nevertheless, these structures focus on the establishment of an agents context and not on dynamic changes and therefore they stay out of our focus in this paper. Mentioning another related work, Müller introduced transitions, in the basic case transitions from a point-event to its immediate outcomes, as a framework to investigate causation and probability theory. The approach uses consistency as a way how to define causal possibility. Still, transitions are again built on the usual BST foundations and transitions are more like tool choosing selected point-events rather than dynamically constructing them.

3. Branching space-times

BST is important for us to keep in mind as a reference theory whose expressiveness we would like to match but do so in a different way. Let us have a look at how time and possibilities are treated there.

---

2 The notion that two spatially separated events happen at the same time.

3 Endurantism expresses the view that objects are 3D material objects wholly present at every moment of their existence. Perdurantism is the somewhat opposite view that objects are 4D objects whose existence extends over a period of time, i.e. the objects have different temporal parts.

4 Precisely an agent-related.

5 Branching time + instants + agents and choices.

Branching space-times were introduced by Belnap in (Belnap, 1992). However, we use the concise formulation from Placek and Wroński article (Placek & Wroński, 2009). The basic definitions connected to BST follow.

Definition 1 (Placek & Wroński, 2009)

The set OW called Our World , is composed of point-events e ordered by $\leq$.

A set $h \subseteq OW$ is upward-directed iff $\forall e_1, e_2 \in h\ \exists e \in h$ such that $e_1 \leq e$ and $e_2 \leq e$.

A set h is maximal with respect to the above property iff $\forall g \in OW$ such that $h \subset g$, g is not upward-directed.

A subset h of OW is a history iff it is a maximal upward-directed set.

For histories h_1 and h_2, any maximal element in $h_1 \cap h_2$ is called a choice point for h_1 and h_2.

Histories are meant to capture the familiar notion of possible courses of events. Hence if some event e occurred in one course of events, it is inconsistent with a different event e', its counterpart from a different course of events. We see a history can be a very large set as it would contain a series of point-events from the 'origin' of time until its 'end'. The quotation marks remind us that a model of BST can be without an origin and an end too.

Definition 2 (Placek & Wroński, 2009)

$\langle OW, \leq \rangle$ where OW is a nonempty set and $\leq$ is a partial ordering on OW is a structure of BST iff it meets the following requirements:

(1) The ordering $\leq$ is dense.
(2) $\leq$ has no maximal elements.
(3) Every lower bounded chain in OW has an infimum in OW.
(4) Every upper bounded chain in OW has a supremum in every history that contains it.
(5) (Prior choice principle) For any lower bounded chain $O \subset h_1 - h_2$ there exists a point $e \in OW$ such that e is maximal in $h_1 \cap h_2$ and $\forall e' \in O(e < e')$.

We can see the steps necessary to befriend branching and space-times. The notion of histories and their treatment is reminiscent of earlier branching temporal logic, namely computational tree logic and its fullpaths (Hodkinson & Reynolds, 2006). We see this approach as an eternalist one, assuming we have all point-events in OW and these eternally coexist. A different, not necessarily eternalist reading of BST was presented in (Pooley, 2013), where a non-standard non-eternalist interpretation of BST is presented as the "most promising way to reconcile becoming with relativity".

The language of Dynamic branching logic will use Priorean operators F, P (The meanings are as usual. In the order of the operators 'it will be true', 'it was true'.) and therefore we should also address the question, how are these interpreted in BST. We add one new operator, the '$Sett:$' operator denotes a settled option, i.e. true for all the branches. The $Sett:$ operator allows us to basically describe the lack of options - in all courses of events, the statement will be true (for example if I want to capture the statement "no matter what I do, I will have some typos in this paper").

Definition 3 Point satisfies formula – BST (Placek & Wroński, 2009)

For the model $\mathfrak{M} = \langle OW, \leq, v \rangle$. Where v is the valuation $v :$ Atoms $\rightarrow \mathcal{P}(OW)$. For a given event e and history h, such that $e \in h$:

$$
\begin{array}{llll}
\mathfrak{M}, e, h & \Vdash & p & \text{iff } e \in v(p) \\
\mathfrak{M}, e, h & \Vdash & \neg\varphi & \text{iff not } \mathfrak{M}, e, h \Vdash \varphi \\
\mathfrak{M}, e, h & \Vdash & \varphi \wedge \psi & \text{iff } \mathfrak{M}, e, h \Vdash \varphi \text{ and } \mathfrak{M}, e, h \Vdash \psi \\
\mathfrak{M}, e, h & \Vdash & F\varphi & \text{iff there is } e' \in OW \text{ and } e^* \in h \text{ s.t.} \\
 & & & \quad e' \leq e^* \text{ and } \mathfrak{M}, e', h \Vdash \varphi \\
\mathfrak{M}, e, h & \Vdash & P\varphi & \text{iff there is an } e' \in h \text{ s.t. } e' \leq e \text{ and } \mathfrak{M}, e', h \Vdash \varphi \\
\mathfrak{M}, e, h & \Vdash & Sett : \varphi & \text{iff for all } e' \in h', \text{ for all } h' \text{ such that } e \in h': \\
 & & & \quad M, e', h' \Vdash \varphi
\end{array}
$$

A short note about the future operator, although the expression does not contain the event e, it does contain the event $e*$ from the history h to which also e belongs. Therefore one could rephrase the expression as 'at a future event to e, namely $e*$, we will be able to say that φ is true'. We cannot, however, just pick event e' as it does not have to be necessarily in h and it might be that φ does not hold at $e*$ (i.e. it is not a settled future).

4. Dynamic branching logic

Dynamic branching logic (DBL) will attempt to model a presentist theory with indeterminism. Based on the definitions mentioned earlier, DBL is presentist in the sense that only a set of events (i.e. present events) is considered and possible future (or past) is constructed based only on these present events. Concerning indeterminism, DBL allows for present events to have multiple possible future outcomes (e.g. quantum observations). We present a syntactic and a semantic approach.

4.1 Syntax

Language of DBL is the following atomic statements ($a, b, c, \ldots$), truth and falsity ($\top$, $\bot$), logical connectives ($\wedge$), negation ($\neg$), temporal operators (F, P, G, H), actions ([C], [R], $[\varphi!]$), and the branching modality $Sett$:.

The atomic statements are assumed to be in present tense and simple, for example: 'The doctor is a time-traveling alien.' The temporal operators can be read as it is usual, i.e. $F\varphi$ stands for 'it will be true in the future that φ' and $P\varphi$ stands for 'it was true in the past that φ'. The temporal operators G and H represent then necessary temporal statements, i.e. $G\varphi$ means 'it will always be true that φ'.

DBL has three types of actions, [C] is construction, i.e. creation of the future, [R] is reconstruction, i.e. recreation of the past, and $[\varphi!]$ is claiming that a statement φ is true or becoming true (similar to public announcement from dynamic epistemic logics). The creation of the future and recreation of the past actions are inspired by Barbour's time capsules. The main idea can be summed up in the following way. Every state $s \in S$ contains evidence for the events that were in the past of the state (e.g. fossils, broken glass, contrail). This evidence allows us to reconstruct the events that lead to the given state and determine what possible states preceded the given state s. In a similar manner, the current state of affairs s determines (by physical laws) all the possible states. Therefore, based merely on our current state n, we can make a judgment about our past and future. The actions [C] and [R] then represent an ontological construction, not an epistemic one, from the given state. Because we do not have at this time a more detailed account of the states, the actions can be taken as functions from a state to states.

Because DBL aims to be also a model of branching, the idea of a statement being settled, i.e. true for all branches, is a useful one. We denote 'φ is settled' as $Sett : \varphi$.

The relation between the temporal operators is the following:

(1) $$H\varphi \equiv \neg P \neg \varphi, G\varphi \equiv \neg F \neg \varphi$$

We can then look for axioms of our logic in well known modal logics and their temporal versions. As basis, we can take the minimal tense logic K_t (Hodkinson & Reynolds, 2006) (Goranko & Galton, 2015).

Axiom 4

The axioms for BTL are:

1. all propositional tautologies
2. $G(\varphi \to \psi) \to (G\varphi \to G\psi)$ and $H(\varphi \to \psi) \to (H\varphi \to H\psi)$
3. $G\varphi \to GG\varphi$ and $H\varphi \to HH\varphi$
4. $\varphi \to GP\varphi$
5. $\varphi \to HF\varphi$
6. $[\alpha](\varphi \to \psi) \to ([\alpha]\varphi \to [\alpha]\psi)$

Where α can be any action.

And the rules:

1. Modus Ponens $\varphi, \varphi \to \psi \vdash \psi$
2. $\varphi / G\varphi$
3. $\varphi / H\varphi$
4. $\varphi / [\alpha]\varphi$
5. Substitution in any propositional tautology.

4.2 Semantics

We present a suggestion for DBL semantics. We need to first realize a DBL model represents a configuration of three dimensional objects, entities existing concurrently. Actions then allow us to construct a new model of a new instant discarding the old one. The model always represents the current status of temporal relations. Every model contains in itself also information about the future or past events. For example a model representing a configuration of the first of January 2015 will contain in itself also books describing past events, diaries of people, video tapes, or geological sediments. The configuration does also contain many predispositions for future events, for example an emitted electromagnetic signal or a sent letter. This approach is reminiscent of Barbour's time capsules (Barbour, 2000). We look at time as a static series of 3D images containing no time by their own virtue, but they do contain a lot of evidence connected to processes connected to time. Remember, however, neither past nor future exist. Merely the hints, predispositions, and conditions for the future are actualized.

The usual terminology speaks about states and this term does not part too much from our idea of a configuration. The main change is that we do not have point-events as our basic building blocks but states (configurations) of 3D objects. Therefore a state s truly represents all that is in the chosen 'snapshot' of the universe. It is not merely one point-event.

Definition 5
The relation $\prec$ is defined on S as the *causal ordering* of states.

We want to maintain the idea of branching similar to the one from BST and therefore we introduce also histories. However, these histories can be considered closer to the original histories from Branching time.
The relation $\prec$ is basically the usual accessibility relation. However, in our temporal context it is interpreted as the causal relation between states.

Definition 6
A frame $\mathcal{F} = \langle S, \prec \rangle$ is the *frame for DBL* iff
(1) $\prec$ is transitive
(2) there exists a state $n \in S$ such that it holds for all s from S that $s = n$ or there are $s_1, ..., s_k \in S$ such that $s_1 = s$ and $s_k = n$ and either $s_1 \prec s_2 \prec ... \prec s_k$ or $s_k \prec s_{k-1} \prec ... \prec s_1$
(3) for all $s_{past} \in S$ such that $s_{past} \prec n$ and for all $s_{future} \in S$ such that $n \prec s_{future}$ it holds that $s_{past} \prec s_{future}$

Hence the state n represents a connection point between the past and the future.

Definition 7
A (dynamic) *history* is the set $h \subset S$ such that $n \in h$ and for all $s_1, s_2 \in h$ either $s_1 = s_2$, $s_1 \prec s_2$ or $s_2 \prec s_1$.

Definition 8
A *choice state* between distinct histories $h_1, h_2 \subset \mathcal{F}$ is the maximal state in the intersection of h_1 and h_2.

We can notice that the minimal possible choice state for two distinct histories is the special state n.
The Definition 6 forces every model of DBL to have its states connected in some way to the present. They are either the consequences of the present state or the causes for it. This is related to our earlier claim that all the future and past states are just representations of information contained in the present state.

Definition 9
The model $\mathfrak{M}$ is the pair $\langle \mathcal{F}, v \rangle$ with the frame $\mathcal{F}$ and the valuation of atomic formulas v.

Similarly as in other dynamic logics, we will change the model in course of the evaluation of formulas. The three possible actions present three possible ways how to do so.

The construction C creates from the model $\mathfrak{M}$ a temporally successive model. Reconstruction R, on the other hand, presents a model that is temporally preceding our current model $\mathfrak{M}$. Both steps make use of the temporal operators in the current model.

Definition 10

We define the model $\mathfrak{M} \mid \mathrm{C}$ for a model $\mathfrak{M}$, as the model where we choose one of the $s \in S$, s.t. $n \prec s$, as the new now, i.e. $n_{\mathrm{C}} = s$. The new model contains only histories that contain n_{C}.

Definition 11

We define the model $\mathfrak{M} \mid \mathrm{R}$ for a model $\mathfrak{M}$, as the model where we choose one of the $s \in S$, s.t. $s \prec n$, as the new now, i.e. $n_{\mathrm{R}} = s$. The new model contains all the histories that contain n_{C}.

Definition 12

We define the model $\mathfrak{M} \mid \varphi$ for a model $\mathfrak{M}$, state s, as the model where for every h it holds that $\mathfrak{M}, s, h \Vdash Sett : \varphi$.

Definition 13 Model satisfies formula - DBL

For the model $\mathfrak{M} = \langle S, \leq, v \rangle$. Where v is the valuation $v : \text{Atoms} \to \mathcal{P}(S)$. For a given state s and history h, st. $s \in h$:

$$
\begin{array}{lll}
\mathfrak{M}, s, h & \Vdash p & \text{iff } s \in v(p) \\
\mathfrak{M}, s, h & \Vdash \neg\varphi & \text{iff not } \mathfrak{M}, s, h \Vdash \varphi \\
\mathfrak{M}, s, h & \Vdash \varphi \wedge \psi & \text{iff } \mathfrak{M}, s, h \Vdash \varphi \text{ and } \mathfrak{M}, s, h \Vdash \psi \\
\mathfrak{M}, s, h & \Vdash F\varphi & \text{iff there is } s' \in S \text{ s.t. } s \prec s' \text{ and } \mathfrak{M}, s, h \Vdash \varphi \\
\mathfrak{M}, s, h & \Vdash P\varphi & \text{iff there is an } s' \in S \text{ s.t. } s' \prec s \text{ and } \mathfrak{M}, s, h \Vdash \varphi \\
\mathfrak{M}, s, h & \Vdash G\varphi & \text{iff for all } s' \in S \text{ s.t. } s \prec s' \text{ it holds } \mathfrak{M}, s, h \Vdash \varphi \\
\mathfrak{M}, s, h & \Vdash H\varphi & \text{iff for all } s' \in S \text{ s.t. } s' \prec s \text{ it holds } \mathfrak{M}, s, h \Vdash \varphi \\
\mathfrak{M}, s, h & \Vdash Sett : \varphi & \text{iff there is } s' \in S \text{ such that } s \prec s' \text{, for all } h' \text{ that contain } s' \\
& & \quad \text{it holds } \mathfrak{M}, s', h' \Vdash \varphi \\
\mathfrak{M}, s, h & \Vdash [\mathfrak{C}]\varphi & \text{iff there is a model } \mathfrak{M} \mid \mathrm{C} \text{ such that } \mathfrak{M} \mid \mathfrak{C}, s, h \Vdash \varphi \\
\mathfrak{M}, s, h & \Vdash [\mathfrak{R}]\varphi & \text{iff for the model } \mathfrak{M} \mid \mathrm{R} \text{ it holds that } \mathfrak{M} \mid \mathfrak{R}, s, h \Vdash \varphi \\
\mathfrak{M}, s, h & \Vdash [\psi]\varphi & \text{iff there is a model } \mathfrak{M} \mid \psi \text{ such that } \mathfrak{M} \mid \psi, s, h \Vdash \varphi
\end{array}
$$

We see therefore that semantics for DBL can work in a very similar fashion as those of BST, because it maintains the same branching structure. However, it adds the element of changing a model with actions. We can close our account with two straightforward theorems[6].

Theorem 14

There exists a BST structure that is not a DBL structure.

Proof. In order for a BST structure to be a DBL structure, we need a to fulfill the definition 6. For every BST structure it holds that $\leq$ is transitive because it is a partial ordering on OW. Nevertheless, the second condition does not hold necessarily. Let us have a BST structure with at least two $e_1, e_2 \in OW$ such that $e_1 \not\leq e_2$ $e_2 \not\leq e_1$. Then if

6 The suggestion to enhance the paper with at least some of these theorems comes from the reviewer.

either one of these two point-events is n the second condition from the definition does not hold. Thus we have a structure of BST that is not a DBL structure.

We can also construct such a BST structure that for every e there will be a e' such that they are incomparable by $\leq$. Then at least for one e it will be true that $e = n$ and there will be an e' such that they are incomparable, i.e. it is not a DBL structure. □

This answer is not a big surprise if we realize that simultaneous point-events are a valid possibility on BST models. However, they are a crucial impossibility for DBL.

Theorem 15

There exists a DBL structure that is not a BST structure.

Proof. The relation $\prec$ does not have to be dense and this would contradict the first requirement of BST models. An example of a DBL model that has a discrete $\prec$ relation is the Platonia model presented by Barbour. □

5. Discussion

The comparison of DBL and BST is quite straight forward. We see that DBL attempts to capture the presentist idea by creating new models every time the state of the world changes in time (either progresses or regresses). BST on the other hand has a stable unchanging structure that was fixed at the moment of the first setting up of the model. Both models represent branches and hence an option for indeterminism, there can be states in DBL with multiple future options.

There is no big difference in the approach to branching. Already the original BST claimed clearly that there is no backward branching because it is 'plausible enough to warrant making clear what it comes to' (Belnap, 2003). DBL could in theory have backward branching. Nevertheless, for the same reason as BST, the models were chosen not to contain backward branches. For any event that would lead to two possible futures would have a unique reconstruction and therefore it seems plausible to assume there is only one past. Branching as it was familiar from BST is also present in DBL, however, it is contained in single models and does not represent a single model.

One might ask, why did we start out with BST and basically strip it of all its new features, just to end up with a strange branching time model. The answer is that we tried to keep in mind the fruits that came with BST and attempted to maintain them also in our model. This is visible on the fact, how DBL models are based on observers and do not represent some general branching of time.

6. Summary

We showed how Branching space-times approach the question of time and why they do so as an eternalist theory. We then presented a dynamic branching logic that presents a presentist alternative. The comparison of the two approaches showed that they should be equivalent in their expressive strength.

Further investigation could show whether the closeness of DBL to other modal logics gives it an advantage and allows it to prove (as opposed to BST) completeness and similar properties. Especially a deeper investigation should be done in relation to the more philosophical approach of Belnap (Belnap et al., 2001) and Müller (Müller, 2005).

Acknowledgments

Great thanks are due to the kind reviewer. Although not all of his comments were fully realized in this article, they were all a great contribution and will be part of future work. This output was created within the project 'Proměny koncepce filosofického vnímání', subproject FF_VG_2015_074 'Rozšíření modelů větvícího se času' solved at Charles University in Prague from the Specific university research in 2015.

References

Barbour, J. B. (2000). *The end of time: the next revolution in physics*. Oxford University Press.

Belnap, N. (1992). Branching Space-Time. *Synthese*, 3, 385–434.

Belnap, N. (2003). *Branching space-time*, postprint January, 2003. http://phisci-archive.pitt.edu/1003/.

Belnap, N., Perloff, M., & Xu, M. (2001). *Facing the Future: Agents and Choices in Our Indeterminist World*. Oxford University Press.

Dieks, D. (1988). Special relativity and the flow of time. *Philosophy of Science*, 456–460.

Dorato, M. (2012). Presentism/eternalism and endurantism/perdurantism: why the unsubstantiality of the first debate implies that of the second. *Philosophia naturalis*, 1, 25–41.

Goranko, V., & Galton, A. (2015). Temporal logic. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2015 ed.). http://plato.stanford.edu/archives/winter2015/entries/logic-temporal/.

Hodkinson, I., & Reynolds, M. (2006). Presentism and eternalism in perspective. *Handbook of modal logic*, 6, 655–720.

Müller, T. (2005). Probability theory and causation: A branching space-times analysis. *The British journal for the philosophy of science*, 3, 487–520.

Müller, T., & Placek, T. (2015). Defining determinism. *The British journal for the philosophy of science*, forthcoming.

Placek, T., & Wroński, L. (2009). On infinite EPR-like correlations. *Synthese*, 1, 1–32.

Pooley, O. (2013). Relativity, the Open Future, and the Passage of Time. *Proceedings of the Aristotelian Society*, 3 pt 3, 321–363.

ACTA UNIVERSITATIS CAROLINAE
PHILOSOPHICA ET HISTORICA 1/2015

MISCELLANEA LOGICA X

Editor: Michal Peliš
Cover and layout Kateřina Řezáčová
Published by Charles University in Prague
Karolinum Press, Ovocný trh 3–5, 116 36 Prague 1, Czech Republic
www.karolinum.cz
Prague 2016
Typeset by Lacerta (www.sazba.cz)
Printed by Karolinum Press
ISSN 0567-8293 (Print)
ISSN 2464-7055 (Online)
MK ČR E 18598

Distributed by Faculty of Arts, Charles University in Prague, nám. Jana Palacha 2,
116 38 Prague 1, Czech Republic (books@ff.cuni.cz)