Unterschrift des Betreuers

Technische Universität Wien
Vienna University of Technology

DIPLOMARBEIT

“Choice Sequences - Past and Future”

Thema

Ausgeführt am Institut für

Diskrete Mathematik und Geometrie

der Technischen Universität Wien

unter der Anleitung von

Ao.Univ.Prof. Dr.phil. Matthias Baaz

durch

Mag. Sarah Zobel

Name

Illekgasse 2/8, 1150 Wien

Anschrift

Wien, 20.03.2009

Datum

Unterschrift (Student)

“On est rien d’autre que la somme de ses choix.”
Albert Camus

Acknowledgements

It was quite a peculiar thing to write this thesis, since I actually had fun doing so. I suppose, being interested in the topic you write about is normally the best thing you can hope for. Fun during the writing process is then something really special. This is not to say that there weren't any bumps and breaks along the way. Therefore, I'm even more grateful to the people I want to thank now.

I would like to thank Matthias Baaz for permitting and taking an interest in my topic of choice. Our conversations were always lectures in disguise and I had more than one epiphany. I think, I seldomly learnt so much just by talking to someone.

For their moral support and friendship, I would like to thank Thomas Graf, Constanze Neumann, Viola Schmitt, Magda Schwager and Georg Seitz. I am quite sure, I made each of them suffer at one point or another for which I am very sorry.

And last but not least I want to thank my parents, Wolf and Bernadette Zobel. They are getting quite experienced in giving motivational speeches to help me pull through periods of inspirational draught. Thank you for helping me and supporting me in my wish to finish Mathematics and lead my life the way I do.

Contents

Introduction

Choice sequences are the central concept of intuitionistic mathematics. Every difference that can be pointed out between intuitionistic and classical mathematics is a direct consequence of putting choice sequences at the base of intuitionistic thinking. We will use this introduction to give a short general overview of the logical and historical background of choice sequences. During the main part of this thesis, however, we will focus on a special type of choice sequences, called *lawless sequences*.

Quite generally put, choice sequences are mappings from $\mathbb{N}$ into a set X where the elements or values of the sequence are "chosen" one after another from the elements in X. The action of choosing can be, e.g. completely restricted by a law (*lawlike sequence*) or completely random and unrestricted (*lawless sequence*). Between these two extremes, other types of choice sequences are conceivable. One such type, called *hesitant sequences*, will also be discussed briefly in the following chapter.

Note that no type of choice sequences is a class of static, completely given objects the way sequences are usually conceived in classical mathematics. They are processes of choosing (in some sense, depending on the type of sequence) at every moment a finite number of values such that only a finite number of values is known at each point.

Traditionally, choice sequences are taken as mappings from $\mathbb{N}$ into $\mathbb{N}$. Therefore, throughout this thesis, we will adopt this convention (Kreisel (1968), Troelstra (1977) and others).

Historically, at the beginning of the 20th century, one of L.E.J. Brouwer's concerns when he started to develop his intuitionistic philosophy of mathematics was to find an intuitionistically adequate way to formulate a theory of sets. The utilisation of choice sequences in this matter opened a way to grasp the intuition of the *continuum*, a concept that was heavily discussed at that time. Brouwer represented every real number by a choice sequence evolving in the mind of the mathematician, and he called the introduction of choice sequences the "second act of intuitionism".

The "first act of intuitionism" is to view mathematics as a subjective creation of the mind of a mathematician. An important issue accompanying this philosophical point of view is that properties of mathematical objects, since they are created by the mind of the mathematician, always need to be either provable or refutable, which is only possible if the verification of the property requires just a finite subpart of the mathematical object (Brouwer 1992:21ff). This is reflected in the perception of choice sequences as processes. Thus choice sequences are only *potentially infinite* (at very point only a finite initial segment is known) and are thus the prototypical *incomplete objects*.

How do the differences between intuitionistic and classical mathematics follow from this? The utilisation of choice sequences leads to a different way of "doing mathematics".

In classical mathematics we do not work directly with mathematical objects. Therefore, to gain new information about the objects of interest, new properties are proved from all the properties already known about the objects. More concretely: instead of working with the objects themselves, in classical mathematics we work with all properties known about the objects.

In some sense, classical mathematics tries to approximate the "objects of interest" from above. Obviously, these objects form a subset (or subclass) of all mathematical objects. However, because of the inability to refer to one of these objects directly, the right subset of objects is cut out from the set of all mathematical objects by listing all the properties the objects should have (this leads to the introduction of *axioms*) and provably have (properties obtained from proofs). Each new property specifies an even smaller subset of all mathematical objects. The sequence of all sets gained by adding new properties converges towards the set containing only the objects of interest.

In intuitionistic mathematics, we work with the objects directly by collecting information about the object one by one that not only specifies the object further but also the set of objects it belongs to. In this light, intuitionistic mathematics approximates a mathematical object from below.

The approximation from above is problematic for intuitionistic mathematics since it requires that for every object and every property it is provable that the object has or does not have this property. This requirement is of course the *law of the excluded middle*. As a consequence of the first act, the law of the excluded middle, however, does not hold generally intuitionistically. The perfect specification of a real number needs an infinite amount of information. Since we only ever have a finite amount of information about a mathematical object, some potential property might neither be provable nor refutable. The introduction of choice sequences as the basic mathematical objects as the second act is an answer to this problem.

In classical mathematics the assumption that we have access to the potentially infinite amount of information about a mathematical object is unproblematic in the sense that classical mathematics is based on classical logic for which the law of the excluded middle holds.

Summing up the above discussion, intuitionistic and classical mathematics differ in the treatment of (information about) mathematical objects. Ideologically they form the two extremes on a scale.

Approximation from above as in classical mathematics and approximation from below as in intuitionistic mathematics might well be mixed. The ontology of choice sequences presented below is one immediatee example.[1] Thus the study of choice sequences can not be seen as just an exercise in intuitionistic mathematics.

Now let us return to choice sequences themselves. Above, we introduced three different types of choice sequences: lawlike, hesitant and lawless sequences. Lawlike and lawless sequences are the two extremes with respect to predetermined information, whereas hesitant sequences lie somewhere between them.

Lawlike or *constructive sequences* are mappings from $\mathbb{N}$ into $\mathbb{N}$ restricted or predetermined by some law. This means that every value of the sequence has to be uniquely computable from all previously given values. Therefore, this type of sequence is also often called a *lawlike or constructive function from* $\mathbb{N}$ *into* $\mathbb{N}$.

Hesitant sequences differ from lawlike sequences in the respect that the process starts without any restriction on future values. However, at some moment during the process a defining law determining all future values might become known. Thus for hesitant sequences, a state of lawlikeness might be reached during the process whereas for lawlike sequences the defining law is given from the beginning.

For *Lawless sequences* at no point during the process of choosing values, a restricting law can be found.

As already mentioned above, these three types of choice sequences can be seen as different examples of the determination of a mathematical object by approximation from above or from below. Lawlike sequences are completely predetermined. Therefore, even though only a finite number of values is known at all times, the predetermining law itself can be seen as a complete and precise characterisation of the object which gives a precise approximation from above. Lawless sequences, on the other hand, are completely undetermined. Every new value chosen is new information about the object. Thus lawless sequences are approximated from below. With hesitant sequences, both types of approximation are mixed.

The fact that a classical treatment is possible for some types of choice sequences suggests that classical mathematics is in fact a fragment of intuitionistic mathematics for which the law of the excluded middle holds. Classical mathematics might be called the "fragment of possibility" of intuitionistic mathematics in the sense that for an object of this fragment every property is either provable or refutable.

The theory of lawless sequences as well as the system of intuitionistic logic and intu-

[1] Note, that the formalisation of choice sequences was not done by Brouwer himself. When Brouwer talks about choice sequences in his work, he usually talks about what was later on termed *lawless sequences*.

itionistic analysis were not formalised by Brouwer but have been developed later by other logicians, e.g. Heyting, Troelstra, Kreisel, etc. Thus intuitionistic logic is based on intuitionistic mathematics (the "intuitionistic philosophy"), whereas classical mathematics is based on classical logic. Nevertheless, it is quite useful to point out the main differences between intuitionistic logic and classical logic, since lawless sequences are a useful tool to intuitionistically refute undesirable classical principles.

The main difference between intuitionistic logic and classical logic already pointed out above is that the *principle of the excluded middle* does not hold in intuitionistic logic. A direct consequence of this fact is the invalidity of the *eliminability of double negation*. In general, from an intuitionistic point of view, for a formula to be true means that the formula is true *and* that there is a prove of this fact. A negated statement, therefore, has to be refutable or contradictory. From this point of view, the invalidity of the *elimination of double negation* is apparent. A doubly negated statement is of course weaker than a nonnegated statement. The former means that there is a proof that the statement is not contradictory, whereas the latter means that a proof of the statement can be obtained (Heyting 1956).

Heyting also showed that classical logic is indeed a fragment of intuitionistic logic, paralleling the above conceptual relation between intuitionistic an classical mathematics (Heyting (1956)).

In the course of this thesis we will take a closer look at lawless choice sequences, which are an exclusively intuitionistic concept and are inherently incomplete objects. Especially their incomplete character makes them an ideal and interesting topic and a potential tool to look at. We try to expand the notion of lawless sequences and to apply the newly introduced objects to the linguistic field of natural language semantics.

The first chapter gives a detailed overview of the most important results obtained for lawless sequences in the literature. The four axioms needed to characterize both the sequences and all operations on them are motivated and explained. Furthermore, the bar theorem and its consequence, the fan theorem, are discussed. Lastly we will give a proof of the elimination theorem for lawless sequences. It shows that the addition of lawless sequences to the arithmetical system $\mathbf{IDB}_1$ is a conservative extension.

The second chapter discusses three modifications or generalisations of lawless choice sequences: bilateral choice sequences, indexed choice sequences and bundled choice sequences. For all three types of sequences the technical apparatus is redefined and the compatibility of the new process and the operations on it with the axioms given in chapter 1 for lawless sequences is checked. Additionally, the three types of processes will be compared to each other and we will try to give translations among the various types of processes.

In the third chapter, one of the modifications of chapter 2 is applied to the formal system used in modern linguistics for the treatment of natural language semantics called possible-worlds semantics. The system is a higher-order intensional language with a possible-worlds semantics which is a standard Kripke semantics. The idea is that this Kripke semantics can be substituted by choice sequences that give a potentially infinite but locally finite semantics compatible with intuitionistic reasoning.

Chapter four introduces with short discussions four issues or interesting questions that were opened up by the discussion in the previous three chapters.

Chapter 1

The Theory of Lawless Sequences

Lawless sequences arechoice processes from $\mathbb{N}$ into $\mathbb{N}$, where the choice of values is completely unrestricted, i.e. the values do not adhere to a law. Thus, no value can be computed from previous values, and at every moment only an initial, already "chosen" segment of the sequence is known. As it turns out, lawless sequences are easier to handle conceptually if it is assumed that every finite sequence of values is an initial segment of some lawless sequence. Therefore, it is permitted to specify the initial segment of a sequence by explicitly stating a finite sequence of values. However, no additional information can be assumed.

In the literature, the intuition behind these sequences is usually made graspable by using the picture of casting a dice: the initial segment can be modeled by a finite number of deliberate placings of the dice followed by an infinite number of casts (Kreisel (1968), Troelstra (1977)).[1] The freedom to deliberately place the values of the initial segment also implies that every finite sequence is an initial segment of some lawless sequence.

Troelstra terms unrestricted sequences without a specified initial segment as *proto-lawless sequences*. The universe of proto-lawless sequences is part of the universe of lawless sequences; a proto-lawless sequence is a lawless sequence with the empty tuple $\langle\rangle$ as the specified initial segment.

In light of the above discussion, a short comment about the difference between lawless and hesitant sequences as described in the introduction is in order. From the short characterization of hesitant sequences given, the idea might arise that these sequences form a special class of lawless sequences. That this is not the case and that in fact the set of hesitant sequences and the set of lawless sequences are disjoint, can be seen from a simple consideration. Since lawless sequences are completely unrestricted with respect to all future choices, the following fact is certainly true for all lawless

[1]The picture of casts of a dice should, however, not suggest that choice sequences are stochastic processes.

sequences α: $\neg(\alpha = a)$ for a a lawlike process. $\neg(\alpha = a)$, however, is false for a hesitant sequence α since a construction law might become known at some stage. By similar reasoning, $(\alpha = a)$ for hesitant α and lawlike a is false as a construction law might not yet be known. Therefore, only $\neg\neg(\alpha = a)$ can be stated with complete certainty and the sets $H = \{\alpha |\ \neg\neg(\alpha = a),\ a \text{ lawlike}\}$ of hesitant sequences and $LS = \{\alpha |\ \neg(\alpha = a),\ a \text{ lawlike}\}$ of lawless sequences are trivially disjoint.

In the following section we will give a few definitions and notations for the axiomatization and discussion of the universe of lawless sequences and operations thereupon. The content of this chapter was taken from various articles and books about lawless sequences and intuitionistic mathematics covering more or less the same amount of information - Kreisel (1968), Troelstra (1977), Troelstra (1983), Troelstra and van Dalen (1988a) and Troelstra and van Dalen (1988b) - therefore, only direct adoptions of proofs will be cited.

1.1 Basic Definitions and Notations

Notation: For lawless sequences lowercase Greek letters, α, β, $\gamma \ldots$, will be used as variable names. The values or elements of a lawless sequence α are written as αn for $n \in \mathbb{N}$. Variable names for lawlike processes will be lowercase Latin letters, $a, b, c, \ldots$, from the beginning of the alphabet.

Sequences and tuples may be coded as a natural number with the help of the following pairing function.

Definition 1.1. The function j is a surjective *pairing*[2] from $\mathbb{N}^2 \to \mathbb{N}$ with inverses j_1 and j_2, such that $\forall z \exists x \exists y\ j(x, y) = z$, $\forall x j(j_1 x, j_2 x) = x$ and $j(x, y) = j(x', y') \to x = x' \wedge y = y'$.

The pairing function j can be used recursively to code p-tuples $\nu^p(x_0, \ldots, x_{p-1})$ or sequences as follows.

$$\nu^0(x_0) = x_0 \quad \nu^1(x_0, x_1) = j(x_0, x_1) \quad \nu^p(x_0, \ldots, x_p) = j(x_0, \nu^{p-1}(x_1, \ldots, x_p))$$

The code number (or *sequence number*) n of a tuple or sequence $x_0, \ldots, x_p$ will be written as $n = \langle x_0, \ldots, x_p \rangle$. Since the pairing function is injective, the code number can be identified with the tuple or sequence it codes. The empty sequence $\langle\rangle$ will be identified with 0. As an abbreviation, $\hat{x}$ is written instead of $\langle x \rangle$.

[2] Any such function is adequate. As a natural coding would serve Kleene's primitive recursive coding of finite sequences (Moschovakis 1986).

Definition 1.2. *Concatenation* is indicated by the operator symbol $*$. The concatenation of the sequences $\langle x_0, \ldots, x_p \rangle$ and $\langle x_{p+1}, \ldots, x_{p+n} \rangle$ is thus

$$\langle x_0, \ldots, x_p \rangle * \langle x_{p+1}, \ldots, x_{p+n} \rangle = \langle x_0, \ldots, x_{p+n} \rangle.$$

Definition 1.3. The *length-function* for sequences is given as

$$\mathrm{lth} : \mathbb{N} \to \mathbb{N}, \quad \mathrm{lth}\, \langle\rangle = 0, \quad \mathrm{lth}\, \langle x_0, \ldots, x_p \rangle = p + 1.$$

Notation: An *initial segment* of length n, $\langle \alpha_0, \ldots, \alpha_{n-1} \rangle$, of a sequence α is written as $\overline{\alpha}n$.

Definition 1.4. The *initial segment relation* is well-defined for a pair containing a choice sequence and a sequence number.

$$\alpha \in n \quad \text{iff} \quad \overline{\alpha}(\mathrm{lth} n) = n$$

The intial segment relation can be interpreted topologically. If for all $n \in \mathbb{N}$ the sets of all choice sequences with the initial segment n are defined as the open sets of a topology on the universe of choice sequences, $\alpha \in n$ is readable in the sense that α is an element of the open set generated by the finite sequence n.

We can furthermore define an order on sequence numbers.

Definition 1.5. The *order* $\prec$ on sequence numbers is defined for sequence numbers $n, m \in \mathbb{N} :\ m \prec n := \exists m'\ m * m' = n.$

Now let us turn to the axiomatization of lawless sequences, starting with the axioms specifying the objects themselves.

1.2 Axioms Specifying the Objects

Since lawless sequences are characteristically unrestricted, only two axioms stating the most basic requirements are formulated. For one, that every finite sequence of natural numbers is an initial segment of some lawless sequence, and secondly, that equality for lawless sequences is decidable.

1.2.1 Density

The density axiom formalises the condition that every finite sequence of values in $\mathbb{N}$ is an initial segment of some lawless sequence.

$LS1 \qquad \forall n \exists \alpha (\alpha \in n)$

As a direct consequence, there are infinitely many sequences for each finite initial segment, i.e. $\forall\alpha\forall x(\alpha \in n * \hat{x} \rightarrow \alpha \in n)$. Also note, that $LS1$ is the only existential axiom proposed for lawless sequences.

1.2.2 Decidability of Equality

The following axiom stating the *decidability of equality* for lawless sequences needs a short preparatory explanation. The only information known about a certain lawless sequence α at every moment is a finite initial segment $\overline{\alpha}n$ and its "individuality" as a process. Therefore, intensional equality, i.e. equivalence, and extensional equality of two lawless sequences mutually entail each other.

Proposition 1.6. $\alpha \equiv \beta \leftrightarrow \alpha = \beta$ where $\alpha = \beta :\Leftrightarrow \forall n(\alpha n = \beta n)$

Proof. The implication $\alpha \equiv \beta \rightarrow \alpha = \beta$ is selfevident: if two lawless sequences are equivalent, i.e. if they are the same process, all of their elements are identical. The other direction, $\alpha = \beta \rightarrow \alpha \equiv \beta$, follows from the information available about lawless sequences. Since only a finite initial segment is known at every moment and all future values are left unrestricted, a statement like $\forall n(\alpha n = \beta n)$ concerning *all* values of two possibly distinct processes α and β (including future values!) implies either that the two processes are actually the same process, i.e. $\alpha \equiv \beta$, or that one of the two processes is in fact not a lawless sequence since all its values are predetermined by the other process. Thus $\alpha = \beta \rightarrow \alpha \equiv \beta$. $\square$

Therefore, intensional and extensional equality are decidable: either two lawless sequences are given as the same process or they are not.

$$LS2 \qquad \alpha \equiv \beta \vee \neg\alpha \equiv \beta$$
$$LS2' \qquad \alpha = \beta \vee \neg\alpha = \beta$$

1.3 Axioms Specifying the Operations

The nature of the operations on lawless sequences is the main focus of the theory of lawless sequences, since the objects themselves do not offer an elaborate structure. Studying the operations on lawless sequences leads to the observation that all such operations are constant on an open set (with respect to the topology on Baire space generated by sets of lawless sequences sharing the same initial segments) of the universe of lawless sequences. A direct consequence is that all operations on lawless sequences have to be continuous.

1.3.1 Operations on Choice Sequences

For choice sequences in general and for lawless sequences in particular, there are two kinds of functionals that need to be looked at: on the one hand functionals from sequences into natural numbers, $\Phi : \mathbb{N}^{\mathbb{N}} \to \mathbb{N}$, and on the other hand functionals from sequences into sequences, $\Psi : \mathbb{N}^{\mathbb{N}} \to \mathbb{N}^{\mathbb{N}}$. This of course does not exhaust all possible functionals definable for choice sequences. The two types of functionals discussed here play an important role in the formalisation of intuitionistic analysis. The following definitions and formulae are valid for all types of sequences but will be given specifically for lawless ones.

A class of functionals $\Phi : \mathbb{N}^{\mathbb{N}} \to \mathbb{N}$ can be defined from two different perspectives. One possibility is to look at all *continuous* functionals from lawless sequences into natural numbers. The second possibility is to define a class of functionals with the desired properties inductively.

Following from $LS3$, the axiom of *open data* (see section 1.3.2), functionals are continuous if and only if they depend only on a finite initial segment. Thus, the value of the functional for a lawless sequence α can be computed from a sufficiently long initial segment. By making use of sequence coding, the information contained in the initial segment can be compressed into a $n \in \mathbb{N}$. This means that the class of continuous functionals $Cont_{LS}$ can be induced by a class of *neighbourhood functions* K_0 from $\mathbb{N}$ into $\mathbb{N}$.

Definition 1.7. A function $\xi : \mathbb{N} \to \mathbb{N}$ is in K_0 iff $k \in \mathbb{N}$ is a sequence number, i.e. $k = \langle \alpha_0, \ldots, \alpha_p \rangle$ an initial segment of a sequence α, and there is a functional $\Phi_\xi : \mathbb{N}^{\mathbb{N}} \to \mathbb{N} \in \ Cont_{LS}$ such that

$$\xi k = \begin{cases} 0 & \text{the value of } \Phi_\xi(\alpha) \text{ is not yet determined} \\ x+1 & \Phi_\xi(\alpha) = x \end{cases}$$

$\xi k = 0$ can equivalenty be interpreted as the initial segment k being too short to compute a value for Φ_ξ.

The functions in K_0 have to obey two conditions, *consistency* and *totality*.

Consistency: $\forall m \forall n \forall x(\xi m = x + 1 \to \xi(m * n) = x + 1)$

Totality: $\forall \alpha \exists m \exists x(\overline{\alpha}(\text{lth } m) = m \wedge \xi m = x + 1)$

A handy abbreviation that also expresses the fact that ξ induces a functional Φ_ξ is the notation

$$(F1) : \xi(\alpha) = x :\Leftrightarrow \exists y(\xi(\overline{\alpha}y) = x + 1)$$

The second possibility is to define functionals from $\mathbb{N}^{\mathbb{N}}$ into $\mathbb{N}$ inductively. The class of functionals defined in this way is called the class of *Brouwer operations Ind.*

Definition 1.8.

$$
\begin{array}{lll}
Ind & (i) & \lambda\alpha.n \in Ind \quad \forall n \in \mathbb{N} \\
 & (ii) & \Phi^{(0)}, \Phi^{(1)}, \Phi^{(2)} \ldots \in Ind \ \Rightarrow \Phi \in Ind,\ \Phi(\alpha) = \Phi^{(\alpha 0)}(\lambda x.\alpha(x+1)) \\
 & & \text{or in other words: } \forall x(\Phi_{\langle x \rangle} \in Ind) \rightarrow \Phi \in Ind \quad \text{with } \Phi_n(\alpha) := \Phi(n * \alpha)
\end{array}
$$

Since *Ind* is the smallest class of functionals satisfying the above closure conditions, we can formulate a principle of induction.

$$\forall n(\lambda\alpha.n \in X) \wedge [\forall x(\Phi_{\langle x \rangle} \in X) \rightarrow \Phi \in X] \rightarrow \ Ind \subset \ X$$

Same as for the class of continuous functionals, the class of inductively defined functionals on (lawless) sequences can be represented by a class of neighbourhood functions K. Parallel to *Ind*, the class K is also defined inductively, and, thus, all elements of K are lawlike operations.[3]

Definition 1.9. K is the smallest class of functions $\xi : \mathbb{N} \rightarrow \mathbb{N}$ that is closed under $K1$ and $K2$.

$$
\begin{array}{ll}
K1 & \lambda n.y + 1 \in K \quad \forall y \in \mathbb{N} \\
K2 & \xi 0 = 0 \wedge \forall x(\lambda n.\xi(\hat{x} * n) \in K) \rightarrow \xi \in K
\end{array}
$$

The principle of induction can be adapted to K. Let

$$A(\alpha, Q) :\Leftrightarrow \alpha = \lambda n.y + 1 \vee [\alpha 0 = 0 \wedge \forall x(\lambda n.\alpha(\hat{x} * n) \in K)]$$

be the conjunction of the closure conditions, then

$$K3 \quad \forall\alpha[A(\alpha, Q) \rightarrow Q\alpha] \rightarrow \forall[K\alpha \rightarrow Q\alpha]$$

states that K is the minimal class of functionals satisfying the conditions $K1$ and $K2$. Now that the two classes of functionals on lawless sequences are defined, the central issue is whether those two classes coincide, i.e. whether $Ind = Cont_{LS}$ and equivalently whether $K = K_0$.

[3]Functions and functionals also have to be interpreted as processes rather than completely given objects. They are given by an algorithm and not by their graph.

The two subset relations $Ind \subset Cont \subset Cont_{LS}$, with $Cont$ being the class of continuous operations defined on all types of sequences, are shown by induction over Ind (Troelstra and van Dalen 1988a:226).

Proposition 1.10. $K \subset K_0$

Proof. We apply $K3$ with $\alpha \in K_0$ for $Q(\alpha)$. We have to verify the premiss of $K3$. $\lambda n.x + 1 \in K_0$ is obvious; and if for all x $\lambda m.\alpha(\hat{x} * m) \in K_0$, and $\alpha 0 = 0$, then also $\alpha \in K_0$. For let β be any sequence and let $\Delta\beta := \lambda x.\beta(x + 1)$. Then $\beta = \langle\beta 0\rangle * \Delta\beta$, and $\alpha(\overline{\beta}(x + 1)) = \alpha(\langle\beta 0\rangle * \overline{\Delta\beta}x)$; now $\lambda m.\alpha(\langle\beta 0\rangle * m) \in K_0$, hence for some x $\alpha(\langle\beta 0\rangle * \overline{\Delta\beta}x) \neq 0$. So $\alpha(\overline{\beta}(x + 1)) \neq 0$. It is also easy to see that $\alpha n \neq 0 \rightarrow \alpha(n * m) = \alpha n$; we only have to note that $\alpha n \neq 0$ means $n = \hat{x} * n$ for some x, n', so $\alpha(\hat{x} * n') \neq 0$, hence also $\alpha(\hat{x} * n') = \alpha(\hat{x} * n' * m) = \alpha(n * m)$. This establishes the premiss of $K3$ for this Q, so $K \subset K_0$. ☐

The problematic direction is $Cont_{LS} \subset Ind$. Brouwer gave a proof for the latter direction, which, however, is not unproblematic. The central assumption for the proof is equivalent to the possibility of cut-elimination for the system LS (Troelstra 1977). That $K_0 \subset K$ holds classically is shown by the following proposition (Troelstra and van Dalen 1988a:227).

Proposition 1.11. (classical logic + axiom of dependend choice[4]) $K_0 \subset K$

Proof. Assume $\gamma \in K_0 \setminus K$, and let $\gamma n := \lambda m.\gamma(n + m)$, then $\gamma_{\langle\rangle} = \gamma \in K_0 \setminus K$, so $\gamma(\langle\rangle) = 0$ (otherwise γ would be constant and greater than zero, and therefore belong to K). For some x_0 $\gamma_{\langle x_0\rangle} \in K_0 \setminus K$, for if $\forall x(\gamma_{\langle x\rangle} \in K)$, then also $\gamma \in K$. Repeating this we find successively $x_1, x_2, \ldots$ such that $\gamma_{\langle x_0, x_1\rangle} \in K_0 \setminus K$, $\gamma_{\langle x_0, x_1, x_2\rangle} \in K_0 \setminus K$, etc. So there is a sequence α, with $\alpha i = x_i$ such that for all y $\gamma_{\overline{\alpha}y} \in K_0 \setminus K$. But since $\gamma \in K_0$, thee is a y for which $\gamma_{\overline{\alpha}y}(\langle\rangle) > 0$, i.e. $\gamma_{\overline{\alpha}y}$ is non-zero constant and therefore belongs to K; we have thus obtained a contradiction. ☐

The above proof uses the *principle of the excluded middle* at each step and is therefore not an intuitionistically valid proof.

The problem of $K = K_0$ will be discussed further in section 1.4 since the equality implies Brouwer's *bar theorem* and vice versa.

[4]The axiom of dependend choice $DC - N$ has the following form

$$DC - N \quad \forall x \exists y A(x, y) \rightarrow \forall x \exists \phi \in \mathbb{N}^{\mathbb{N}}\ [\phi 0 = 0 \wedge \forall n A(\phi n, \phi(n + 1))]$$

and is an equivalent but specialized form of the axiom of countable choice for two variables in $\mathbb{N}$, $AC - NN$. The related $AC - NF$ for a number valued and a function valued argument of A is given in section 1.6.1.

A simple residue of the bar theorem is the *extension principle*: each $\Gamma \in Cont_{LS}$ can be extended to a continuous operation defined on all sequences, i.e. each $\Gamma \in Cont_{LS}$ has an extension $\Gamma' \in Cont$.

One way to picture this extension is by using a process *Abstr* that is applied to non-lawless sequences. *Abstr* "forgets" all additional information beyond the given initial segment. However, *Abstr* is not a "good" operation on sequences, since $Abstr(b) = b$ cannot be proved even though extensional equality is obviously given. The preparational reasoning for $LS2$, i.e. $\alpha = \beta \rightarrow \alpha \equiv \beta$, requires $Abstr(b)$, the pseudo-lawless version of b, to still be intensionally equivalent to the non-lawless b, which is false, since all intensional information has been "forgotten".

The construction of neighbourhood functions for functionals $\Psi : \mathbb{N}^{\mathbb{N}} \rightarrow \mathbb{N}^{\mathbb{N}}$ can be done in analogy to the construction of the functions for $Cont_{LS}$ and Ind. In fact, the class of neighbourhood functions for $Cont_{LS}$ and Ind also represent the class $Cont^1$ of continuous functionals. (We assume that $K = K_0$.)

We can also formulate a biconditional relation between the neighbourhood function $e \in K$ and the induced functional $(e|\alpha) : \mathbb{N}^{\mathbb{N}} \rightarrow \mathbb{N}^{\mathbb{N}} \in Cont^1$.

$$(e|\alpha)(x) = y \leftrightarrow \exists z(e(\hat{x} * \overline{\alpha}z) = y + 1)$$

This biconditional relation completely defines the elements of $Cont^1$ in terms of elements in K.

1.3.2 Open Data

Since the only information about a lawless sequence that can be "worked with" is a finite initial segment, all operations on lawless sequences have to depend only on such a segment. This of course means that a function from lawless sequences into $\mathbb{N}$ has to map all lawless sequences with the same initial segment onto the same natural number. In the case of one lawless argument this observation can be put as follows.

$$A(\alpha) \rightarrow \exists n(\alpha \in n \wedge \forall \beta \in n\ A(\beta))$$

The above can be generalized for a function of higher arity in the form given below as $LS3$. $LS3$ is called the axiom of *open data*.

$$\begin{aligned} LS3 \qquad & (\neq (\alpha, \alpha_0, \ldots, \alpha_p) \wedge A(\alpha, \alpha_0, \ldots, \alpha_p)) \rightarrow \\ & \exists n(\alpha \in n \wedge \forall \beta \in n(\neq (\beta, \alpha_0, \ldots, \alpha_p) \wedge A(\beta, \alpha_0, \ldots, \alpha_p)) \end{aligned}$$

The notation $\neq (\alpha, \alpha_0, \ldots, \alpha_p)$ is the usual abbreviation for the situation

$$\bigwedge_{i=0}^{p} \alpha \neq \alpha_i.$$

Note that the axiom of open data turns out be false if an operation uses any other (intensional) information about lawless sequences instead of some arbitrary finite initial segment, e.g. the deliberately placed initial segment. Let Φ_I be the operation that assigns to each lawless sequence the length of its deliberately placed initial segment. Then, obviously,

$$\Phi_I(\alpha) \rightarrow \exists n(\alpha \in n \wedge \forall \beta \in n \ \Phi_I(\beta))$$

is false, since β might have been started with a different deliberate initial segment than α even though they both start with the initial segment having the sequence number n. A related proposition that follows from the axiom of open data shows that of two interdependent sequences at least one sequence can not be lawless (Troelstra 1977:16).

Proposition 1.12. For α, β and γ choice sequences, such that $\gamma(2k) = \alpha(k)$ and $\gamma(2k+1) = \beta(k)$, it follows that α, β and γ cannot be simultaneously lawless.

Proof. Let $\alpha \neq \beta$. Assume $\gamma = \alpha$, then $\gamma(2k) = \alpha(2k) = \alpha(k)$ and thus $\alpha(2^n) = \alpha(1)\ \forall n \in \mathbb{N}$. Open data for one lawless argument applied to $\forall n(\alpha(2^n) = \alpha(1))$ gives $\exists y \forall \xi \in \overline{\alpha}\ \forall n\ (\xi(2^n) = \xi(1))$, which is false. Therefore, $\gamma \neq \alpha$ and analogously $\gamma \neq \beta$. So $\neq (\gamma, \alpha, \beta)$ is satisfied and $LS3$ can be applied to $\forall x(\gamma(2x) = \alpha(x) \wedge \gamma(2x+1) = \beta(x))$ which also gives the false conclusion $\exists y \forall \xi \in \overline{\gamma}y\ \forall x(\neq (\xi, \alpha, \beta) \wedge \xi(2x) = \alpha(x) \wedge \xi(2x+1) = \beta(x))$. □

Two consequences of $LS1 - LS3$ are given in the next two propositions from Troelstra and van Dalen (1988b:650).

Proposition 1.13. $\forall \alpha \neg \forall x(\alpha x \neq 0)$

Proof. Assume $\forall x(\alpha x \neq 0)$, then by $LS3$ follows that $\forall \beta \in \overline{\alpha}n \forall x(\beta x \neq 0)$ for some $n \in \mathbb{N}$. This is refuted by taking a $\beta \in \overline{\alpha}n * \langle 0 \rangle$, which is possible by $LS1$. □

Proposition 1.14. Identity is the only lawlike operation under which the universe of lawless sequences is closed.

Proof. Suppose $\alpha = \Gamma\beta, \alpha \neq \beta$. Then by $LS3$ $\forall \gamma \in \overline{\alpha}x(\gamma \neq \beta \rightarrow \gamma = \Gamma\beta)$ for some x, which is clearly false by $LS1$: choose $\gamma_0 \in \overline{\alpha}x * \langle \beta x + 1 \rangle, \gamma_1 \in \overline{\alpha}x * \langle \beta x + 2 \rangle$, and the contradiction is immediate. □

1.3.3 Continuity

Weak continuity for operations on lawless sequences directly follows from the axiom of open data, $A(\alpha) \to \exists n(\alpha \in n \wedge \forall \beta \in n\ A(\beta))$.

$$WC-N \qquad \forall\alpha\exists x A(\alpha, x) \to \forall\alpha\exists x\exists n(\alpha \in n \wedge \forall\beta \in n\ A(\beta, x))$$

By making use of the equivalence $(A \vee B) \leftrightarrow \exists x[(x = 0 \wedge A) \vee (x \neq 0 \wedge B)]$, a version of $WC-N$ for disjunctions can be derived.

$$WC-N^{\vee} \qquad \forall\alpha(A\alpha \vee B\alpha) \to \forall\alpha\exists x(\forall\beta \in \overline{\alpha}x A\beta \vee \forall\beta \in \overline{\alpha}x B\beta)$$

A nice consequence of $WC-N^{\vee}$ is that it refutes the principle of the excluded middle $(\forall - PEM)$ for universally closed formulas. The proof is taken from Troelstra and van Dalen (1988a).

Proposition 1.15. $WC-N^{\vee}$ refutes $\forall - PEM$.

Proof. Assume $\forall - PEM$, i.e. $\forall\alpha(\forall x(\alpha x = 0) \vee \neg\forall x(\alpha x = 0))$, then by $WC-N^{\vee}$ we obtain $\forall\alpha\exists y(\forall\beta \in \overline{\alpha}y\forall x(\beta x = 0) \vee \forall\beta \in \overline{\alpha}y\neg\forall x(\beta x = 0))$. Now specialize α to $\lambda x.0$, then for some y, with $n = \overline{(\lambda x.0)}(y)$

$$\forall\beta \in n\forall x(\beta x = 0) \vee \forall\beta \in n\neg\forall x(\beta x = 0)$$

The first disjunct is false, as may be seen by taking any $\beta \in n * \langle 1 \rangle$; and the second disjunct is also false, as follows by choosing $\beta = \alpha = \lambda x.0$. □

The above proposition is one of the intuitionistically provable refutations of classical principles by using choice sequences.

A stronger form of continuity can be expressed with the help of continuous lawlike operations on lawless sequences, as defined in section 1.3.1. The intuitionistic interpretation of the quantifier combination $\forall\alpha\exists x$ requires a method for finding x to be given, i.e. a proof of $\forall\alpha\exists x\ A(\alpha, x)$ contains a way of calculating x for each α. This leads to an axiom of choice or selection principle.

$$\forall\alpha\exists x A(\alpha, x) \to \exists\Theta\forall\alpha A(\alpha, \Theta\alpha)$$

The combination of this axiom of choice with $LS3$, which requires Θ as an operation on lawless sequences to be continuous, implies the following stronger form of continuity.

$$LS4_1 \qquad \forall\alpha\exists x A(\alpha, x) \to \exists\Gamma\forall\alpha A(\alpha, \Gamma\alpha) \quad \text{with } \Gamma \in \mathbb{N}^{\mathbb{N}} \to \mathbb{N}, \text{ continuous}$$

The same reasoning leads to continuity for assigning any type of lawlike object, e.g. a lawlike sequence a, to a lawless sequence by a functional $\Psi : \mathbb{N}^{\mathbb{N}} \to \mathbb{N}^{\mathbb{N}}$.

$$LS4_2 \qquad \forall\alpha\exists a A(\alpha, a) \to \exists\Psi\forall\alpha A(\alpha, \Psi\alpha) \quad \text{with } \Psi \in \mathbb{N}^{\mathbb{N}} \to \mathbb{N}^{\mathbb{N}}, \text{ continuous}$$

$LS4_1$ and $LS4_2$ can also be stated by using the inducing neighbourhood functions.

$$LS4_1{}^* \qquad \forall\alpha\exists x A(\alpha, x) \to \exists\xi \in K_0\forall\alpha A(\alpha, \xi(\alpha))$$
$$LS4_2{}^* \qquad \forall\alpha\exists a A(\alpha, a) \to \exists\zeta \in K_0\forall\alpha A(\alpha, (\zeta|\alpha))$$

The strongest form of continuity expressible, is for $\xi = e$ in $LS4_1{}^*$ and $e \in K$. Since K is the class of inductively defined neighbourhood functions, e is a lawlike function. The substitution of e for ξ requires the assumption, that the class of neighbourhood functions for continuous functionals K_0 and the class of neighbourhood functions for inductively defined functionals K are in fact equal. For a discussion, see sections 1.3.1 and 1.4.

$$LS4_1{}^{**} \qquad \forall\alpha\exists x A(\alpha, x) \to \exists e \in K\forall\alpha A(\alpha, e(\alpha))$$
$$LS4_2{}^{**} \qquad \forall\alpha\exists a A(\alpha, a) \to \exists e \in K\exists b\forall\alpha A(\alpha, b_{e(\alpha)})$$
$$\text{where } b_x := \lambda y.bj(x, y)$$

If we generalize $LS4_2{}^{**}$ for operations with more than one lawless argument, we obtain the most general form of $LS4$,

$$LS4 \qquad \forall\alpha_1 \ldots \forall\alpha_p\exists a(\#(\alpha_1, \ldots, \alpha_p) \wedge A(\alpha_0, \ldots, \alpha_p, a)) \to$$
$$\exists e \in K\exists b\forall\alpha_1 \ldots \forall\alpha_p(\#(\alpha_1, \ldots, \alpha_p) \to A(\alpha_1, \ldots, \alpha_p, b_{e(\nu_p(\alpha_1,\ldots,\alpha_p))}))$$

where $\nu_p(\alpha_1, \ldots, \alpha_p) := \lambda x.\nu_p(\alpha_1 x, \ldots, \alpha_p x)$ and $\#(\alpha_1, \ldots, \alpha_p) := \bigwedge(\alpha_i \neq \alpha_j)$ for $i \neq j$. This means that p-tuples of independent lawless sequences, $\alpha_1, \ldots, \alpha_p$ with $\#(\alpha_1, \ldots, \alpha_p)$, “behave” like a single lawless sequence with respect to e.
The special case, i.e. the generalization of $LS4_1{}^{**}$, is

$$\forall\alpha_1 \ldots \forall\alpha_p\exists x(\#(\alpha_1, \ldots, \alpha_p) \wedge A(\alpha_0, \ldots, \alpha_p, x)) \to$$
$$\exists e \in K\forall\alpha_1 \ldots \forall\alpha_p(\#(\alpha_1, \ldots, \alpha_p) \to A(\alpha_1, \ldots, \alpha_p, e(\nu_p(\alpha_1, \ldots, \alpha_p))))$$

with the same notation as for $LS4$.
Some caution is required when continuity schemata for other types of choice sequences are formulated. Since other types of information besides a finite initial segment might be used to compute functional values, extensional equality of two sequences α and β might not result in equality of the values assigned by a functional Γ, i.e. $\alpha = \beta \not\to \Gamma\alpha =$

$\Gamma\beta$. However, as long as the functionals on other types of sequences can be guaranteed to be extensional, which more or less means that they only operate on a finite initial segment, then the continuity schemata for lawless sequences can be adopted fully.

1.4 The Bar Theorem

Even though Brouwer took the *bar theorem* as a consequence of his theory of choice sequences. The theorem does not follow from the axioms $LS1 - LS4$ formulated in the previous sections. Thus, in the literature, it is usually given in form of an induction schema, BI_D, with additional comments regarding its status as a theorem.

$$BI_D \quad \forall n(Pn \vee \neg Pn) \wedge \forall\alpha\exists x P(\overline{\alpha}x) \wedge \forall n(Pn \to Qn) \wedge \forall n(\forall y Q(n * \hat{y}) \to Q(n)) \to Q(\langle\rangle)$$

As always, $\forall\alpha\exists x P(\overline{\alpha}x)$ or more generally $P(\overline{\alpha}x)$ are well-formed if computation of the value of P for a sequence α requires only a finite initial segment.

BI_D can be interpreted as an induction principle over *trees*, thus the name BI_D (*decidable bar induction*, Troelstra and van Dalen (1988a:224ff)).

Definition 1.16. A set T is called a *tree* iff $\langle\rangle \in T, \forall n(n \in T \vee n \notin T)$ and $\forall nm(n \in T \wedge m \prec n \to m \in T)$

For every tree T, $T^{\sim} := \{n \in T |\ \forall m \prec n(m \in T)\}$ can be defined. A tree is *well-founded* iff $\forall\alpha\exists x \neg(\overline{\alpha}x \in T)$, i.e. for each sequence only a certain initial segment is in the tree.

If T is well-founded, $T^{\sim} \setminus T$ are the terminal nodes of $T^{\sim}$.

Definition 1.17. $T^{\sim} \setminus T$ is called a *bar* for T. For each α we can find an x such that $\overline{\alpha}x \in T$ and $\forall n \prec \overline{\alpha}x(n \in T)$ but $\overline{\alpha}x + 1 \notin T$.

Now, BI_D can be interpreted with respect to well founded trees. From $\forall n(Pn \vee \neg Pn)$ and $\forall\alpha\exists x P(\overline{\alpha}x)$ it follows that $\{n |\ \forall m \prec n\ \neg Pm\}$ is a well-founded tree with $\{n |\ PN \wedge \forall m \prec n\ \neg Pm\}$ as a bar. $\forall n(Pn \to Qn)$ and $\forall n(\forall y\ Q(n * \hat{y}) \to Q(n))$ means Q holds on a bar and if Q holds for all successors of a node n it holds for n, i.e. Q is "inherited" upwards. As the conclusion BI_D states that Q holds for the empty sequence which is the root of the tree.

It was already stated in section 1.3.1 that the bar theorem implies $K = K_0$ and vice versa. Thus if a proof can be given for BI_D, $K = K_0$ follows immediately. Brouwer claimed to have a proof of BI_D. The proof, however, is not unproblematic, which will be discussed after establishing the implicational relation between the bar theorem and $K = K_0$.

First, it is shown (informally) that BI_D implies $K = K_0$ or equivalently $Ind = Cont_{LS}$.

Proposition 1.18. BI_D implies $Cont_{LS} = Ind$

Proof. Since $Ind \subset Cont_{LS}$ was already shown in section 1.3.1, only $Cont_{LS} \subset Ind$ is left to show. Take for P and Q in BI_D the predicates

$$Pn := \text{'n is sufficiently long to compute } \Phi\alpha \text{ for } \alpha \in n\text{'}$$
$$Qn := \Phi_n \in Ind$$

where Φ is any functional of $Cont_{LS}$ (representable by a neighbourhood function of K_0) and $\Phi_n(\xi) := \Phi(n * \xi)$. P and Q satisfy all the premises of BI_D: P is decidable, since the algorithm of computing Φ should be able to determine if an initial segment is long enough for the computation; $\forall\alpha\exists x P(\overline{\alpha}x)$ follows from the fact that Φ is total; $\forall n(\forall y Q(n * \hat{y}) \rightarrow Q(n))$ is the second closure condition for Ind. Finally, $Q(\langle\rangle)$ just means that $\Phi \in Ind$. Since this holds for any Φ of $Cont_{LS}$, $Cont_{LS} \subset Ind$. □

Let us now look at Brouwer's argument for the bar theorem, i.e. his proof. The main question and the central issue when proving the bar theorem is how to be sure that $\forall\alpha\exists x P(\overline{\alpha}x)$ is valid. We can reformulate this question by introducing two new terms: we say a predicate P is a *bar* iff $\forall\alpha\exists x P(\overline{\alpha}x)$ and *n is P-barred* iff $\forall\alpha \in n\exists x P(\overline{\alpha}x)$. Then the main issue becomes: how can we be certain that $\langle\rangle$ is P-barred?

Brouwer's basic assumption is that from any "fully analysed" proof of "$\langle\rangle$ is P-barred" a proof-tree can be extracted with "$\langle\rangle$ is P-barred" as its root that contains only three types of inferences:

(I) $P(n)$, hence n is barred (immediate inference), i.e.

$$Pn \Rightarrow \forall\alpha \in n\exists x P(\overline{\alpha}x)$$

(D) for all x, $n * \hat{x}$ is barred, hence n is barred (downward inference), i.e.

$$\forall x\forall\alpha \in n * \hat{x}\exists y P(\overline{\alpha}y) \Rightarrow \forall\alpha \in n\exists y P(\overline{\alpha}y)$$

(U) n is barred, hence $n * \hat{x}$ is barred (upwards inference), i.e.

$$\forall\alpha \in n\exists y P(\overline{\alpha}y) \Rightarrow \forall\alpha \in n * \hat{x}\exists y P(\overline{\alpha}y)$$

Under suitable assumptions for P, all (U) inferences can be eliminated, e.g. if P is monotone[5] or decidable. For the resulting proof-tree after the elimination of all (U) inferences, Brouwer's argument for BI_D continues as follows: assume the premise of

[5] A predicate P is *monotone* iff $n \in P \wedge n \prec m \rightarrow m \in P$.

BI_D to be given and establish Qn everytime $\forall\alpha \in n\exists y P(\overline{\alpha}y)$ occurs in the proof-tree. This effectively means, that one associates $Pn \to Qn$ and $\forall n(\forall y Q(n * \hat{y}) \to Q(n))$ with (I) and (D) inferences respectively.

The crucial problem of the idea for a proof given above is the assumption that for any P there can be found a "fully analysed" proof of "$\langle\rangle$ is P-barred". As stated in the last paragraph, (U) inferences can only be eliminated under suitable assumptions for P. Therefore, if P does not meet these requirements, the needed proof-tree might not exist. An example due to Kleene (cf. Troelstra and van Dalen (1988a:233)) shows that for arbitrary P, (U) inferences are not eliminable.

Thus, Brouwer's proof of BI_D can not be accepted as such. However, addition of the bar theorem as an axiom or induction schema to $LS1-4$ does not lead to an "incoherent" theory.

A more general version of BI_D is the *monotone bar induction* BI_M.

$$BI_M \quad \forall\alpha\exists x P(\overline{\alpha}x) \wedge \forall nm(Pn \to P(n * m)) \wedge \forall n(\forall y P(n * \hat{y}) \to Pn) \wedge \to P(\langle\rangle)$$

The assumption of BI_M is equivalent to the assumption of continuity for lawless sequences and BI_D (Troelstra and van Dalen 1988a:231).

1.5 The Fan Theorem

The *fan theorem* is a direct consequence of the bar theorem restricted to a certain kind of finitely branching trees, called *fan*.

Definition 1.19. A set of finite sequences of natural numbers T is a *fan* iff $\langle\rangle \in T, \forall n(n \in T \vee n \notin T), \forall nm(n \in T \wedge m \prec n \to m \in T), \forall n \in T\exists x(n * \hat{x} \in T)$ and $\forall n \in T\exists z\forall x(n * \hat{x} \in T \to x \leq z)$.

In other words, a fan is an inhabited, decidable set of sequence numbers closed under predecessor where each node has at least one successor and for each node there are only a finite amount of possible successors. I.e. a fan is a tree with two additional requirements.

Definition 1.20. A sequence α is called a *branch* of a fan T iff all initial segments belong to T:

$$\alpha \in T := \forall x(\overline{\alpha}x \in T)$$

Like the bar theorem, the fan theorem should rather be called "fan axiom". For a decidable property A, the fan thorem states that, if for each infinite branch of a fan T, there is a finite initial segment satisfying A, then there exists a uniform upper bound to the length of the satisfying initial segment.

$$FAN_D(T) \qquad \forall n \in T(An \vee \neg An) \wedge \forall \alpha \in T \exists x A(\overline{\alpha}x) \rightarrow \exists z \forall \alpha \in T \exists y \leq z A(\overline{\alpha}y)$$

The fan theorem in the above form is valid for all fans, thus it can be generalised to

$$FAN_D: \ fan(T) \rightarrow FAN_D(T)$$

where $fan(T)$ expresses that the set T is a fan, i.e. that T satisfies the formulas given in the definition of a fan.

FAN_D can be seen as the intuitionistic version of Koenig's lemma. FAN_D is classically equivalent to Koenig's lemma. One direction can be seen b< considering the set $T_A := \{m | \ \forall m' \preceq m \neg Am'\}$ of all inital segments whose predecessors in T do not satisfy A for a fan T and any decidable A. T_A is a tree with finite branches. Suppose that $\neg\exists z \forall \alpha \in T \exists y \leq z A(\overline{\alpha}y)$ or the classically equivalent $\exists z \forall \alpha \in T \exists y \leq z \neg A(\overline{\alpha}y)$, then T_A contains arbitrarily long branches and thus by Koenig's lemma an infinite branch, which contradicts $\forall \alpha \in T \exists x A(\overline{\alpha}x)$. Hence, Koenig's lemma implies FAN_D.

Another interpretation of FAN_D is *restricted compactness* for the fan T with the topology generated by the basis $V_n := \{\alpha \in T | \ n \in T\}$ for $n \in T$: any decidable set of basis elements $\{V_n | \ An \wedge n \in T\}$ covering T has a finite subcover $\{V_n | \ An \wedge n \in T \ \wedge \mathrm{lth}(n) \leq z\}$.

FAN_D can be strengthened to apply also to non-decidable predicates, which (following up on the above compactness interpretation) expresses *full compactness* of a fan T.

$$FAN(T) \qquad \forall \alpha \in T \exists x A(\overline{\alpha}x) \rightarrow \exists z \forall \alpha \in T \exists y \leq z A(\overline{\alpha}y)$$

That $FAN(T)$ means compactness for T can be seen as follows. For any cover of T $\{W_i | \ i \in I\}$ and any inital segment $\overline{\alpha}x$ of a branch α, we know there is a W_i containing the basis element $V_{\overline{\alpha}x}$. Now $FAN(T)$ implies $\exists z \forall \alpha \in T \exists y \leq z \exists i \in I(V_{\overline{\alpha}y} \subset W_i)$. Set all $y = z$. Since T is a fan, there are only finitely many initial segments of length z, $n_1, \ldots, n_p$. For each n_i and V_{n_i} choose a W_j such that $V_{n_i} \subset W_j$. The resulting set of elements of the inital cover is the finite subcover. Thus T is compact. Of course compactness conversely implies $FAN(T)$.

Again, the general form of the fan theorem is

$$FAN: \ fan(T) \rightarrow FAN(T)$$

FAN is classically valid.

Two more generalizations can be formulated for FAN. The first is for predicates $A(\alpha, x)$:

$$fan(T) \wedge \forall\alpha \in T \exists x A(\alpha, x) \rightarrow \exists z \forall\alpha \in T \exists y \leq z A(\alpha, y)$$

The second generalization is a combination of FAN with weak continuity of lawless sequences.

$$FAN^* \qquad fan(T) \wedge \forall\alpha \in T \exists x A(\alpha, x) \rightarrow \exists z \forall\alpha \in T \exists x \forall\beta \in \overline{\alpha}x A(\beta, x)$$

Neither the first generalization nor FAN^* are classically valid.

1.6 The System LS* and the Elimination Theorem

1.6.1 EL and EL$_1$

The system of elementary analysis formalises basic arithmetic and constructive functions on natural numbers.

Definition 1.21. Elementary Analysis - **EL:**

1. variables in $\mathbb{N}$; notation: x,y,z,u,v,w,n,m,...
2. variables for constructive functions (lawlike sequences); notation: a,b,c,d, ...
3. constants: 0, successor S, equality $=$, abstraction operator λ, recursor R, pairing function with inverses j, j_1, j_2, application of functions to $\mathbb{N}$ Φ
4. logical constants: $\wedge, \vee, \rightarrow, \forall, \exists$

For better readability the following abbreviations are adopted: $\neg A :\Leftrightarrow A \rightarrow (S0 = 0)$ and $\phi t :\leftrightarrow \Phi\phi t$ for a function symbol ϕ and a numerical term t.

For the successor, equality, pairing functions, induction with respect to formulae, conversion and primitive recursion the usual axioms are assumed. Furthermore, the quantifier-free axiom of choice is added.

$$QF-AC \qquad \forall x \exists y A(x, y) \rightarrow \exists a \forall x A(x, ax) \quad (A \text{ quantifier-free})$$

If the choice schema for numbers is replaced with the axiom of choice for one number valued and one function valued argument, the resulting system is called **EL**$_1$.

$$AC-NF \qquad \forall x \exists a A(x, a) \rightarrow \exists b \forall x A(x, (b)_x)$$
$$\text{where } (b)_x := \lambda y.bj(x, y)$$

The function b induces the choice function $(b)_x$.

A word has to be said about choice axioms and choice schemata in intuitionistic mathematics. Let us look at $QF - AC$. In intuitionistic mathematics stating $\forall x \exists y A(x, y)$ implies that the proof of the existence of a value y for x such that $A(x, y)$ contains a way of computing said y. This "way of computing y" already given in the proof of the premiss of the choice schema can be naturally represented as a choice function a. Analogous reasoning applies for $AC - NF$ and all other choice axioms. Therefore, the respective choice functions are not conjured up from thin air, as it might seem, but follow from the premisses.

1.6.2 IDB$_1$, LS and LS*

The system **IBD**$_1$ is an extension of **EL**$_1$ by variables for elements of K, the inductively defined neighbourhood functions (Brouwer operations) discussed in section 1.3.1. This means that, additionally to the constructive functions, the functionals induced by the neighbourhood functions in K are available for computation.

Definition 1.22. Inductively defined Brouwer operations - **IDB**$_1$**:**

1. system **EL**$_1$
2. variables for neighbourhood functions (Brouwer operations); notation: e,f, ...

For the Brouwer operations the axioms K1 to K3, the abbreviations F1 and F2, a second abstraction symbol λ' and new rules of term formation are added.

$$
\begin{array}{ll}
K1 & \lambda n.y + 1 \in K \\
K2 & \alpha 0 = 0 \wedge \forall x(\lambda n.\alpha(\hat{x} * n) \in K) \rightarrow \alpha \in K \\
K3 & \forall \alpha[A(\alpha, Q) \rightarrow Q\alpha] \rightarrow \forall[K\alpha \rightarrow Q\alpha] \\
 & \text{with } A(\alpha, Q) :\Leftrightarrow \alpha = \lambda n.y + 1 \vee [\alpha 0 = 0 \wedge \forall x(\lambda n.\alpha(\hat{x} * n) \in K)] \\
F1 & e(\alpha) = x \leftrightarrow \exists y(e(\overline{\alpha}y) = x + 1) \\
F2 & (e|\alpha)(x) = y \leftrightarrow \exists z(e(\hat{x} * \overline{\alpha}z) = y + 1)
\end{array}
$$

The conversion axiom for the abstractions symbol λ' is formulated analogously to the abstraction over constructive functions in **EL** and **EL**$_1$.

A further extension of **IDB**$_1$ by adding variables, axioms and term formation rules for lawless sequences results in the system **LS**. Adding quantifier symbols for lawless sequences gives the system **LS***.

Definition 1.23. Lawless Sequences - **LS*:**

1. system $\mathbf{IDB}_1$
2. variables for lawless sequences; notation: α, β, γ, ...
3. logical constants: $\underline{\forall}, \underline{\exists}$

For $\underline{\forall}\alpha \in t$ and $\underline{\exists}\alpha \in t$ we need to add the following axioms.

$$\begin{array}{ll}
1) & \underline{\forall}\alpha_1 \in t_1 \ldots \underline{\forall}\alpha_p \in t_p\ \underline{\forall}\beta \in t(P(\alpha_1, \ldots, \alpha_p) \to Q(\alpha_1, \ldots, \alpha_p, \beta)) \Rightarrow \\
& \underline{\forall}\alpha_1 \in t_1 \ldots \underline{\forall}\alpha_p \in t_p\ (P(\alpha_1, \ldots, \alpha_p) \to \underline{\forall}\beta \in t\ Q(\alpha_1, \ldots, \alpha_p, \beta)) \\
& \underline{\forall}\alpha_1 \in t_1 \ldots \underline{\forall}\alpha_p \in t_p\ \underline{\exists}\beta \in t(\underline{\forall}\gamma \in t\ P(\alpha_1, \ldots, \alpha_p, \gamma) \to P(\alpha_1, \ldots, \alpha_p, \beta)) \\
2) & \underline{\forall}\alpha_1 \in t_1 \ldots \underline{\forall}\alpha_p \in t_p\ \underline{\forall}\beta \in t(Q(\alpha_1, \ldots, \alpha_p, \beta) \to Q(\alpha_1, \ldots, \alpha_p,)) \Rightarrow \\
& \underline{\forall}\alpha_1 \in t_1 \ldots \underline{\forall}\alpha_p \in t_p\ (\underline{\exists}\beta \in t\ Q(\alpha_1, \ldots, \alpha_p, \beta) \to P(\alpha_1, \ldots, \alpha_p)) \\
& \underline{\forall}\alpha_1 \in t_1 \ldots \underline{\forall}\alpha_p \in t_p\ \underline{\forall}\beta \in t(P(\alpha_1, \ldots, \alpha_p, \beta) \to \underline{\exists}\gamma \in t\ P(\alpha_1, \ldots, \alpha_p, \gamma)) \\
3) & \underline{\forall}\alpha_1 \ldots \underline{\forall}\alpha_n\ \underline{\forall}\alpha \in t\ A(\alpha_1, \ldots, \alpha_n, \alpha) \leftrightarrow \underline{\forall}\alpha_1 \ldots \underline{\forall}\alpha_n\ \underline{\forall}\alpha(\alpha \in t \to A(\alpha_1, \ldots, \alpha_n, \alpha)) \\
& \underline{\forall}\alpha_1 \ldots \underline{\forall}\alpha_n\ \underline{\exists}\alpha \in t\ A(\alpha_1, \ldots, \alpha_n, \alpha) \leftrightarrow \underline{\forall}\alpha_1 \ldots \underline{\forall}\alpha_n\ \underline{\exists}\alpha(\alpha \in t \wedge A(\alpha_1, \ldots, \alpha_n, \alpha))
\end{array}$$

For maximal writing and reading comfort, we will introduce abbreviations.

Notation:

$$\begin{array}{l}
\dot{\forall}\alpha A(\alpha, \vec{\beta}) := \underline{\forall}\alpha(\neq (\alpha, \vec{\beta}) \to A(\alpha, \vec{\beta})) \\
\dot{\exists}\alpha A(\alpha, \vec{\beta}) := \underline{\exists}\alpha(\neq (\alpha, \vec{\beta}) \wedge A(\alpha, \vec{\beta})) \\
\dot{\forall}\alpha A(\vec{\alpha}) := \dot{\forall}\alpha_1 \ldots \dot{\forall}\alpha_p A(\vec{\alpha}) \\
\dot{\exists}\alpha A(\vec{\alpha}) := \dot{\forall}\alpha_1 \ldots \dot{\forall}\alpha_p A(\vec{\alpha}) \\
\vec{\alpha} \in \vec{n} := \alpha_1 \in n_1 \wedge \ldots \wedge \alpha_p \in n_p \\
\dot{\forall}\vec{\alpha} \in \vec{n} A := \dot{\forall}\alpha_1 \in n_1, \ldots, \dot{\forall}\alpha_p \in n_p A \\
\vec{n} * \vec{m} := (n_1 * m_1, \ldots, n_p * m_p)
\end{array}$$

Then the axioms LS 1-4 can be written as:

$$\begin{array}{ll}
LS1 & \forall n \underline{\exists}\alpha(\alpha \in n) \\
LS2 & \underline{\forall}\alpha\underline{\forall}\beta(\alpha = \beta \vee \neg\alpha = \beta) \\
LS3 & \dot{\forall}[A(\alpha, \vec{\beta}) \to \exists n(\alpha \in n \wedge \underline{\forall}\gamma \in nA(\gamma, \vec{\beta}))] \\
LS4 & \dot{\forall}\vec{\alpha}\exists a A(\vec{\alpha}, a) \to \exists e \in K \forall n(en \neq 0 \to \exists a \dot{\forall}\vec{a} \in nA(\vec{\alpha}, a))
\end{array}$$

The elimination result that will be proved in the following section implies that the system **LS** is a conservative extension of $\mathbf{IDB}_1$, i.e. that quantification over lawless sequences can be regarded as a "figure of speech", as Troelstra puts it, or a convenient conceptional point of view instead of a further strengthening of the theory.

1.6.3 The Elimination Theorem

The elimination theorem as formulated by Kreisel (1968) shows that for theorems containing quantification over lawless sequences provable in **LS*** in fact a logically equivalent theorem not containing quantification over lawless sequences is provable in the smaller system $\mathbf{IDB}_1$. In Troelstra's words: quantification over lawless sequences can be treated as a "figure of speech".

Theorem 1.24. (Kreisel 1968) There exists a mapping τ of formulae of **LS** without free lawless variables onto the formulae of $\mathbf{IDB}_1$ such that
1) $\tau(A) = A$ for a formula of $\mathbf{IDB}_1$
2) **LS*** $\vdash$ A $\leftrightarrow \tau(A)$
3) **LS*** $\vdash$ A $\Leftrightarrow$ $\mathbf{IDB}_1 \vdash \tau(A)$

The proof given here for 1) and 2) of the theorem is taken from Troelstra and van Dalen (1988b) with some completions and explanations where necessary.

The logical equivalences needed to define the mapping τ are first proved as lemmas.

Lemma 1.25. **LS*** $\vdash \dot{\forall}\vec{\alpha}\exists x A(\vec{\alpha}, x) \rightarrow \exists e \in K \forall n(en \neq 0 \rightarrow \exists x \dot{\forall}\vec{\alpha} \in n A(\vec{\alpha}, x))$

Proof. $\dot{\forall}\vec{\alpha}\exists x A(\vec{\alpha}, x) \rightarrow \exists e \in K \forall n(en \neq 0 \rightarrow \exists x \dot{\forall}\vec{\alpha} \in n A(\vec{\alpha}, x))$ is an easy consequence of $LS4$. □

The above consequence of $LS4$ has in fact already been mentioned in section 1.3.3 as the special case of $LS4$, where the lawlike element is a number instead of a sequence or function.

Lemma 1.26.
(i) **LS*** $\vdash \dot{\forall}\vec{\alpha} \in \vec{n}(A(\vec{\alpha}, \vec{\beta}) \rightarrow B(\vec{\alpha}, \vec{\beta})) \leftrightarrow \forall m(\dot{\forall}\vec{\alpha} \in \vec{n} * m A(\vec{\alpha}, \vec{\beta}) \rightarrow \dot{\forall}\vec{\alpha} \in \vec{n} * m B(\vec{\alpha}, \vec{\beta}))$
(ii) **LS*** $\vdash \dot{\exists}\alpha \in n B(\alpha, \vec{\beta}) \leftrightarrow \exists m \dot{\forall}\alpha \in n * m B(\alpha, \vec{\beta})$

Proof. (i) The direction from the left to the right is immediate. For the converse, suppose first $\vec{\alpha} \equiv \alpha, \vec{n} \equiv n$ and assume (1) $\forall m(\dot{\forall}\alpha \in n * m A(\alpha, \vec{\beta}) \rightarrow \dot{\forall}\alpha \in n B(\alpha, \vec{\beta})$. Let $\neq (\alpha, \vec{\beta}), \alpha \in n, A((\alpha, \vec{\beta}))$; then, by $LS3$, $\alpha \in m'$ and $\dot{\forall}\gamma \in m'(\gamma \in n \wedge A(\gamma, \vec{\beta}))$ for some m'. Without loss of generality we can assume $m' = n * m$, hence $\dot{\forall}\gamma \in n * m \wedge A(\gamma, \vec{\beta})$; therefore by our assumption (1) $\dot{\forall}\gamma \in n \wedge B(\gamma, \vec{\beta})$, and so $B(\gamma, \vec{\beta})$. Thus $\dot{\forall}\alpha \in n(A(\alpha, \vec{\beta} \rightarrow B(\alpha, \vec{\beta}))$. For a vector $\vec{\alpha}$ of length p in the statement of the lemma we have to apply this argument p times.

(ii) For the direction from left to right, assume $\dot{\exists}\alpha \in nB(\alpha, \vec{\beta})$ and $\alpha \in n$. Now, n doesn't have to be long enough to compute $B(\alpha, \vec{\beta})$. However, there exists an m such that $\alpha \in n * m$ and $n * m$ is long enough to compute $B(\alpha, \vec{\beta})$. Then, by $LS3$, $\exists m \dot{\forall} \alpha \in n * mB(\alpha, \vec{\beta})$.
For the converse assume $\exists m \dot{\forall} \alpha \in n * mB(\alpha, \vec{\beta})$ and $\alpha \in n * m$ then $\alpha \in n$ and $B(\alpha, \vec{\beta})$, thus $\exists \alpha \in nB(\alpha, \vec{\beta})$. □

Lemma 1.27. Let a be either a number variable or a variable for lawlike sequences, then

$$\mathbf{LS^*} \vdash \dot{\forall} \vec{\alpha} \in \vec{n} \exists a A(\vec{\alpha}, a) \leftrightarrow \exists e \in K \forall m (em \neq 0 \rightarrow \exists a \dot{\forall} \vec{\alpha} \in \vec{n} * mA(\vec{\alpha}, a))$$

Proof. The direction from right to left follows directly from the definition of a in the left hand side as either $a := e(\nu_p(\alpha_1, \ldots, \alpha_p))$ or $a := b_{e(\nu_p(\alpha_1, \ldots, \alpha_p))}$ with a lawlike b.
For the converse, let $\vec{n} \equiv \langle n_1, \ldots, n_p \rangle$, and assume the left-hand side, then for some $f \in K \quad \forall n(fm \neq 0 \rightarrow \exists a \dot{\forall} \vec{\alpha} \in \vec{n} * mA(\vec{\alpha}, a))$.
We define e by

$$en = y + 1 := \exists n'(n_1 * k_1^p n \succeq k_1^p n' \wedge \ldots \wedge n_p * k_p^p n \succeq k_p^p n' \wedge fn' = y + 1)$$

It is easy to see that $e \in K$ and that e satisfies the right-hand side of the statement of the lemma. □

Lemma 1.28.
(i) $\mathbf{LS^*} \vdash \dot{\forall} \alpha \in n(A(\alpha, \vec{\beta}) \leftrightarrow \dot{\forall} \alpha \in n[A(\alpha, \vec{\beta}) \wedge \gamma = \gamma]$
(ii) $\mathbf{LS^*} \vdash \dot{\exists} \alpha \in n(A(\alpha, \vec{\beta}) \leftrightarrow \dot{\exists} \alpha \in n[A(\alpha, \vec{\beta}) \wedge \gamma = \gamma]$

Proof. We prove (i) and (ii) by simultaneous induction on the logical complexity of A, where (ii) at each step is deduced from (i). We treat only some typical cases, all other cases follow analogously.
Case 1. $A(\alpha, \vec{\beta}) \equiv B(\alpha, \vec{\beta}) \wedge C(\alpha, \vec{\beta})$. Then

$$\begin{aligned}
&\dot{\forall} \alpha \in n(A(\alpha, \vec{\beta}) \wedge \gamma = \gamma) \\
&\leftrightarrow \dot{\forall} \alpha \in n((B(\alpha, \vec{\beta}) \wedge \gamma = \gamma) \wedge (C(\alpha, \vec{\beta}) \wedge \gamma = \gamma)) \\
&\leftrightarrow \dot{\forall} \alpha \in n(B(\alpha, \vec{\beta}) \wedge \gamma = \gamma) \wedge \dot{\forall} \alpha \in n(C(\alpha, \vec{\beta}) \wedge \gamma = \gamma) \\
&\leftrightarrow \dot{\forall} \alpha \in nB(\alpha, \vec{\beta}) \wedge \dot{\forall} \alpha \in nC(\alpha, \vec{\beta}) \\
&\leftrightarrow \dot{\forall} \alpha \in nA(\alpha, \vec{\beta})
\end{aligned}$$

Case 2. Let $A(\alpha, \vec{\beta}) \equiv \underline{\forall}\alpha' B(\alpha, \alpha', \vec{\beta})$. Without loss of generality we can suppose $\neq (\gamma, \vec{\beta})$. Then the following are equivalent:

$$\dot{\forall}\alpha \in n(\underline{\forall}\alpha' B(\alpha, \alpha', \vec{\beta}) \wedge \gamma = \gamma)$$
$$\dot{\forall}\alpha \in n(\gamma = \gamma \wedge \dot{\forall}\alpha' B(\alpha, \alpha', \vec{\beta}) \wedge B(\alpha, \alpha, \vec{\beta}) \wedge \ldots \wedge B(\alpha, \beta_i, \vec{\beta}) \wedge \ldots)$$
$$\dot{\forall}\alpha \in n[\dot{\forall}\alpha'(B(\alpha, \alpha', \vec{\beta}) \wedge \gamma = \gamma) \wedge$$
$$B(\alpha, \alpha, \vec{\beta}) \wedge \ldots \wedge B(\alpha, \beta_i, \vec{\beta}) \wedge \ldots \wedge B(\alpha, \gamma, \vec{\beta}))]$$
$$\dot{\forall}\alpha \in n\dot{\forall}\alpha'(B(\alpha, \alpha', \vec{\beta}) \wedge \gamma = \gamma)$$
$$\dot{\forall}\alpha \in n\dot{\forall}\alpha' B(\alpha, \alpha', \vec{\beta})$$

The first equivalence holds by the meaning of $\dot{\forall}\alpha'$ and $LS2$. The second equivalence holds because the outer quantifier ensures $\neq (\gamma, \alpha)$ and the inner quantifier ensures $\forall\alpha'(\neq (\alpha', \alpha, \vec{\beta}) \to B(\alpha', \alpha, \vec{\beta}))$. The remaining equivalences are immediate.

Case 3. Suppose (i) has been proved; then by (ii) of lemma 1.26 $\dot{\exists}\alpha \in nA(\alpha, \vec{\beta}) \leftrightarrow \exists m\dot{\forall}\alpha \in n{*}mA(\alpha, \vec{\beta})$, therefore $\dot{\exists}\alpha \in n(A(\alpha, \vec{\beta}) \wedge \gamma = \gamma) \leftrightarrow \exists m\dot{\forall}\alpha \in n{*}m(A(\alpha, \vec{\beta}) \wedge \gamma = \gamma) \leftrightarrow \exists m\dot{\forall}\alpha \in n * mA(\alpha, \vec{\beta}) \leftrightarrow \dot{\exists}\alpha \in n(A(\alpha, \vec{\beta})$. $\square$

Lemma 1.29. **LS*** $\vdash \dot{\forall}\vec{\alpha} \in \vec{n}t_1[\vec{\alpha}] = t_2[\vec{\alpha}] \leftrightarrow \dot{\forall}\vec{a} \in \vec{n}t_1[\vec{a}] = t_2[\vec{a}]$, where $\forall\vec{a} \in \vec{n}$ is defined the same way as $\forall\vec{\alpha} \in \vec{n}$.

Proof. By induction on the complexity of a numerical term one can prove:

$$\exists e \in K\forall n(en \neq 0 \to \exists x\underline{\forall}\vec{\alpha} \in \vec{n}\forall\vec{a} \in \vec{n}(t[\vec{\alpha}] = t[\vec{a}] = x))$$

where **LS***$\vdash e(\alpha_1, \ldots, \alpha_p) = t[\alpha_1, \ldots, \alpha_p]$ and **IDB**$_1$$\vdash e(a_1, \ldots, a_p) = t[a_1, \ldots, a_p]$. The induction requires checking all possible ways to form terms. For illustration, two cases are considered.

Case 1. Let e_1 and e_2 for $t_1[\alpha_1, \alpha_2, x]$ and $t_2[\alpha_1, \alpha_2]$ be known already. We wish to construct the corresponding f for

$$t_3 := t_1[\alpha_1, \alpha_2, t_2[\alpha_1, \alpha_2]]$$

e_1 is a functor $\Phi[x]$; let e_1' be defined such that

$$e_1'0 = 0, \quad e_1'(\hat{x} * n) = \Phi[x]n = e_1 n.$$

Then take for f: $f := \lambda' n.e_1'(sg(e_2 n)(\langle e_2 n - 1\rangle * n))$.

Case 2. Assume t to be defined by primitive recursion from t_1 and t_2, i.e.

$$t[\alpha_1, \alpha_2, x] = Rt_1t_2x, \text{ i.e.}$$
$$t[\alpha_1, \alpha_2, 0] := t_1[\alpha_1, \alpha_2] \text{ and } t[\alpha_1, \alpha_2, Sz] := t_2[j(t[\alpha_1, \alpha_2, z], z), \alpha_1, \alpha_2, x]$$

and assume e_1 and e_2 to be already constructed. Note also, that the addition of a constant R to the language satisfying the above is a conservative extension, since it is already proved by induction that $\forall x(\lambda n.Ref(\hat{x} * n) \in K)$.

Now take $f := Re_1e_2'$ where e_2' is defined analogously to e_1' above.

The lemma follows directly from the construction of the appropriate $e \in K$. Since the construction has already been proved, we only need to prove

$$\dot{\forall}\vec{\alpha} \in \vec{n}(e(\alpha_1, \ldots, \alpha_p) = f(\alpha_1, \ldots, \alpha_p)) \leftrightarrow \forall \vec{a} \in \vec{n}(e(a_1, \ldots, a_p) = f(a_1, \ldots, a_p))$$

Assume $\dot{\forall}\vec{\alpha} \in \vec{n}(e(\alpha_1, \ldots, \alpha_p) = f(\alpha_1, \ldots, \alpha_p))$ and let $\vec{\alpha} \in \vec{n}$. We can find an m such that $\nu_p(a_1, \ldots, a_p) \in m, em \neq 0, fm \neq 0$. Take $\vec{\alpha} \in \vec{n}, \nu_p(\alpha_1, \ldots, \alpha_p) \in m$ and $\#(\alpha_1, \ldots, \alpha_p)$. Then $e(\alpha_1, \ldots, \alpha_p) = f(\alpha_1, \ldots, \alpha_p) = em-1$, hence also $e(a_1, \ldots, a_p) = f(a_1, \ldots, a_p)$ by $LS3$.

The other direction follows analogously. □

Now, let us define the mapping τ by giving a step by step rewriting algorithm that transforms formulas of **LS*** into formulas of $\mathbf{IDB}_1$. The following definition is taken nearly verbatim from Troelstra and van Dalen (1988b:663).

Definition 1.30. *The mapping τ:*

Step 1. The first step in defining $\tau(A)$ consists of rewriting A in terms of quantifiers $\dot{\forall}\alpha, \dot{\exists}\alpha$ using the equivalences

$$\underline{\forall}\alpha A(\alpha, \vec{\beta}) \leftrightarrow \dot{\forall}\alpha A(\alpha, \vec{\beta}) \wedge \bigwedge_{i \leq p} A(\beta_i, \vec{\beta})$$
$$\underline{\exists}\alpha A(\alpha, \vec{\beta}) \leftrightarrow \dot{\exists}\alpha A(\alpha, \vec{\beta}) \wedge \bigvee_{i \leq p} A(\beta_i, \vec{\beta})$$

The result is uniquely determined if $\alpha, \vec{\beta}$ is a list of *all* the lawless variables actually free in A. However in the steps described below we also need dummy variables. This is uniquely determined modulo logical equivalence by lemma 1.28; the replacement transforms any formula into an **LS**-provably equivalent formula.

We shall now define the effect of τ on formulas written with propositional operators and $\dot{\forall}\alpha \in t, \dot{\exists}\alpha \in t$ (t without lawless variables); these operations are regarded as logical primitives in the syntactic definition. $\dot{\forall}\alpha, \dot{\exists}\alpha$ are regarded as synonymous with

$\dot{\forall}\alpha \in \langle\rangle$, $\dot{\exists}\alpha \in \langle\rangle$. Disjunction is treated as defined, and we take $0 = 1$ as falsum.
Step 2. We first eliminate all occurences of $\dot{\exists}\alpha \in n$ by replacements proved in lemma 1.26

$$\dot{\exists}\alpha \in nA(\alpha, \vec{\beta}) \mapsto \exists m \dot{\forall}\alpha \in n * mA(\alpha, \vec{\beta}).$$

Step 3. Next we show how to transform formulas of the form $\dot{\forall}\vec{\alpha} \in \vec{n}A(\vec{\alpha})$ into formulas with fewer logical operations within the scope of blocks $\dot{\forall}\vec{\alpha} \in \vec{t}'$. This process can be continued till we arrive at $\dot{\forall}\vec{\alpha} \in \vec{t}(P)$ for P prime. The necessary replacements given by

$$\dot{\forall}\vec{\alpha} \in \vec{n}(A \wedge B) \mapsto \dot{\forall}\vec{\alpha} \in \vec{n}A \wedge \dot{\forall}\vec{\alpha} \in \vec{n}B \text{ by lemma 1.28}$$
$$\dot{\forall}\vec{\alpha} \in \vec{n}(A \rightarrow B) \mapsto \forall \vec{m}(\dot{\forall}\vec{\alpha} \in \vec{n} * \vec{m}A \wedge \dot{\forall}\vec{\alpha} \in \vec{n} * \vec{m}B) \text{ by lemma 1.26}$$
$$\dot{\forall}\vec{\alpha} \in \vec{n}\exists a A(\vec{\alpha}, a) \mapsto \exists e \in K \forall m(em \neq 0 \rightarrow \exists a \dot{\forall}\vec{\alpha} \in \vec{n} * mA(\vec{\alpha}, a)) \text{ by lemma 1.27}$$
$$\dot{\forall}\vec{\alpha} \in \vec{n}\forall a A(\vec{\alpha}, a) \mapsto \forall a \dot{\forall}\vec{\alpha} \in \vec{n}A(\vec{\alpha}, a)$$

where a is a numerical or lawlike sequence variable.
Step 4. Finally we have to show how to eliminate quantifier strings $\dot{\forall}\vec{\alpha} \in \vec{t}$ in front of prime formulas, where it is assumed that $\vec{t}$ does not contain lawless variables. We use

$$\dot{\forall}\vec{\alpha} \in \vec{t}(s_1[\vec{\alpha}] = s_2[\vec{\alpha}]) \mapsto \forall\vec{a} \in \vec{t}(s_1[\vec{a}] = s_2[\vec{a}]) \text{ by lemma 1.29}$$

and $\dot{\forall}\vec{\alpha} \in \vec{t}(\phi \in K) \longmapsto \phi \in K$ if no variable of $\dot{\forall}\vec{\alpha}$ occurs in ϕ, and

$$\dot{\forall}\vec{\alpha} \in \vec{t}(\phi[\vec{\alpha}] \in K) \mapsto \dot{\forall}\vec{\alpha} \in \vec{t}\exists e(e \in K \wedge \forall x(\phi[\vec{\alpha}](x) = ex))$$

in all other cases, also by lemma 1.29.
Since it was shown in lemmas 1.26 - 1.29 that the substitutions and rewritings employed in defining τ replace a formula by an **LS***-provably equivalent one, the definition of τ can also be read as a proof of points 1) and 2) of the elimination theorem. Point 3) will not be proved here but can be found in Kreisel and Troelstra (1970)[6].
One important consequence of point 3) of the elimination theorem is

Proposition 1.31. **LS*** is conservative over $\mathbf{IDB}_1$.
which will also stated without proof.

[6] The proof uses induction on the length of derivations.

1.7 Summary

In this chapter the theory of lawless sequences as developed for the foundation of intuitionistic analysis has been discussed. The four axioms, $LS1 - LS4$, specify the lawless sequences themselves and possible types of operations on lawless sequences. Furthermore, the bar theorem and the fan theorem which both are ultimately added to the theory as axioms were formulated. Lastly, the formal system, **LS***, for arithmetic with lawless sequences was defined and the elimination theorem was proved. The theorem states that every provable statement containing quantification over lawless variables has an equivalent formulation not containing lawless variables which is moreover provable in a subsystem of **LS***.

As a last remark it needs to be said that the actual sequences used for intuitionistic analysis (cf. Troelstra and van Dalen (1988a) and Troelstra and van Dalen (1988b)) are a derived, more restricted type of choice sequence which only retain some of the properties of lawless sequences.

Chapter 2

Three New Variants of Choice Sequences

This chapter is dedicated to three kinds of generalisations and modifications of choice sequences. Choices sequences, as already mentioned in the introduction, are the prototypical unfinished objects. We will explore the possibilities provided by lawless sequences a little bit more in this direction.

The basic ideas behind the three modifications or generalisations presented in this section are that one might want to consider either more than one class of information or allow one class of information to develop in more than one direction.

The first subpart discusses a possible formalisation of the latter idea termed *bilateral choice sequences*. *Bilateral sequences* are specifically taylored to model bilateral growth (i.e. in two different directions) which is modelled by a mapping from an infinite set with a linear order with no smallest element to some set of values.

The second subpart is dedicated to *indexed choice sequences*, which are proposed to deal with different classes of information. Each value of an indexed choice sequence, hence, is "flavoured" with an index from an index set representing the different classes of information. The other properties of choice sequences are left untouched.

The third subpart explores another take at the idea of having multiple classes of information called *bundled choice sequences*. However, contrary to indexed choice sequences, at every moment each branch of information of a bundled sequence is increased by the same amount of values, modelling a uniform increase of information. To capture the uniformity, bundled sequences are formalised with tuples as values.

The type of choice process for all three new sequence types can either be lawlike, hesitant or lawless. However, only lawless bilateral choice sequences, lawless indexed choice sequences and lawless bundled choice sequences will be of interest in this section. Also, since the names of the different types start to get quite long, we will use the term *lawless sequences* for the choice sequences defined in chapter 1 and abbreviate *lawless indexed/bilateral/bundled choice sequences* with *indexed/bilateral/bundled (choice) sequences.*

2.1 Bilateral Choice Sequences

The modification of choice sequences in this section, the *bilateral choice sequences*, formalise the idea that a strand or class of information can develop in more than one direction. Not to break with the picture of a choice sequence as a (linear) sequence of values, we will only consider two directions of growth which results in a regular choice sequence with two open ends.

Definition 2.1. *Bilateral choice sequences* α_B are mappings from a linearly ordered set X with no smallest element into a (possibly denumerably infinite) set Y.

For example, for this section we will chose $\mathbb{Z}$ for X and $\mathbb{N}$ for Y, i.e. $\alpha_B : \mathbb{Z} \to \mathbb{N}$.

The choice process for bilateral sequences works similarly to the one for lawless sequences. After a finite initial segment has been set deliberately, a finite amount of values can be added at each moment to the initially set segment. The main difference between lawless sequences and bilateral sequences at this point is that the values can be added either to the left or the right of the previously chosen values, creating a two-ended process.

A consequence of this way of chosing and adding new values is that bilateral sequences cannot record the order of information gain. After two values are added successively, once on the left and once on the right of the finite existing segment, there is no way to determine which of the end values was added first. In a strictly linear structure with a fixed root (i.e. an order with a smallest element) like regular choice sequences and indexed/bundled choice sequences, this is not the case (cf. sections 2.2 and 2.3).

For example the sequence starts with a deliberately placed initial segment

$$\ldots, 2, 6, 8, 6, 55, \ldots$$

After 4 was added at the left side and 7 was added on the right side

$$\ldots, 4, 2, 6, 8, 6, 55, 7, \ldots$$

there is no way of determining the order in which 4 and 7 were added by simply looking at the new finite segment.

Another consequence is that the initial segment relation and the notion of an initial segment of length n can not be defined. Even though the sequence coding is easily adaptable to bilateral sequences since there is no change of make up for finite segments, the sequence coding depends on a fixed point of reference - a starting point of the sequence - to be at its left edge. Of course it is possible to arbitrarily set a starting point, e.g. at $\alpha_B(0)$ parallel to lawless sequences, but just setting a starting point does not ensure that a definition can be given.

Let $\alpha_B 0$ be the fixed root of the sequence. How to define $\alpha_B \in n$ or $\overline{\alpha_B}n$? For $\alpha_B \in n$ it does not suffice that there is some finite segment around $\alpha_B 0$ that is identical to n. Different initial segments around $\alpha_B 0$ give different values for operations on α_B and both continuity and consistency are lost. The situation also does not improve if the requirement that $\alpha_B 0$ divides the finite sequence n into two equally long parts is posed. The adoptions of this requirement restricts the sequences and all operations on them too severely: Equally long parts on both sides of the root imply a uniform growth in both directions of the sequence. This means that only bilateral sequences that can be written as bundled choice sequences with tuple length two would be allowed (cf. section 2.3).

It seems that what we need is to also have a reference point inside a finite segment that gets mapped onto $\alpha_B 0$. Hence, we will have to split (finite) sequences in two: a part that lies to the left of the reference point and a part that lies to the right of the reference point.

Definition 2.2. $\underleftarrow{\alpha_B 0} := \langle \alpha_B 0, \alpha_B(-1), \alpha_B(-2), \ldots \rangle$ is the *leftward monolateral sequence* induced by a bilateral sequence at point $\alpha_B 0$.

Definition 2.3. $\underrightarrow{\alpha_B 0} := \langle \alpha_B 0, \alpha_B 1, \alpha_B 2, \ldots \rangle$ is the *rightward monolateral sequence* induced by a bilateral sequence at point $\alpha_B 0$.

$\underleftarrow{\alpha_B 0}$ and $\underrightarrow{\alpha_B 0}$ divide the bilateral sequence into two normal choice sequences. These two sequences are lawlike since they fully depend on $\alpha_B 0$ (cf. proposition 12 in section 1.3.2). This does not matter for the definition of an inital segment for these sequences, however. As discussed in section 1.3.1, according to the *extension principle*, the operations on lawless sequences can be extended to lawless sequences. For this purpose a metatheoretical operator *Abstr* which forgets intensional information about sequences was introduced. It was shown, however, that *Abstr* can not be defined in the system and should only be used as an informal device.

Nevertheless, $Abstr(\underrightarrow{\alpha_B 0})$ and $Abstr(\underleftarrow{\alpha_B 0})$ can be treated as lawless sequences and the inital segment of length n for $\underleftarrow{\alpha_B 0}$ and $\underrightarrow{\alpha_B 0}$ is written as usual $\overline{\underleftarrow{\alpha_B 0}}n$ and $\overline{\underrightarrow{\alpha_B 0}}n$, respectively, and is well-defined.

The question at this point is whether it is possible to define $\alpha_B \in n$ with the help of the initial segments. Even though the initial segment relation is well-defined for the two monolateral sequences, no satisfactory definition can be given for the original bilateral sequence.

Consider the following definition.

Definition 2.4. The initial segment relation (to be discarded):
$\alpha_B \in n$ iff $\exists k \exists m(k * m = n \wedge \underleftarrow{\alpha_B 0} \in k \wedge \underrightarrow{\alpha_B 0} \in k0 * m)$

This definition suffers from the same problem discussed above: with more than one possible initial segment, operations on these sequences do not obey open data and are hence not continuous. To give an illustration let us consider the following two initial segments of two bilateral sequences. The roots $\alpha_{B,1}0$ and $\alpha_{B,2}0$ are underlined.

$$\alpha_{B,1} := \ldots, 1, \underline{2}, 1, 2, \ldots$$
$$\alpha_{B,2} := \ldots, 1, \underline{2}, 1, 3, \ldots$$

Under the definition above, $\alpha_{B,1}$ and $\alpha_{B,2}$ both have (amongst others) the initial segments $\langle 1, 2\rangle$ and $\langle 2, 1\rangle$. Now let an operation on bilateral sequences be $\Phi(\alpha_B) = 4$ for $\alpha_B \in \langle 1, 2\rangle$ and $\Phi(\alpha_B) = 5$ for $\alpha_B \in \langle 2, 1\rangle$. The axiom of open data, which we want to maintain from lawless sequences, says: $\Phi(\alpha) = x \rightarrow \exists n(\alpha \in n \wedge \forall \beta \in n\ \Phi(\beta)) = x$. From this we expect both $\alpha_{B,1}$ and $\alpha_{B,2}$ to get the same value assigned by Φ. However, since Φ is defined for both possible initial segments, it can not be ensured that Φ assigns the same value to both sequences. A related problem for operations on bilateral sequences is that they can in principle chose every initial segment around $\alpha_B 0$. This means that the value assigned to a sequence varies depending on which initial segment is chosen.[1]

Therefore, $\alpha_B \in n$ is also not defineable with the help of monolateral sequences without

[1]Restricting operations to finite initial segments that all differ in length and that are pairwise prefix-free does also not give a solution to the problem. Open data still fails. Consider

$$\Phi(\langle 1, 2, 3\rangle) = 5$$
$$\Phi(\langle 2, 3, 5, 7\rangle) = 9$$
$$\Phi(\langle 0, 9, 5, 6, 1\rangle) = 70$$

and

$$\alpha_{B,1} := \ldots, 0, 9, 5, 6, \underline{1}, 2, 3, \ldots$$
$$\alpha_{B,2} := \ldots, \underline{1}, 2, 3, 5, 7, \ldots$$

$\alpha_{B,1}$ and $\alpha_{B,2}$ share the initial segment $\langle \underline{1}, 2, 3\rangle$ and thus are in the open set generated by this finite sequence. However, for both sequences there is also another possibility for Φ to assign a value. Therefore, $\alpha_{B,1}$ and $\alpha_{B,2}$ are expected to both be assigned the same value and different values always relative to which initial segment happened to be chosen.

a root being fixed in finite sequences as well, i.e. the lengths of the inital segments of the monolateral sequences need to have a constant predefined length. This requires a new sequence coding and a general restriction on the form of bilateral sequences which is decidedly undesireable.

Apart from the problem with the initial segment relation and the definition of an initial segment of length n, to set a unique starting point for bilateral sequences does not lead to the most general definition possible. The entire discussion about the definition of an initial segment up to this point has been led with the assumption that we have *one* fixed root at $\alpha_B 0$. The fact that the fixed root is at $\alpha_B 0$ (and nowhere else) has never been used at all. Hence, the position of the root can be abstracted from and the discussion itself does not lose a crucial point.

Abstracting from the position of the root leads to the notion of a *relativised root*. Each and every position inside a bilateral sequence can be set as the general point of reference from which to look at the sequence.

No problems arise if relativised roots are worked with, i.e. if the main point of reference is changed during the choice process. Every *finite* subpart of a choice process can be an initial segment of a new choice process. The new process is still lawless, since only finitely many values are determined by the previously started choice process. Thus, adopting finitely many values of another sequence is the same as chosing these values as the deliberately placed initial segment. In other words: the choice of a new root can always be interpreted as starting a new process. This would eliminate a slightly ugly consequence of relativised roots, namely that operations on bilateral sequences are also relative to this root. I.e. a non-relativised operation will return different values for different roots and would thus be inconsistent (see below).

The possibility of using relativised roots also facilitates the characterisation of bilateral sequences. Since each and every value of a bilateral sequence can be chosen as root, the directions the information develops into can not be properly divided; each new choice of a root might change the position of a value from being part of the "left branch" to being part of the "right branch".

How can we work with bilateral sequences with relativized roots if such crucial notions as the initial segment relation can not be defined? A way forward is to restrict operations on bilateral sequences to operations on monolateral sequences belonging to a relativised root since the initial segment relation is in this case well-defined. The restriction of operations to monolateral sequences of course means that one entire direction of development is always ignored when a value of the operation is computed. Another consequence already mentioned above is that the operations on the monolateral branch are always relativised to the chosen root to maintain continuity.

Since the initial segment relation for bilateral sequences can not be defined, operations on bilateral sequences do not exist as such. Therefore, axioms $LS3$ (*open data*) and $LS4$ (*continuity*) can just be checked for operations on the monolateral parts. $LS1$ is also problematic since it uses the inital segment relation.

$$LS1 \qquad \forall n \exists \alpha (\alpha \in n)$$

Intuitively, $LS1$ (*density*) is still true for bilateral sequences. Any finite sequence can be set as block of starting values for the bilateral sequence just as for lawless sequences. There is a way, however, to restate $LS1$ for bilateral sequences using its monolateral parts.

$$LS1_B \qquad \forall n(\exists \alpha \exists j \exists \underleftarrow{\alpha j}(\underleftarrow{\alpha j} \in n) \wedge \exists \alpha \exists j \exists \underrightarrow{\alpha j}(\underrightarrow{\alpha j} \in n))$$

$LS1_B$ states that for every finite sequence there is a leftward and a rightward monolateral sequence of some bilateral sequence relative to some root such that this sequence is an initial segment of this sequence.

$LS2$ is entirely unproblematic for bilateral sequences and can be adapted straightforwardly. For the *decidability of equality*

$$LS2 \qquad \alpha \equiv \beta \vee \neg \alpha \equiv \beta$$
$$LS2' \qquad \alpha = \beta \vee \neg \alpha = \beta$$

the argumentation in chapter 1 can be adoped without change. $LS2$ is a direct consequence of the lawlessness of the choice process and does not depend on the type of choice process.

To discuss $LS3$ and $LS4$, it was already stated that it suffices to look at the operations on the monolateral parts of a bilateral sequence. The monolateral sequences are just like lawlike choice sequences. Therefore, the discussion about operations on lawless sequences can be adopted thanks to the *extension principle.*

A final word of caution concerns talking about the application of operations on bilateral sequences. Even though the operations on the monolateral parts can be seen as being applied to the bilateral sequences themselves - the operations are extensional - one functional can give potentially infinitly many different values for one bilateral sequence. This is the case since different relativised roots give different monolateral sequences which in turn all most probably differ in their initial segments.

2.2 Indexed Choice Sequences

Indexed choice sequences model choice processes where an index representing a class of information is assigned to each new value or piece of information added to the sequence. Thus the indices partition the continually increasing information into various independent branches or classes of information about the same object.

Definition 2.5. *Indexed choice sequences* α_I are choice sequences from $\mathbb{N}$ into $X \times I$, where X is a set of values and I is a set of indices.

Let X be $\mathbb{N}$ in this chapter. The set of indices can be potentially (countably) infinite; for simplicity's sake we will work with an enumeration of the elements of I, equating each element of I with a $n \in \mathbb{N}$.

In case the set of indices contains only one element, the indexed sequence is identical to a normal choice sequence. For cardinalities of I equal to or higher than two, the indexed sequences model choice processes with two or more branches where each index determines a "strand of information".

On the technical side, the addition of indices requires a new sequence coding that takes into account the different "flavours" given by the indices. A function on indexed sequences is given below.

Definition 2.6. Sequence coding for indexed sequences:

$$\phi(\alpha_I) := \prod_{\substack{j \in \mathbb{N} \\ \alpha_I(j) = \langle x_j, i_j \rangle}} p_{2j}^{x_j} p_{2j+1}^{i_j} \text{ with } p_k \in \mathbb{P}$$

ϕ is recursively defineable and thus compatible with the requirements put forward for a sequence coding in the first chapter. We also set $\phi(\langle\rangle) = 0$. By this definition, the value of an indexed choice sequence at 0 can be interpreted as the root of the sequence. For indexed sequences this is not of importance per se. The definition of a root will be needed in section 2.4, where the correlation of indexed sequences and bilateral sequences is discussed.

The notion of an initial segment of length n and the initial segment relation can be retained with the help of the above sequence coding. $\overline{\alpha_I}n$ is defined as the finite product

$$\prod_{\substack{j = 1 \ldots n \\ \alpha_I(j) = \langle x_j, i_j \rangle}} p_{2j}^{x_j} p_{2j+1}^{i_j}.$$

The initial segment relation

$$\alpha_I \in n \quad \text{iff} \quad \overline{\alpha_I}(\mathrm{lth} n) = n$$

is well-defined by virtue of the sequence coding being one-one.

To give a feeling for indexed sequences, without loss of generality, let the set of indices be $\{0, 1\}$. Then, an indexed choice sequence is a mapping $\alpha_I : \mathbb{N} \to \mathbb{N} \times \{0, 1\}$. We give an example of an indexed choice sequence in general and with concrete values:

$$\langle \langle x_0, i_0 \rangle , \langle x_1, i_1 \rangle , \langle x_2, i_2 \rangle , \ldots \rangle$$
$$\langle \langle 3, 0 \rangle , \langle 5, 1 \rangle , \langle 4, 0 \rangle , \langle 76, 1 \rangle , \langle 9, 1 \rangle , \langle 91, 0 \rangle \ldots \rangle$$

In the above example with concrete values there are two strands of information, one belonging to the index 0

$$3, 4, 91 \ldots$$

and one belonging to the index 1,

$$5, 76, 9, \ldots$$

The value $\langle 3, 0 \rangle$ is the root of the sequence.

We can also give a picture of casting a dice analogous to the one given for lawless sequences. For indexed sequences we have $|I|$ many different colored dices. After deliberately choosing an initial segment of value-color pairs, one after another a dice is chosen from the set of dices and is cast. The color of the dice and the value cast are then recorded together.

Let us take a look at the axioms formulated for lawless sequences $LS1 - LS4$ and show their compatibility with indexed sequences.

The first axiom, *density*, is valid for indexed sequences.

$$LS1 \qquad \forall n \exists \alpha (\alpha \in n)$$

n as (a sequence number of) a finite sequence is uniquely determined by the coding, and, since we again allow a deliberate choice of a finite initial segment, each such n occurs as one such segment. Thus the initial segment relation is well-defined and can be used as a definition for open sets of a topology on indexed choice sequences (cf. section 1.1).

For the *decidability of equality*

$$LS2 \qquad \alpha \equiv \beta \vee \neg\alpha \equiv \beta$$
$$LS2' \qquad \alpha = \beta \vee \neg\alpha = \beta$$

the argumentation can be adoped without change from the previous section.

Before $LS3$ and $LS4$ can be discussed, the type of operations compatible with indexed choice sequences need to be looked at. In the case of normal choice sequences, the operations on lawless sequences were continuous functionals $\Phi : \mathbb{N}^{\mathbb{N}} \to \mathbb{N}$ or $\Psi : \mathbb{N}^{\mathbb{N}} \to \mathbb{N}^{\mathbb{N}}$ represented by neighbourhood functions $a : \mathbb{N} \to \mathbb{N}$. The functionals (and respective neighbourhood functions) were both characterised in two different ways: as the set of continuous operations on lawless sequences or as a certain inductively defined class of mappings. As a consequence of the *bar theorem*, these two different strategies are shown to generate the same set of operations and the neighbourhood functions turn out to be in fact lawlike processes.

What are possible operations on indexed sequences? Parallel to lawless sequences, we can look at functionals $\Phi_I : (\mathbb{N} \times I)^{\mathbb{N}} \to \mathbb{N} \times I$ or $\Psi_I : (\mathbb{N} \times I)^{\mathbb{N}} \to (\mathbb{N} \times I)^{\mathbb{N}}$ and look whether we can find neighbourhood functions of the type $a_I : \mathbb{N} \times I \to \mathbb{N} \times I$. Or maybe the operations on lawless sequences can be adapted to indexed sequences, giving operations of the type $\Phi : (\mathbb{N} \times I)^{\mathbb{N}} \to \mathbb{N}$ or $\Psi : (\mathbb{N} \times I)^{\mathbb{N}} \to (\mathbb{N} \times I)^{\mathbb{N}}$ with neighbourhood functions $a : \mathbb{N} \to \mathbb{N}$.

The matter of operations on indexed sequences becomes clearer if the operations on lawless sequences are considered more thoroughly. The beauty of the operations on lawless sequences is that the neighbourhood functions inducing these operations turn out to be in fact lawlike choice sequences themselves. Of course it is desirable to reproduce this fact to be able further along to adopt results for operations on lawles sequences proved in chapter 1. Thus, neighbourhood functions need to have the same domain and range as indexed choice sequences. Consequently, functionals on indexed sequences are best of the following form.

$$\Phi_I : (\mathbb{N} \times I)^{\mathbb{N}} \to \mathbb{N} \times I$$
$$\Psi_I : (\mathbb{N} \times I)^{\mathbb{N}} \to (\mathbb{N} \times I)^{\mathbb{N}}$$

We will also define the functionals on indexed sequences relative to the two possible directions taken for functionals on lawless sequences in chapter 1.

Let $Cont_I$ be the class of continuous functionals on indexed sequences of the above form. The neighbourhood functions $K_{0,I}$ inducing these functionals are defined analogously to the ones for lawless sequences.

Definition 2.7. A function $\xi : \mathbb{N} \to (\mathbb{N} \times I)$ is in $K_{0,I}$ iff $k \in \mathbb{N}$ is a sequence number, i.e. $k = \langle \langle \alpha_0, i_0 \rangle, \ldots, \langle \alpha_p, i_p \rangle \rangle$ an initial segment of an indexed sequence α_I, and there is a functional $\Phi_\xi : (\mathbb{N} \times I)^{\mathbb{N}} \to \mathbb{N} \times I \in \ Cont_I$ such that

$$\xi k = \begin{cases} \langle 0, 0 \rangle & \text{the value of } \Phi_\xi(\alpha_I) \text{ is not yet determined} \\ \langle x+1, i \rangle & \Phi_\xi(\alpha_I) = \langle x, i \rangle \end{cases}$$

For the functions of $K_{0,I}$ the requirements of *consistency* and *totality* are fulfilled.

The second possibility to define functionals on indexed sequences is by an inductive definition. The class Ind_I of inductively defined functionals on indexed sequences is defined as follows.

Definition 2.8.

$$\begin{array}{lll} Ind_I & (i) & \lambda\alpha.\, \langle n, i \rangle \in Ind_I \quad \forall n \in \mathbb{N}, \forall i \in I \\ & (ii) & \Phi^{\langle 0,i \rangle}, \Phi^{\langle 1,i \rangle}, \Phi^{\langle 2,i \rangle} \ldots \in Ind_I \ \Rightarrow \Phi \in Ind_I,\ \Phi(\alpha) = \Phi^{(\alpha_I 0)}(\lambda x.\alpha_I(x+1)) \\ & & \text{or in other words:} \\ & & \forall x \forall i \in I(\Phi_{\langle x,i \rangle} \in Ind_I) \to \Phi \in Ind_I \quad \text{with } \Phi_n(\alpha_I) := \Phi(n * \alpha_I) \end{array}$$

The inductive definition of Ind_I can be packed into a principle of induction.

$$\forall n \forall i(\lambda\alpha.\, \langle n, i \rangle \in X) \wedge [\forall x \forall i(\Phi_{\langle x,i \rangle} \in X) \to \Phi \in X] \to Ind_I \subset X$$

The neighbourhood functions K_I of the inductively defined functionals are defined parallely.

Definition 2.9. K_I is the smallest class of functions $\xi : \mathbb{N} \to (\mathbb{N} \times I)$ that is closed under $K1_I$ and $K2_I$.

$$\begin{array}{ll} K1_I & \lambda n.\, \langle y+1, i \rangle \in K_I \quad \forall y \in \mathbb{N}, \forall i \in I \\ K2_I & \xi 0 = \langle 0, 0 \rangle \wedge \forall x \forall i \in I(\lambda n.\xi(\langle x, i \rangle \hat{} * n) \in K_I) \to \xi \in K_I \end{array}$$

Note, that $\xi 0$ is the value of the neighbourhood function assigned to the empty sequence which has sequence number 0. Setting this value equal to 0 ensures that the empty sequence is always too short for a functional to return a value.

The proof of proposition 1.10 which states that $K \subset K_0$ for operations on lawless sequences can be adapted to the newly defined continuous and inductively defined neighbourhood functions for indexed sequences since the actual form of the elements of the lawless sequnces is not used. Thus, with the appropriate changes to the proof of proposition 1.10, $K_I \subset K_{0,I}$. Whether $K_{0,I} \subset K_0$ holds depends on the applicability of the bar theorem to indexed sequences.

The class of continuous functionals $Cont^1_I$ of the form $\Psi_I : (\mathbb{N} \times I)^{\mathbb{N}} \rightarrow (\mathbb{N} \times I)^{\mathbb{N}}$ are also induced by $K_{0,I}$ parallel to the functionals $Cont^1$ for lawless sequences.

Having defined the functionals on indexed sequences, we return to the discussion of the axioms. $LS3$, the *axiom of open data*, is repeated with the usual notation.

$$\begin{aligned} LS3 \qquad & (\neq (\alpha, \alpha_0, \ldots, \alpha_p) \wedge A(\alpha, \alpha_0, \ldots, \alpha_p)) \rightarrow \\ & \exists n(\alpha \in n \wedge \forall \beta \in n(\neq (\beta, \alpha_0, \ldots, \alpha_p) \wedge A(\beta, \alpha_0, \ldots, \alpha_p)) \end{aligned}$$

The axiom is valid for the operations defined above. The neighbourhood functions inducing the functionals on indexed sequences are total and consistent and moreover depend only on a finite sequence of values. Consistency ensures that any value other than 0 returned by the function is and stays the same for all sequences belonging to the open set belonging to the shortest sequence for which this value is returned.

Since all inductively defined functionals are total and depend only on a finite initial segment and since the initial segment relation is well-defined, weak continuity is given. All other forms of continuity follow from the combination of $LS3$ with a choice principle that is in general intuitionistically valid (cf. section 1.3.3). Thus $LS4$ can adopted as well.

2.3 Bundled Choice Sequences

Bundled choice sequences are another take at modelling more than one strand of information about an object. Contrary to indexed sequences, the strands of bundled sequences are not time-independent. At every moment in time the same (finite) amount of information is added to all strands. This simultaneity and parallelism of the process is formalised by chosing tuples as values.

Definition 2.10. A bundled sequence $\alpha_T : \mathbb{N} \rightarrow X^m$ with $m \in \mathbb{N}$ is a sequence of tuples of elements of X an enumerable set.

For X we chose $\mathbb{N}$. Each element of the tuple belongs to a different strand of information; m thus codes the number of strands.

An example of a bundled sequence $\alpha_T : \mathbb{N} \to \mathbb{N}^m$, $m = 3$:

$$\langle 1,3,6\rangle, \ldots \quad \text{next choice: } \langle 10,5,7\rangle$$
$$\langle 1,3,6\rangle, \langle 10,5,7\rangle, \ldots \quad \text{next choice: } \langle 0,43,19\rangle, \langle 2,4,1\rangle$$
$$\langle 1,3,6\rangle, \langle 10,5,7\rangle, \langle 0,43,19\rangle, \langle 2,4,1\rangle, \ldots \quad \text{etc.}$$

Note that even though the values of each branch are chosen simultaneously, the values are not codependent. Bundled choice sequences are just like tuples of choice sequences, where each "dimension" is a separate process. Unfortunately, bundled sequences can not be defined by using lawless sequences. The proposition 12 proved in section 1.3.2 prevents any sequences defined in this way to be lawless themselves. The above definition is conceptually identical but does not suffer from this shortcoming.

To give a picture: bundled sequences come close to the simultaneous cast of m differently coloured dices. The succession of cast values is written up for each value separately. This analogy captures the fact that, even though each branch of information itself behaves like a lawless sequence, all m branches form one process.

Since bundled sequences are sequences of tuples as values, a new sequence coding is required.

$$\langle \alpha 0, \ldots, \alpha n\rangle := \sum_{j=1..n} p_{mj}^{\pi_1^m(\alpha j)} p_{mj+1}^{\pi_2^m(\alpha j)} \ldots p_{mj+m-1}^{\pi_m^m(\alpha j)}$$

with $p_i \in \mathbb{P}$ the ith prime number and π_i^m the projection onto the ith component of a tuple with m elements.

The values of the bundled sequence, $\alpha j = \langle x_{j,1}, \ldots, x_{j,m}\rangle$, are themselves coded as a natural number with the sequence coding used for lawless sequences defined in section 1.1.

With the new sequence coding, the initial segment relation and the initial segment of length n can be easily defined. For the initial segment of length n see the definition of the sequence coding above. The initial segment relation $\alpha_T \in n$ is defined as follows.

$$\alpha_T \in n \text{ iff } \langle \langle x_{01}, \ldots, x_{0m}\rangle, \ldots, \langle x_{\mathrm{lth}n,1}, \ldots, x_{\mathrm{lth}n,m}\rangle\rangle = n$$

Therefore, we expect all axioms for lawless sequences to be adoptable for bundled sequences and operations thereupon.

The first two axioms $LS1$ and $LS2$ are applicable to bundled sequences. The initial segment relation needed for *density* to be meaningful is well-defined and *decidability of identity* (or equality) is a direct consequence of lawlessness.

For the definition of operations on bundled sequences we follow the same line of argumentation given for indexed sequences in the previous section. To reproduce the desirable property of neighbourhood functions being lawlike bundled sequences, the functionals are of the following form.

$$\Phi_T : \mathbb{N}^{m\mathbb{N}} \to \mathbb{N}^m$$
$$\Psi_T : \mathbb{N}^{m\mathbb{N}} \to \mathbb{N}^{m\mathbb{N}}$$

The class of continuous functionals $Cont_T$ on bundled sequences is induced by the class of neighbourhood functions $K_{0,T}$.

Definition 2.11. A function $\xi : \mathbb{N} \to \mathbb{N}^m$ is in $K_{0,T}$ iff $k \in \mathbb{N}$ is a sequence number, i.e. $k = \langle\langle x_{00}, \ldots, x_{0,m-1}\rangle, \ldots, \langle x_{p0}, \ldots, x_{p,m-1}\rangle\rangle$ is an initial segment of a bundled sequence α_T, and there is a functional $\Phi_\xi : \mathbb{N}^{m\mathbb{N}} \to \mathbb{N}^m \in \ Cont_T$ such that

$$\xi k = \begin{cases} \langle 0, \ldots, 0\rangle & \text{the value of } \Phi_\xi(\alpha_T) \text{ is not yet determined} \\ \langle x_0 + 1, \ldots, x_{m-1} + 1\rangle & \Phi_\xi(\alpha_T) = \langle x_0, \ldots, x_{m-1}\rangle \end{cases}$$

The functions of $K_{0,T}$ fulfill both *consistency* and *totality*.

The inductively defined functionals Ind_T are formed by similar conditions as the functionals on indexed sequences.

Definition 2.12.

$$\begin{array}{lll} Ind_T & (i) & \lambda\alpha.\langle n_0, \ldots, n_{m-1}\rangle \in Ind_T \quad \forall \langle n_0, \ldots, n_{m-1}\rangle \in \mathbb{N}^m \\ & (ii) & \Phi^{\langle 0,\ldots,0\rangle}, \ldots, \Phi^{\langle 1,\ldots,1\rangle}, \ldots \in Ind_T \ \Rightarrow \Phi \in Ind_T \\ & & \Phi(\alpha_T) = \Phi^{(\alpha_T 0)}(\lambda x.\alpha_T(x+1)) \\ & & \text{or in other words:} \\ & & \forall \langle n_0, \ldots, n_{m-1}\rangle \, (\Phi_{\langle n_0,\ldots,n_{m-1}\rangle} \in Ind_T) \to \Phi \in Ind_T \\ & & \text{with } \Phi_n(\alpha_T) := \Phi(n * \alpha_T) \end{array}$$

The corresponding principle of induction is

$$\forall \vec{n}(\lambda\alpha.\vec{n} \in X) \wedge [\forall \vec{n}(\Phi_{\vec{n}} \in X) \to \Phi \in X] \to Ind_T \subset X$$

with $\vec{n} := \langle n_0, \ldots, n_{m-1}\rangle$ a value of a bundled sequence of dimension m.

K_T is defined as follows.

Definition 2.13. K_T is the smallest class of functions $\xi : \mathbb{N} \to (\mathbb{N} \times I)$ that is closed under $K1_T$ and $K2_T$.

$K1_T$ $\lambda n. \langle n_0 + 1, \ldots, n_{m-1} + 1 \rangle \in K_T \quad \forall \langle n_0, \ldots, n_{m-1} \rangle \in \mathbb{N}^m$

$K2_T$ $\xi 0 = \langle 0, \ldots, 0 \rangle \wedge \forall \langle n_0, \ldots, n_{m-1} \rangle (\lambda n. \xi(\langle n_0, \hat{\ldots}, n_{m-1} \rangle * n) \in K_T) \to \xi \in K_T$

Proposition 1.10 is also adaptable again, giving $K_T \subset K_{0,T}$.

The discussion of the last two axioms runs analogously to the discussion given in the previous section. $LS3$ and $LS4$ can both be adapted to the operations on bundled sequences.

2.4 Comparison

This part of chapter 2 aims to give a comparison of the three new types of choice sequences introduced in this chapter. The points of interest are:

- Are there translations from one type of sequence into the other types of sequences?
- What do the sequences model and capture idea-wise?
- Are there differences between the operations?

2.4.1 Indexed and Bilateral Sequences

At first we will investigate how indexed sequences and bilateral sequences are linked.

Let us start by giving a translation $t_{I \to B}$ from indexed sequences into bilateral sequences. With the choice of interpreting $\alpha_I(0)$ of an indexed choice sequence as root, sequences with an index set of cardinality $|I| = 2$ can be written parallel to bilateral sequences.

For technical reasons, we first need to define a projection π_i^j.

Definition 2.14. The function π_i^j for $i \leq j$ such that $\pi_i^j(\langle x_1, \ldots, x_j \rangle) = x_i$ is a projection onto the i-th component of a j-tuple.

The translation is then

$$t_{I \to B}(\alpha_I 0) = \alpha_B(0)$$

$$t_{I \to B}(\overline{\alpha_I} k * \alpha k) = \begin{cases} t_{I \to B}(\overline{\alpha_I} k), \pi_1^2 \alpha k & \pi_2^2 \alpha k = 1 \\ \pi_1^2 \alpha k, t_{I \to B}(\overline{\alpha_I} k) & \pi_2^2 \alpha k = 0 \end{cases}$$

The indices of the sequence values 0 and 1 are interpreted as “add the value to the left” and “add the value to the right”, respectively. The adresses of the values of the bilateral sequence are determined during the course of the recursion. This translation is not one-one, though, suggesting that the two types of sequences can not code the same amount of information about the choice process. The only requirement for indexed sequences to be mapped onto the same bilateral sequence is that the relative order of the values bearing the same index has to be identical. To illustrate this, we will look at indexed choice sequences of the type $\mathbb{N} \to \{0,1\} \times \{0,1\}$.

> The indexed choice sequences
> $\langle\langle 1,1\rangle, \langle 1,0\rangle, \langle 0,1\rangle, \ldots\rangle$
> and
> $\langle\langle 1,0\rangle, \langle 1,1\rangle, \langle 0,1\rangle, \ldots\rangle$
> can both be written as the following bilateral sequence
> $\ldots, 1, 1, 0, \ldots$

As the above example shows, if the values are recoded respecting only this translation rule, there is no one-to-one relation between indexed choice sequences and bilateral choice sequences. The translation rule ignores both, the order of information gain encoded in indexed choice sequences and the position of the root of these sequences. The latter information about the sequence can be included in the translation if the root of the indexed choice sequence is mapped to the relativised root of the bilateral choice sequence. The order of information gain, however, cannot be coded in bilateral sequences, as already stated in section 2.1. Therefore, the best translation available is as follows.

$$t_{I\to B}(\alpha_I 0) = \underline{\alpha_B(i)}$$

$$t_{I\to B}(\overline{\alpha_I}k * \alpha k) = \begin{cases} t_{I\to B}(\overline{\alpha_I}k), \pi_1^2\alpha k & \pi_2^2\alpha k = 1 \\ \pi_1^2\alpha k, t_{I\to B}(\overline{\alpha_I}k) & \pi_2^2\alpha k = 0 \end{cases}$$

where $\underline{\alpha_B(i)}$ is the relativised root of the bilateral sequence

As an example:

> The indexed choice sequence
> $\langle\langle 1,1\rangle, \langle 1,0\rangle, \langle 0,1\rangle, \ldots\rangle$
> is written as
> $\ldots, 1, \underline{1}, 0, \ldots$

and another indexed choice sequence

$\langle\langle 1, 0\rangle, \langle 1, 1\rangle, \langle 0, 1\rangle, \ldots\rangle$

is written as

$\ldots, \underline{1}, 1, 0, \ldots$

where the underlined value is the relativised root of the sequence.

Since the notion of relativised root is inherently dynamic, the addition of the position of the root in bilateral sequences is only a minor addition; the improved translation is still not one-one. Two or more indexed sequences are translated into the same bilateral sequence if they have the same root and if the relative order of the values bearing the same index is identical.

By way of the translation from indexed sequences to bilateral sequences we now know that operations on bilateral sequences (i.e. the operations on the monolateral parts) can be adapted to the translated indexed sequence. The translation is in fact a nice way of getting hold of all values of the same index. The operations on the monolateral sequences can then be used to map only one branch of information onto some desired value.

The translation from bilateral sequences to indexed sequences $t_{B\to I}$ is a little trickier. Bilateral sequences do not code the order of information gain. Therefore, there are many possible translation from a given bilateral sequence of values (i.e. from already chosen values) into indexed sequences. The only exact way to translate bilateral sequences into indexed sequences is thus online, i.e. translating the values as they are chosen. Another (less accurate) way to go would be to fix an order of translation.

$t_{B\to I}(\alpha_B n) = \alpha_I x$ where x is the position of n in the translation order and

$\alpha_I x = \langle \alpha_B n, 0\rangle$ iff $\alpha_B n$ is left to the relativised root

$\alpha_I x = \langle \alpha_B n, 1\rangle$ iff $\alpha_B n$ is right to the relativised root

When the online translation method is chosen, the translation is one-one and onto since the online translation ensures that the order of information gain is reflected in the indexed sequence. A previously fixed arbitrary translation order makes the translation lose these properties.

The online translation then obviously allows operations on indexed sequences to be applied to the translation of the bilateral sequence. If a fixed translation order is chosen, the product of the translation is still an indexed sequence in its make up. The order of information gain has however been destroyed. Operations on indexed sequences can of course be applied to these sequences. It is, however, unclear what the the meaning of such an application is and whether the resulting values reflect any property of the indexed or the bilateral sequence.

2.4.2 Indexed and Bundled Sequences

Bundled and indexed sequences both describe a simultaneous development of a sequence in m different directions/ belonging to m different classes of information. The crucial difference is that indexed sequences allow all branches to develop independently whereas for bundled sequences all branches develop uniformly.

Can a translation from bundled sequences to indexed sequences be given or vice versa? Take $\alpha_T : \mathbb{N} \to \mathbb{N}^m$. Then α_T can be translated into an indexed sequence $\alpha_I : \mathbb{N} \to \mathbb{N} \times I$ with $|I| = m$. The translation $t_{T \to I}$ is as follows:

$$t_{T \to I}(\alpha_T j) = \alpha_I(mj), \ldots, \alpha_I(mj + m - 1)$$

where

$$\alpha_I(mj) = \langle \pi_0^m(\alpha_T j), 0 \rangle \; \alpha_I(mj + 1) = \langle \pi_1^m(\alpha_T j), 1 \rangle \text{ etc.}$$

The translation does not map bundled sequences to general indexed sequences as defined in section 2.2. I.e. not every indexed sequence has a corresponding bundled sequence. The indexed sequences into which the bundled sequences are translated are those where in an initial segment of length $m \cdot j$ all m different indices appear j times.[2]An example:

$$\langle 1, 2, 3 \rangle, \langle 5, 6, 7 \rangle, \langle 9, 10, 11 \rangle, \ldots$$

is translated into

$$\langle 1, 0 \rangle, \langle 2, 1 \rangle, \langle 3, 2 \rangle, \langle 5, 0 \rangle, \langle 6, 1 \rangle, \langle 7, 2 \rangle, \langle 9, 0 \rangle, \langle 10, 1 \rangle, \langle 11, 2 \rangle, \ldots$$

As a translation can be given from bundled sequences into indexed sequences, the operations on indexed sequences are also applicable to bundled sequences.

Can an inverse translation from indexed sequences into bundled sequences be given? As it was argued above, the translation of bundled sequences only targets a subset of all indexed sequences. Therefore it is to be expected that a ready translation from indexed sequences into bundled sequences can not be given. The main issue for giving a general translation is that the amounts of values of each index in an indexed sequence

[2]The fact that the translation orders the indices inside the indexed sequence is neglectable but not without importance. Differently ordered values of indexed sequences model a different process of information gain. Every reordering changes the initial segment of the sequence which determines the position of the sequence in the universe of indexed sequences. Nevertheless the ordering of the indices is neglectable since the observation that the translation of bundled sequences into indexed sequences only targets a subset of all indexed sequences remains intact when different orderings are considered.

differ whereas for bundled sequences all amounts of values are required to be the same. Thus "filling in" the values of the indexed sequence in tuples "as they come" does not guarantee well-formed bundled sequences.

The following example illustrates this.

> $\langle 7,1\rangle\,,\langle 5,2\rangle\,,\langle 9,0\rangle\,,\langle 10,0\rangle\,,\langle 11,1\rangle\,,\langle 13,0\rangle\,,\langle 4,0\rangle\,,\ldots$
>
> results in the following bundled sequence by the filling in method:
>
> $\langle 9,7,5\rangle\,,\langle 10,11,?\rangle\,,\langle 13,?,?\rangle\,,\langle 4,?,?\rangle\,,\ldots$

Since the last three tuples stay unfinished, they are not well-formed values of a bundled sequence and hence useless for operations. We can also not be certain that these tuples might get "filled up" because of the lawlessness of the process.

Therefore, operations on bundled sequences are not readily applicable to indexed sequences, as expected.

The picture the above discussion paints is, thus, that bundled sequences are indexed sequences with a restriction on the occurrence of indices.

2.4.3 Bilateral and Bundles Sequences

In this section we will take advantage of the results of the previous two sections.

Concerning the translation $t_{T\to B}$ from bundled into bilateral sequences, we already know the translations $t_{T\to I}$ and $t_{I\to B}$. Therefore, we easily obtain $t_{T\to B}$ by a successive application of $t_{T\to I}$ and $t_{I\to B}$. A direct consequence is that the operations on bilateral sequences can be adapted to bundled sequences.

We also expect that there should be no translation $t_{B\to T}$ from bilateral sequences into bundled sequences since there is no $t_{I\to T}$. To see that this is really the case, remember that choice sequences are processes and that operations on processes depend on a finite initial segment.

Let α_B be the following process at a fixed moment in time:

$$\ldots,5,4,\underline{1},2,3,\ldots$$

Then the first idea for a translation from bilateral to bundled sequences could be to decide that the relativised root of the bilateral sequence and everything to its left belongs to the first branch and everything to its right belongs to the second branch of the bundled sequence, and then to form pairs to give the actual values of the bundled sequence.

$$t_{B\to T}(\ldots,5,4,\underline{1},2,3,\ldots)=\langle 1,2\rangle\,,\langle 4,3\rangle\,,\langle 5,?\rangle\,,\ldots$$

Parallel to the issue that was discussed for the translation from indexed sequences into bundled sequences, it can not be ensured that the missing values will be filled.

A consequence concerns the operations on bilateral sequences. We cannot adapt an operation on bundled sequences to bilateral sequences. Since the translation from bilateral sequences into bundled sequences does not guarantee a well-formed initial segment (i.e. the translation can result in a long succession of unfinished tuples), an operation might never become applicable to the translated bilateral sequence (the neighbourhood function of this operation always gives the value 0). This violates both totality and continuity of the operation.

2.5 Summary

In this chapter we have looked at three different modifications or generalisations of lawless choice sequences. We have briefly discussed their compatibility with the axioms for lawless sequences defined in chapter 1. For this matter we also had to consider operations on the different new types of sequences and their properties.

We then furthermore compared the three new types of sequences and tried to give translations from one type to another, where possible. A consequence of these translations was that operations of one sequence type could be adapted to all other types of sequences the former type could be translated into. Even though we are always talking about "adapting" an operation to another class of sequences, we actually want to have a translation from the various operation types into the other types, when possible. That this translation between operations can be given with the help of the translations given for the sequences should be quite obvious.

Chapter 3

Choice Sequences and Possible Worlds

The property of lawless choice sequences of being incomplete, potentially infinite objects make them an interesting tool to apply to problems in other branches of science. The objective for this chapter will be to use lawless sequences in a framework for possible world semantics of natural languages and to take a look at the analysis of modality.

3.1 Possible-Worlds Semantics in Linguistics

In formal linguistics today, there are quite a few different ways to deal with natural language semantics. One rather common approach goes back to Richard Montague, who gave a formal system for a fragment of English (Montague (1973) amongst others) This system encompasses both the syntax and the semantics of this fragment. His work on semantics had the most impact on modern day linguistics and provided the basis for an influential framework in modern formal natural language semantics (Partee 2005).

When Montague's work was introduced into linguistics, it not only provided interesting new insights in the semantics of natural language sentences but also a way to deal with the denotation of words without getting caught in prototype theory.[1]

What is the goal of this framework? Older approaches just wanted to give a translation of the content of a natural language expression into a formal language (mostly first order classical logic). Montague not only parted with first order logic, he also capitalized Frege's *principle of compositionality* which states that the meaning of a complex expression is built up from the meanings of its parts: the meaning of a sentence like **Peter sleeps** is built up from the meaning of **Peter** and **sleeps** by some

[1] Prototype theory deals with graded categorization. Some members of a category expressed by some linguistic expression are more typical for this category than other members. E.g. a chair might be a more typical piece of furniture than a bed. Judgements of this kind make it difficult to define what the meaning of "furniture" is if typical features are taken as the basis of a theory of meaning.

kind of combination mechanism. By providing the structure of models for his typed intensional language together with an interpretation function, he gave an algorithmic way of deriving the *truth condition* and (given a specific model) the *truth value* of a sentence. The truth condition of a sentence is the minimal set of requirements for the world (given by a model) such that the sentence is true in this world.

We will first give Montague's intensional language as it stands in "Proper Treatment of Quantification in Ordinary English" (Montague 1973). Then we will discuss a few additions and changes that were made up until now that will be needed for the discussion of Kratzer's analysis of modality (Kratzer 1991).

3.1.1 Montague's Intensional Language

In "Proper Treatment of Quantification in Ordinary English", Montague independently sets up the syntax and the semantics for his chosen fragment of English and formulates translation rules that map English expressions from his syntax onto expressions of his semantics. As a semantic language, he formulates a higher-order typed intensional language with a possible-worlds model-theoretic semantics, based on the idea of intensional languages as put forth by Frege (1892), Church (1951) and Carnap (1947) amongst others.

The idea of intensional languages is that even though some natural language expressions or expressions in Mathematics can denote the same thing, their meaning can still be different. For example $5-2$, $2+1$ and 3 denote the same number *three*. The mathematical operations performed, however, differ greatly. In Frege's terminology $5-2$, $2+1$ and 3 have the same *Bedeutung* (reference) but not the same *Sinn* (sense). Frege's *Sinn* of an expression was called the *intension* of the expression, Frege's *Bedeutung* the *extension* of an expression by Carnap. In the standard interpretation of the numbers and operations it is true to write $5-2=2+1=3$. Thus, in the standard model of natural numbers, the three expressions have the same extension. However, it is easy to think of an algebra where $5-2$, $2+1$ and 3 denote quite different values. So the expressions do not have the same intension.

The above idea is formalised by an intensional language that distinguishes the two layers of extensionality and intensionality and that has an appropriate model structure for this task. Intensional contexts in natural language, such as "Peter believes that..." or "it is necessary that...", motivate a possible-worlds semantics. In intensional contexts truth is not preserved when an expression is substitued by another expression with the same extension, e.g. even if "Peter believes that 2+1 = 3" is true "Peter believes that 5-2 = 3" does not have to be, if Peter can add but not subtract properly.

Possible worlds model more or less conceivable alternative states of affairs. With possible worlds, it is possible e.g. to refer to all the worlds that are compatible with Peter's belief concerning basic calculations or to model necessary truths or possible truths. *Montague's intensional language* (IL_M)[2] incorporates a typed lambda-calculus as proposed by Church and Kripke frame semantics as used in modal or intuitionistic logic. The following definitions are taken from Montague (1973:227ff).

Definition 3.1. The *type system* of IL_M:
Basic types are e for entity expressions (individuals) and t for truth value expressions. The set *Type* is the smallest set Y such that (1) $e, t \in Y$, (2) if $a, b \in Y$ then $\langle a, b\rangle \in Y$ and (3) if $a \in Y$ then $\langle s, a\rangle \in Y$.

Montague differentiates between intensional and extensional (non-intensional) types. The basic types and the types built from clause (2) are the *extensional fragment* of the type system that applies to all those natural language expressions that have no reference to possible worlds, e.g. connectives like **and** and **or**[3] that are constant across worlds. Intensional types are used for all those expressions that depend on possible worlds like modal expessions such as modal verbs **must** and **can**. [4]

Definition 3.2. The *meaningful expressions* of IL_M of type a, ME_a:

1. Every variable and constant of type a is in ME_a.
2. If α is in ME_a and u is a variable in ME_b then $\lambda u\alpha$ is in $\mathrm{ME}_{\langle b,a\rangle}$ (λ-abstraction)
3. If α is in $\mathrm{ME}_{\langle b,a\rangle}$ and β is in ME_b then $\alpha(\beta)$ is in ME_a (function application)
4. If α, β are in ME_a then $\alpha = \beta$ is in ME_t
5. If ϕ, ψ are in ME_t and u is a variable in ME_a then also $\neg\phi$, $\phi \vee \psi$, $\phi \wedge \psi$, $\phi \rightarrow \psi$, $\phi \leftrightarrow \psi$, $\forall u_a \phi$, $\exists u_a \phi$, $\Box\phi$, $W\phi$, $H\phi$
6. If α is in ME_a then $\hat{}\alpha$ is in $\mathrm{ME}_{\langle s,a\rangle}$
7. If α is in $\mathrm{ME}_{\langle s,a\rangle}$ then $\check{}\alpha$ is in ME_a

$\lambda u\alpha$ denotes the function that takes arguments of the same type as the variable u and has as a value the expression α with all occurrences of u substituted by the argument. $\Box\phi$, $W\phi$ and $H\phi$ are the necessity, future and past operators. The $\hat{}$-operator returns the intension of its argument; the $\check{}$-operator returns the extension of its argument and is thus only well-defined in cases where the argument is an intension.

Definition 3.3. The *model structure* of IL_M:
A model of IL is a quintuple $\mathfrak{A} = (A, I, J, \leq, F)$, where A is a set of entities, I a set of possible worlds, J a set of moments of time, $\leq$ an ordering on J and F an

[2]Monatgue's intensional language is usually abbreviated IL. I changed it to IL_M since intuitionistic logic is also abbreviated as IL and keeping both unchanged might lead to confusion.

[3]Occurrences of natural language as object language will be signalled by **bold face**.

[4]In principle it is possible to just work with the extensional fragment in natural language semantics. However, as Montague shows in his works himself, pure extensionality is the exception in natural language and reference to possible worlds the rule.

interpretation function assigning all non-logical constants their interpretation relative to a world-time-pair.

Since intensional models like above make use of possible worlds, this type of model theoretic semantics is called a *possible-worlds semantics*.[5] A model for the extensional fragment would consist only of A and F since there is no reference to worlds, i.e. $\mathfrak{A} = (A, F)$.

For all non-logical constants of each type there is a domain of possible denotations in the model. The domain of possible denotations for constants of type a corresponding to A, I and J will be written as $D_{a,A,I,J}$: $D_{t,A,I,J}$ is the set of truth values $\{0, 1\}$[6] , $D_{e,A,I,J}$ is the set of entities A, $D_{\langle b,a\rangle,A,I,J}$ is the set of functions from $D_{a,A,I,J}$ into $D_{b,A,I,J}$ and $D_{\langle s,a\rangle A,I,J}$ is the set of functions from the cartesian product of worlds and moments $I \times J$ into $D_{a,A,I,J}$.

For the interpretation of IL_M, the interpretation function F is extended to a function $[\![.]\!]$ from complex well-formed expressions into the set of all possible domains of denotations.[7]

Definition 3.4. The *interpretation function* $[\![.]\!]^{\mathfrak{A},i,j,g}$:
$[\![\alpha]\!]^{\mathfrak{A},i,j,g}$ denotes the meaning of α relative to a model $\mathfrak{A}$, an assignment function g from variables of type a into $D_{a,A,I,J}$, a world i and a time j.

1. If α is a non-logical constant then $[\![\alpha]\!]^{\mathfrak{A},i,j,g} = F(\alpha)(\langle i, j\rangle)$
2. If α is a variable then $[\![\alpha]\!]^{\mathfrak{A},i,j,g} = g(\alpha)$.
3. If α is in ME_a and u is a variable in ME_b then $[\![\lambda u\alpha]\!]^{\mathfrak{A},i,j,g} = h : D_{b,A,I,J} \to D_{a,A,I,J}$ such that for all x in that domain $h(x) = [\![\alpha]\!]^{\mathfrak{A},i,j,g'}$ where g' is the assignment function g except for possibly $g'(u) = x$ (λ-abstraction)
4. If α is in $\mathrm{ME}_{\langle b,a\rangle}$ and β is in ME_b then $[\![\alpha\beta]\!]^{\mathfrak{A},i,j,g} = [\![\alpha]\!]^{\mathfrak{A},i,j,g}([\![\beta]\!]^{\mathfrak{A},i,j,g})$ (function application)
5. If α, β are in ME_a then $[\![\alpha = \beta]\!]^{\mathfrak{A},i,j,g} = 1$ iff $[\![\alpha]\!]^{\mathfrak{A},i,j,g} = [\![\beta]\!]^{\mathfrak{A},i,j,g}$
6. If ϕ, ψ are in ME_t then $[\![\neg\phi]\!]^{\mathfrak{A},i,j,g}$, $[\![\phi \vee \psi]\!]^{\mathfrak{A},i,j,g}$, $[\![\phi \wedge \psi]\!]^{\mathfrak{A},i,j,g}$, $[\![\phi \to \psi]\!]^{\mathfrak{A},i,j,g}$, $[\![\phi \leftrightarrow \psi]\!]^{\mathfrak{A},i,j,g}$ are defined as usual
7. If ϕ is in ME_t and u is a variable in ME_a then $[\![\forall u_a\phi]\!]^{\mathfrak{A},i,j,g} = 1$ iff there is an $x \in D_{a,A,I,J}$ such that $[\![\phi]\!]^{\mathfrak{A},i,j,g'} = 1$ where g' is as above; similarly $[\![\exists u_a\phi]\!]^{\mathfrak{A},i,j,g}$
8. If ϕ is in ME_t then $[\![\Box\phi]\!]^{\mathfrak{A},i,j,g} = 1$ iff $[\![\phi]\!]^{\mathfrak{A},i',j',g} = 1$ for all $i' \in I$ and $j' \in J$.
9. If ϕ is in ME_t then $[\![W\phi]\!]^{\mathfrak{A},i,j,g} = 1$ iff $[\![\phi]\!]^{\mathfrak{A},i,j',g} = 1$ for some $j' \in J$ such that $j \leq j'$ and $j' \neq j$.
10. If ϕ is in ME_t then $[\![H\phi]\!]^{\mathfrak{A},i,j,g} = 1$ iff $[\![\phi]\!]^{\mathfrak{A},i,j',g} = 1$ for some $j' \in J$ such that $j' \leq j$ and $j' \neq j$.
11. If α is in ME_a then $[\![{}^\wedge\alpha]\!]^{\mathfrak{A},i,j,g} = h(\langle i, j\rangle)$ such that for all $\langle i, j\rangle \in I \times J$ $[\![{}^\wedge\alpha]\!]^{\mathfrak{A},i',j',g} = h(\langle i', j'\rangle)$
12. If α is in $\mathrm{ME}_{\langle s,a\rangle}$ then $[\![{}^\vee\alpha]\!]^{\mathfrak{A},i,j,g} = [\![\alpha]\!]^{\mathfrak{A},i,j,g}(\langle i, j\rangle)$

[5]The term *possible-worlds semantics* is also used for the branch of natural language semantics that uses this type of model.

[6]The truth values 0 and 1 are logical constants. Instead of $\{0, 1\}$, $\{\top, \bot\}$ can be used, but is unusual in the formal semantics tradition in linguistics.

[7]Montague does not use the double brackets. We will use them for readability reasons.

The parameters i and j of the interpretation function are a world from the set of possible worlds and a moment of time, i.e. a world-time pair $\langle i, j \rangle \in I \times J$ in the model $\mathfrak{A}$, relative to which the interpretation is started. The pair $\langle i, j \rangle$ usually is the *actual world*, i.e. the real world in which the interpreted sentence was uttered at the time it was uttered. The entire cartesian product $I \times J$ is called the *logical space*.

The rule of *function application* is the central rule of IL_M's syntax and semantics. With this rule Montague formalises the *principle of compositionality*.

Definition 3.5. Truth in IL_M:
if ϕ is in ME_t, i.e. a formula, then ϕ is true with respect to $\mathfrak{A}, i$ and j iff $[\![\phi]\!]^{\mathfrak{A},i,j,g}$ is true for all assignments g.

Meaningful expressions of the intensional logic denote concepts needed for the interpretation of natural language. Expressions of type $\langle a, t \rangle$ are characteristic functions of subsets and thus denote sets of objects of type a. Similarly, expressions of type $\langle a, \langle b, t \rangle \rangle$ can be seen as two-place relations between objects of type a and b, etc.

Why are these expressions interesting for natural languages? As stated above Montague gave a nice formalisation of natural language expressions. The central idea of the formalisation is that an expression denotes the set of objects it successfully describes. For example: The noun **dog** denotes the set of dogs in the model of choice. The intransitive verb **sleep** denotes the set of sleeping individuals in the model. Transitive verbs like **love** denotes the set of pairs of lovers and loved ones in the model and so on. Thus the meaning of **dog** is the characteristic function of the set of dogs relative to a model.[8] In terms of types this means that the meanings of **sleep** and **dog** are expressions of type $\langle \langle s, e \rangle, t \rangle$.

We won't discuss the details of the application to natural language as the main interest for this thesis is the structure of the intensional model, which will be discussed in detail in section 3.2.

3.1.2 Modernisations of IL_M

After the works of Montague were introduced into linguistic research the above system was partly modified to fit the needs of linguists. Barbara Partee (Partee 1973) laid the groundwork to adapting the intensional language as a semantic framework compatible with the type of syntax done in the generative grammar tradition, which in the 1970s was the last phase of transformational grammar (Chomsky 1957).

[8] Montague already builds this idea into the categorial grammar modelling the English syntax. However, his system of categories is (apart from the intensional types) identical to the type system of the intensional language. The translation rules from syntax into semantics only build intensions from the exclusively extensional categories. For the detailed formulation see Montague (1973).

During the 1970s and 80s the intensional language (or rather its application to natural language and theories of natural language syntax) was occasionally enriched with new rules and types but also intensionality was made easier to handle and some types proposed for of verbal categories were altered. We will only focus on changes made by semanticists working in or closely to generative grammar.

For example, a new intensional type i analogous to s was introduced to get direct access to the time parameter of the model, used usually when only matters of tense were important and reference to possible worlds was irrelevant. Also, for exclusively modal purposes, the model structure is usually simplified to a quadruple $M = (D, W, \leq, F)$ where D is a set of individuals, W a set of possible worlds, $\leq$ an order on W giving a temporal ordering on possible worlds and F an interpretation function. Possible worlds in this model structure are something akin to the world-time pairs in IL_M. However, in cases when even the temporal ordering $\leq$ is left away from the model, possible worlds are interpreted as worlds specified for all past and future moments in time.[9] The parameters of the interpretation function $[\![\alpha\beta]\!]^{\mathfrak{A},i,j,g}$ are now usually $[\![\alpha\beta]\!]^{w,M,g}$; a world $w \in W$, the model M and an assignment function g. In some modern works a time parameter t is again added.

Concerning intensionality per se, there are currently two different opinions about using intensional types and reference to possible worlds. One possibility is to see intensions as the default semantic value of an expression in the spirit of Montague. These accounts also differ with respect to the possibility of having intensional types as parts of complex types or not. The other possibility is to work extensionally most of the time and extend the system in a conservative way to account for intensional contexts. Both treatments of intensionality give the desired results. For both treatments also the question of overt world variables in the system is discussed. For reasons of space, the discussion can not be repeated here.

A new principle of composition called *θ-identification* was first introduced by Higginbotham (1985) (later in a restricted form it was called *predicate modification* in Heim and Kratzer (1998)).

Definition 3.6. *θ-identification*:

- If α is in $\mathrm{ME}_{\langle e,a\rangle}$ and β is in $\mathrm{ME}_{\langle e,a\rangle}$ then also $\alpha\beta$ is in $\mathrm{ME}_{\langle e,a\rangle}$
- If α is in $\mathrm{ME}_{\langle e,a\rangle}$ and b is in $\mathrm{ME}_{\langle e,a\rangle}$ then $[\![\alpha\beta]\!]^{w,M,g} = \lambda u_e[[\![\alpha]\!]^{w,M,g}(u) \wedge [\![\beta]\!]^{w,M,g}(u)]$

Predicate modification as defined in Heim and Kratzer (1998) is the restriction of θ-identification to cases of $a = t$.

Now, to give a better idea of how the typed intensional language is used in formal

[9]This can be seen as a reflex to Montague's definition of necessity as "necessarily always".

semantics, we will list the extensional and intensional types of some verbal categories. Remember that natural language expressions denote the set of objects they describe.

- sentence: t (truth value)
- proper name: e (an individual, e.g. **Peter**, **Rome**...)[10]
- noun: $\langle e, t \rangle$ (a set of individuals, e.g. **dog**, **table**...)
- intransitive verb: $\langle e, t \rangle$ (a set of individuals, e.g. **sleep**, **sneeze**...)
- transitive verb: $\langle e, \langle e, t \rangle \rangle$ (a set of pairs of individuals, e.g. **kiss**, **love**...)
- relational noun: $\langle e, \langle e, t \rangle \rangle$ (a set of pairs of individuals, e.g. **mother**, **brother**...) etc.

Nouns and verbs among others are called the *predicates* of the extensional fragment. The types listed here are as given in Heim and Kratzer (1998).

Interestingly, the extensional denotation of a sentence is its truth value. This is an artefact of function application and the choice of letting verbs denote sets of individuals (or tuples of individuals). Since the verb meaning is given as the characteristic function of the set it denotes, once the function has been applied to all arguments (i.e. once the sentence is complete) the verb meaning returns either true or false if the tuple of individuals is in the set. Since natural language sentences are only well-formed when all verb arguments are given, the extensional denotation will always be a truth value.

Next we list the intensional types built from the extensional types above.

- sentence: $\langle s, t \rangle$ (set of worlds)
- proper name: $\langle s, e \rangle$ (a relation between worlds and individuals)
- noun: $\langle s, \langle e, t \rangle \rangle$ (a relation between worlds and sets of individuals)
- intransitive verb: $\langle s, \langle e, t \rangle \rangle$ (a relation between worlds and sets of individuals)
- transitive verb: $\langle s, \langle e, \langle e, t \rangle \rangle \rangle$ (a relation between worlds and sets of pairs of individuals)
- relational noun: $\langle s, \langle e, \langle e, t \rangle \rangle \rangle$ (a relation between worlds and sets of pairs of individuals) etc.

The intensional denotation of a sentence - called a *proposition* - is the set of worlds for which the sentence is true. The intensional denotations of the other elements also have traditional names. The intensional denotation of proper names is called an *individual concept*, the one of nouns and verbs is called a *property*.

Note that the rule of function application, as it stands now, is not suitable for the intensional types listed above.[11] Therefore it will be reformulated as follows.

[10] For Monatgue proper names have the type $\langle \langle e, t \rangle, t \rangle$, same as quantifiers like **every** and **no**. For various reasons the type of proper names was changed into e.

[11] Montague's formulation still works for the extensional fragment.

Definition 3.7. Intensional function application: If α is in $\mathrm{ME}_{\langle s,\langle b,a\rangle\rangle}$ and β is in $\mathrm{ME}_{\langle s,b\rangle}$ then $[\![\alpha\beta]\!]^{w,M,g} = {}^{\wedge}[\![{}^{\vee}\alpha]\!]^{w,M,g}([\![{}^{\vee}\beta]\!]^{w,M,g})$ is in $\mathrm{ME}_{\langle s,a\rangle}$

For the definition the $^{\wedge}$-operator and the $^{\vee}$-operator were used. Intensional function application could also be defined without these two operators by using overt world variables. The result is the same.

The goal of this framework of semantics is to algorithmically derive the truth conditions of a sentence. Thus, we will derive the truth condition of the sentence **Peter can see Susi**. The types of the meanings of the lexical entries are: $D_{\langle s,e\rangle}$ for **Peter**,$D_{\langle s,e\rangle}$ for **Susi** and $D_{\langle s,\langle e,\langle e,t\rangle\rangle\rangle}$ for **see**. For the modal expression Montague's treatment of modality is kept for the moment and **can** is translated as the diamond $\Diamond$, which is taken as the usual abbreviation of $\neg\Box\neg$.

$$\Diamond\mathbf{see}(\mathbf{Susi})(\mathbf{Peter})$$

[12] The meanings of the parts is then as follows. Overt world variables are used to make the types of the expressions visible.

$[\![\mathbf{Peter}]\!]^{w,M,g} = \lambda w.\ \text{Peter}$

$[\![\mathbf{Susi}]\!]^{w,M,g} = \lambda w.\ \text{Susi}$

$[\![\mathbf{see}]\!]^{w,M,g} = \lambda w.\lambda y.\lambda x.\ x \text{ sees } y \text{ in } w$

Then the truth condition of the entire sentence is:

$[\![\Diamond\mathbf{Peter\ see\ Susi}]\!]^{w,M,g} = 1$ iff $[\![\mathbf{Peter\ see\ Susi}]\!]^{w',M,g} = 1$ for some $w' \in W$

$[\![\mathbf{Peter\ see\ Susi}]\!]^{w',M,g} = 1$ iff $[\![\mathbf{see(Susi)(Peter)}]\!]^{w',M,g} = 1$

$[\![\mathbf{see(Susi)(Peter)}]\!]^{w',M,g} = 1$ iff Peter sees Susi in w'

thus:

$[\![\Diamond\mathbf{Peter\ see\ Susi}]\!]^{w,M,g} = 1$ iff Peter sees Susi in w' for some $w' \in W$

The first step was application of the interpretation of the diamond, the second step was giving the functional structure of the sentence and the third and last step was function application of the meaning of **see** to the meanings of **Susi** and **Peter**. Note that the meaning of proper names is invariant across worlds.[13]

[12]For natural language syntactic reasons the direct object **Susi** is the first argument of see and the subject **Peter** the second.

[13]We choose to make proper names world invariant to escape the problem of transworld identity.

3.1.3 Kratzer's analysis of modality

In this section the most prominent and natural application of possible-worlds semantics, the analysis of modal expressions like e.g. modal verbs (e.g. **must**, **can**...) or modal adverbs (e.g. **possibly**, **necessarily**...) will be given. Examples of sentences that will have to be analysed and given an appropriate representation are in (1).

(1) a. Peter must read a book.
  b. It is possible that Peter reads a book.

In Montague's system IL_M the sentences in (1) would have the following meaning.

(2) a. $[\![\Box$ Peter read a book$]\!]^{\mathfrak{A},i,j,g} = 1$ iff for all $\langle i', j' \rangle \in I \times J$such that$[\![$ Peter reads a book $]\!]^{\mathfrak{A},i',j',g} = 1$
  b. $[\![\Diamond$ Peter reads a book$]\!]^{\mathfrak{A},i,j,g} = 1$ iff there is a $\langle i', j' \rangle \in I \times J$such that$[\![$ Peter reads a book $]\!]^{\mathfrak{A},i',j',g} = 1$

The above meanings of the two sentences seem - at first glance - quite alright. Montague's account, however, cannot distinguish between different types of modality.
Modal words are ambiguous between many types of modality (*epistemic, deontic, circumstantial ...*). Therefore the fact that all the examples in (3) have the same interpretation presents a serious shortcoming.

(3) a. Jockl must have been the murderer. (epistemic - in view of the available evidence)
  b. Jockl must go to jail. (deotic - in view what the law provides)
  c. Jockl must sneeze. (circumstantial - in view of his current disposition)
  (Kratzer 1991:640)

Let us look at two different treatments of modality that take into account the different varieties of modality in natural language (Kratzer 1991). The first is directly adapted from modal logic. The second analysis was developed by Angelika Kratzer for natural language semantics and is usually called *Graded Modality.* Both accounts will be given in the modernized version of IL_M in a model structure $M = (D, W, \leq, F)$.
For the first account we need a few definitions (Kratzer 1991:641f).

Definition 3.8. *Truth of a proposition*: A proposition p is true in a world w iff $w \in p$.

Definition 3.9. *Logical consequence*: A proposition p follows from a set of propositions A iff p ist true in all worlds of W in which all propositions of A are true.

Definition 3.10. *Consistency*: A set of propositions A is consistent iff there is a world in W where all propositions of A are true.

Definition 3.11. *Logical Compatibility*: A proposition p is compatible with a set of propositions A iff $A \cup \{p\}$ is consistent.

Definition 3.12. *Conversational background*: A conversational background is a function $f : W \rightarrow \mathcal{P}(\mathcal{P}(W))$ that assigns to each world a subset of the powerset of worlds, i.e. a set of propositions, relative to which the modality is expressed.

The phrases "in view of..." next to the examples in (3) are conversational backgrounds for these sentences. A conversational background can either be given overtly in the sentence or can be supplied from the context. In both cases a new parameter f for conversational backgrounds has to be introduced.[14]

The main idea of the analysis is that the conversational background determines an accessibility relation on the set of worlds W. The modal operator can then only access the worlds that are connected via this relation to the salient world, i.e. the world with respect to which the meaning of the sentence is determined.

Definition 3.13.
$[\![\ \mathbf{must}\ \alpha]\!]^f = \{w \in W : w' \in [\![\alpha]\!]^f,\ \text{for all } w' \text{ such that } wR_fw'\}$
$[\![\ \mathbf{can}\ \alpha]\!]^f = \{w \in W : w' \in [\![\alpha]\!]^f,\ \text{for some } w' \text{ such that } wR_fw'\}$
where wR_fw' iff $w' \in \bigcap f(w)$[15]

In terms of compatibility and logical consequence this means that **must** α picks the set of worlds such that α follows from the conversational background and **can** α picks the set of worlds such that α is consistent with the conversational background. In the worlds picked by the modals the proposition α - the argument of **must/can** - is true.

The incorporation of accessibility relations into the analysis of modal expressions solves the problem of the ambiguity of modals. Nevertheless, the analysis is not optimal. Kratzer shows that it cannot deal with inconsistencies in the conversational background or with conditional clauses (cf. Kratzer (1991:642f)). Therefore, she proposes her system *Graded Modality*.

Modals in natural languages vary with respect to three different dimensions: *modal force*, *modal bases* and *ordering sources*. The modal force determines the grade of necessity or possibility of the modal. The modal base and the ordering source are given via two conversational background operators f and g. The first conversational

[14] Another possibility is to let the assignment function g pick out the right conversational background from the context.

[15] Note that only for simplicity the parameter f is the only parameter of the interpretation function that is given.

background f picks out the set of worlds that are accessible from the salient world analogously to the accessibility relation above. The second conversation background g gives a set of propositions relative to which the worlds in the modal base are ordered.

Definition 3.14. *Partial ordering $\leq_{g(w)}$ on the modal base $f(w)$:*
For all $w, w' \in f(w)$ and for $g(w) \subseteq \mathcal{P}(W)$: $w \leq_{g(w)} w'$ iff $\{p : p \in g(w) \text{ and } w' \in p\} \subseteq \{p : p \in g(w) \text{ and } w \in p\}$

The modal base together with the ordering source determines the set of accessible worlds. Kratzer (1991) gives two types of modal bases, *circumstantial* (in view of the relevant facts) and *epistemic* (in view of the available information). Ordering sources can be *deontic* (what the law provides), *bouletic* (what we want), *stereotypical* (what is normal), *doxastic* (what we believe), *teleological* (what our aims are) etc. Not every modal base can combine with every ordering source. Epistemic modal bases take ordering sources based on information, whereas circumstantial modal bases take ordering sources based on laws, aims, plans or wishes. With regard to terminology, modality with respect to a circumstantial modal base is also called *root modality.* Modality with respect to an epistemic modal base is of course called *epistemic modality.*

Kratzer distinguishes seven grades of modal force: *necessity, possibility, good possibility, at least as good a possibility, better possibility, weak necessity and slight possibility.* We will only give the new definition of necessity (**must**) and possibility (**can**), the two extremes of the scale. Both, necessity and possibility, can only access the optimal worlds with respect to the ordering source.

Definition 3.15. *Optimal worlds O(f,g,w)*:
$O(f, g, w) = \{z \in f(w) : \text{ for all } u \in f(w) \text{ such that } u \leq_{g(w)} z : z \leq_{g(w)} u\}$

The new meanings of **must** and **can** are thus:

Definition 3.16.
$[\![\textbf{ must } \alpha]\!]^{f,g} = \{w \in W : w' \in [\![\alpha]\!]^{f,g} \text{ for all } w' \in O(f, g, w)\}$
$[\![\textbf{ can } \alpha]\!]^{f,g} = \{w \in W : w' \in [\![\alpha]\!]^{f,g} \text{ for some } w' \in O(f, g, w)\}$
(Schwager 2008:3f)

As an example, we look at the truth condition of **Peter must kill someone** and **Paul can shoot a dog**.

The modal force of **must** in the first sentence is necessity. Say the situation in/about which the sentence is uttered is a life and death situation for Peter and Peter obviously wants to stay alive. The modal base is circumstantial and the ordering source bouletic. Therefore, f picks out all the worlds that are compatible with the circumstances in the actual world, i.e. in which Peter is in the same situation, and g picks out all the worlds in which Peter lives. Thus,

$$[\![\textbf{Peter must kill someone}]\!]^{f,g} = 1 \text{ iff } (\forall w' \in O(f,g,w))[\text{Peter kills someone at } w']$$

The optimal worlds are those in which Peter survives the life and death situation.

The modal force of **can** in the second sentence is possibility. For the second context take a situation in which there is a dog plague and the government changed the law such that it is okay to shoot a dog on the street. Then, if the second sentence is uttered in/about this situation, the base is also circumstantial but the ordering source is deontic. So, the truth condition for the second sentence is,

$$[\![\textbf{Paul can shoot a dog}]\!]^{f,g} = 1 \text{ iff } (\exists w' \in O(f,g,w))[\text{Paul shoots a dog at } w']$$

In this example the optimal worlds are those in which there is a dog plague and in which the dog shooting permission has been given.

Summing up section 3.1, we have given an a short introduction into the formal aspects of possible-worlds semantics as used in modern linguistics. In the next section, we will look at this system from a different point of view and propose a new model for possible-worlds semantics. We will give sample translations compatible with the new model for the three possible treatments of modality discussed.

3.2 Using Choice Sequences

In this section we now take a different point of view on possible worlds by modelling the logical space with the help of choice sequences.

In IL_M the cartesian product of the set of possible worlds and the set of moments of time, $I \times J$, forms the logical space. Thus a possible world in IL_M can be interpreted as a complete state of events at one instant of time in which every sentence is either true or false. Another possibility is that only the pair $\langle i, j \rangle$ completely describes a state of a certain possible world, i.e. that the moment of time j acts like an index on the possible world i picking out the state of events of world i at time j.

The following proposal incorporates both possible interpretations. Let IL'_M be a modernised version of IL_M as discussed in section 3.1.2.

Definition 3.17. A new model structure:
Let $\mathcal{M} = (W, U, F, g)$ be a model of IL'_M, where W is an enumeration of possible worlds, U a finite set of entities forming the "universe", F an enumeration of interpretation functions of IL'_M, and g a global variable assignment. The functions F_i, the values of

F, induce together with U an enumeration M of extensional models extensional model $\mu_i = (U, F_i)$.

The main idea is to model logical space as a finite set of sequences of extensional models at each point in time. For the time being these sequences are taken to be lawless. In other words, possible worlds are sequences of extensional models that grow with time.

In the following sections the make up of possible worlds, the requirements consequently imposed on worlds by the axioms of choice sequences and the modified treatment of modality (for worlds as choice sequences) are discussed.

3.2.1 A New World

Possible worlds are modelled by some type of lawless choice sequences. Let us discuss the problems and merits of the four different types of choice sequences treated in the previous chapters.

The lawless sequences of chapter 1 are ill-suited for possible worlds since they can not describe the flow of time into past and future.

Bilateral lawless sequences can model the entire timeline, but have the decided disadvantage that operations are only defined for their monolateral parts. Even though relaitivised roots can be worked with, the fact that for each operation on the sequence either the entire past or the entire future values relative to this root have to be forgotten, make them ill-suited as well.

Bundled lawless sequences seem better suited for worlds than bilateral sequences. Bidimensional bundled sequences can model one direction of time with each dimension. However, the fact that values of bundled sequences are tuples links the two directions of time in the sense that for each new model added in the "future-dimension" there is a new model added in the "past-dimension". Thus the two directions of time develop symmetrically, which is undesirable for the simple reason that we actually seem to know more about the past than about the future.

This leaves indexed lawless sequences. Indexed lawless sequences can model the entire timeline if the index set has cardinality two. The two directions can develop asymmetrically since they are not linked by any relation and operations on indexed sequences are well-defined. Thus, we choose indexed sequences to represent possible worlds.

Definition 3.18. The parameter W of the new proposed model is an enumeration of indexed sequences w. $W : \mathbb{N} \to (\mathbb{N} \times I)^{\mathbb{N}}$ with $|I| = 2$.

The reason we can take natural numbers as the first dimension of the cartesian product is that the enumeration of assignment functions F in the model quite naturally induces an enumeration on all extensional models. Thus each extensional model can be identified with a $n \in \mathbb{N}$. For the index set of cardinality two, we will take $\{0, 1\}$ just for convenience. According to the translation from indexed sequences into bilateral sequences, 0 will be interpreted as "adding a value to the left", i.e. temporally in the past, and 1 will be interpreted as "adding a value to the right", i.e. temporally in the future. Thus, the index set could also be written as $\{P, F\}$.

A possible world in its new definition therefore looks as in the next example.

> $\langle 7, 1 \rangle , \langle 5, 0 \rangle , \langle 9, 0 \rangle , \langle 10, 0 \rangle , \langle 11, 1 \rangle , \langle 13, 0 \rangle , \langle 4, 0 \rangle , \ldots$
>
> where the information coded is:
>
> $\langle \mu_7, F \rangle , \langle \mu_5, P \rangle , \langle \mu_9, P \rangle , \langle \mu_{10}, P \rangle , \langle \mu_{11}, F \rangle , \langle \mu_{13}, P \rangle , \langle \mu_4, P \rangle , \ldots$
>
> which is in the coded temporal order:
>
> $\ldots , \mu_4, \mu_{13}, \mu_{10}, \mu_9, \mu_5, \mu_7, \mu_{11}, \ldots$

Each $n \in \mathbb{N}$ completely determines an extensional model wn, a value of the possible world w. This address n however does not reflect the real position of the model along the timeline as can be seen in the above example. We will set w_0 as the actual world and s_*0 as the default speech time. This creates no problems linguistically since a sentence is always interpreted relative to one time parameter. Changing the speech time can either be solved internally in setting a new time parameter at the interpretation function or by taking existing values as deliberately placed initial segments for a new enumeration of worlds.

In general, the new definition of a model for IL'_M has to be interpreted dynamically, rather than static as the model for Montague semantics is usually intended. Since both, the set of worlds and the set of extensional models are choice sequences themselves and therefore potentially infinite in the sense that at each moment only finite sequences and thus finite sets of worlds and models are considered which, however, grow in time. This is a necessary consequence of building a model with the help of choice sequences. Even though it is usually assumed that the set of possible worlds is infinite, this fact is, as far as we know, never used explicitly in proposed meanings for lexical elements. Therefore, working with an at each point finite amount of worlds has no impact on the interpretation of sentences.

The assumption that worlds behave like lawless sequences leads to certain predictions when the axioms for choice sequences are considered. The first axiom, $LS1$, adapted to possible worlds, says that every finite sequence of extensional models is a possible initial segment of a world. This is a very peculiar state of affairs when the metaphysics of possible worlds is considered. However, as we have no opinion concerning the metaphysic status of possible worlds, we ignore this peculiarity. The decidability of equality, $LS2$, says that either two worlds that have the same sequence of models are actually the same object or, if they have identical initial segments, they eventually differ in one value. This is also desirable linguistically since we want to be potentially able to identify the actual world.

Operations on possible worlds we have encountered in section 3.1.3 with the linguistic theories of modality. Accessibility relations and conversational backgrounds map possible worlds into sets of possible worlds. The axioms $LS3$ and $LS4$ require from operations on the new worlds that worlds with the same initial segment get assigned the same value, in other words that these operations are continuous. This is desirable since worlds with the same initial segment as the actual world, in which a statement is made, are indistinguishable from the actual world. Therefore, it is expected that mappings that only depend on the state of events of a world should return the same values. The two axioms also state that operations on possible worlds can only use an initial segment of a world as argument. Thus, only given information represented by the initial finite sequence of extensional models can be used to determine the value of an operation.

3.2.2 Checking Lexical Items

In the previous section, the new possible worlds modelled by choice sequences were introduced. A change in the model structure also leads to change in some interpretations of types. The old possible worlds were world-time pairs, where worlds and times were theoretically independent. The new possible worlds are bigger objects - sequences of old world-time pairs. Thus, all linguistic objects depending on possible worlds, need to be looked at anew from the perspective of the new model.

We will start by looking at propositions. Propositions in the old model were sets of world-time pairs, i.e. objects of type $\langle s, t \rangle$. To capture both the modal and the temporal level of natural language, we will add a new intensional type i. The intensional type s for world-time pairs will be used solely for entire worlds. Thus the new definition of types is as follows.

Definition 3.19. The *revised type system* of IL'_M:
Basic types are e for entity expressions (individuals) and t for truth value expressions. The set *Type* is the smallest set Y such that (1) $e, t \in Y$, (2) if $a, b \in Y$ then $\langle a, b \rangle \in Y$ and (3) if $a \in Y$ then $\langle s, \langle i, a \rangle \rangle \in Y$.

Note, that i is *not* the type for instances of time as in section 3.1.2 but the type for addresses of values of a world. The order of information gain and the timeline do not coincide, which causes this slight interpetative flaw.

By application of the new type system, the type of propositions becomes $\langle s, \langle i, t\rangle\rangle$. This means that propositions are functions

$$p : (\mathbb{N} \times I)^{\mathbb{N}} \rightarrow \{0, 1\}^{\mathbb{N}}$$

from worlds into functions of addresses into truth values.

The same changes are made for all other intensional objects: individual concepts have the type $\langle s, \langle i, e\rangle\rangle$ properties have e.g. the type $\langle s, \langle i, \langle e, t\rangle\rangle\rangle$.

The interpretation function has to be modified minimally as well. An address value has to be added to the world, model and assignment function parameters.

Just for technical reasons we also need the projection π_i^j defined in section 2.4.1 since the world values are pairs of models and indices. The interpretation needs to have access to each of the components separately.

The derivation of **Peter sees Susi** is repeated for the new semantics. The structure of the sentence is

$$\mathbf{see}(\mathbf{Susi})(\mathbf{Peter})$$

The meanings of the parts is then as follows.

$$[\![\mathbf{Peter}]\!]^{w_0,s_*,M,g} = \lambda w.\lambda i.\ \text{Peter}$$
$$[\![\mathbf{Susi}]\!]^{w_0,s_*,M,g} = \lambda w.\lambda i.\ \text{Susi}$$
$$[\![\mathbf{see}]\!]^{w_0,s_*,M,g} = \lambda w.\lambda i.\lambda y.\lambda x.\ x \text{ sees } y \text{ in } \pi_1^2(wi)$$

Then the truth condition of the entire sentence is:

$$[\![\mathbf{Peter\ see\ Susi}]\!]^{w_0,s_*,M,g} \Leftrightarrow$$
$$[\![\mathbf{see(Susi)(Peter)}]\!]^{w_0,s_*,M,g} \Leftrightarrow$$
$$[\lambda w.\lambda i.\lambda y.\lambda x.\ x \text{ sees } y \text{ in } \pi_1^2(wi)](\lambda w.\lambda i.\ \text{Susi })(\lambda w.\lambda i.\ \text{Peter}) \Leftrightarrow$$
$$\lambda w.\lambda i.\ [\lambda w.\lambda i.\ \text{Peter }](x) \text{ sees } [\lambda w.\lambda i.\ \text{Susi }](y) \text{ in } \pi_1^2(wi) \Leftrightarrow$$
$$\lambda w.\lambda i.\ \ \text{Peter sees Susi in } \pi_1^2(wi)$$

For $\pi_1^2(wi)$ the parameters $\pi_1^2(w_0(s_*))$ are inserted. Thus:

$$[\![\mathbf{Peter\ see\ Susi}]\!]^{w_0,s_*,M,g} = 1 \text{ iff Peter sees Susi in } \pi_1^2(w_0(s_*))$$

For simple, purely extensional sentences, like **Peter sees Susi**, above no real changes can be felt. In the next section we will discuss modality and redefine the meanings given for the modal operators of the three approaches to modality given in section 3.1.3.

3.2.3 Modality Revisited

As we have seen in the previous section, modelling the logical space with choice sequences has no direct consequence for the interpretation of purely extensional sentences. As soon as intensionality has to be dealt with, changes have to be expected.

The first treatment of modality we will take a closer look at, is the one by Montague (1973). He proposes the meanings for **must** and **can** in (4).

(4) a. $[\![\Box\phi]\!]^{\mathfrak{A},i,j,g} = 1$ iff $[\![\phi]\!]^{\mathfrak{A},i',j',g} = 1$ for all $i' \in I$ and $j' \in J$

b. $[\![\Diamond\phi]\!]^{\mathfrak{A},i,j,g} = 1$ iff $[\![\phi]\!]^{\mathfrak{A},i',j',g} = 1$ for some $i' \in I$ and $j' \in J$

Montague's accessibility relation for $\Box$ and $\Diamond$ is the universal relation on world-time pairs in $I \times J$. This can be remodelled quite easily in terms of worlds as choice sequences. Since we can not quantify over the infinite amount of all world values of all worlds, a different approach has to be found. Let us define the operator $.^*$, which forms a choice sequence from the i-th value of all worlds in the enumeration W.

Definition 3.20. The operator $.^* : \mathbb{N} \rightarrow (\mathbb{N} \times I)^{\mathbb{N}}$ takes address values i and returns lawlike indexed sequences of the form $\lambda n.w_n i$, where w_n is the nth world in the enumeration of worlds W.

The idea is to run through all address slices and for each address slice to look at all initial segments, which amounts to looking at all existing world values of all worlds in W. Hence, we propose the new meanings for **must** and **can** as in (5).

(5) a. $[\![\textbf{must}]\!]^{w_0,s_*,M,g} = \lambda p.\forall i \forall n(i^* \in n \rightarrow p(i^*)(\text{lth}n - 1))$

b. $[\![\textbf{can}]\!]^{w_0,s_*,M,g} = \lambda p.\exists i \exists n(i^* \in n \wedge p(i^*)(\text{lth}n - 1))$

Note that the address slices of course are not time slices. The world values in the functions i^* do not have to be simultaneous with respect to model internal time. Since **must** as well as **can** are in all three treatments of modality proposed to express time-independent necessity and possibility.

By virtue of the substitution of a world value dependent accessibility relation for the universal relation, the second proposal for modals in natural language gives a more differentiated and fine grained picture.

The meanings of **must** and **can** of the second proposal are repeated in (6).

(6) a. $[\![$ **must** $\alpha]\!]^f = \{w \in W : w' \in [\![\alpha]\!]^f,$ for all w' such that $wR_f w'\}$

b. $[\![$ **can** $\alpha]\!]^f = \{w \in W : w' \in [\![\alpha]\!]^f,$ for some w' such that $wR_f w'\}$
where $wR_f w'$ iff $w' \in \bigcap f(w)$.

The parameter f gives a set of propositions relative to the utterance world that determines the accessible worlds. We will identify f with the set of propositions f gives for the current world of evaluation. A world w is accessible if and only if all propositions of f are true in w. Hence, f acts as a filter on W.

This restriction by the propositions in f is easily added to the proposed modification of Montague's modality treatment in (5) to give the meanings of the two modals for the new model.

Definition 3.21. Let $f = \{f_1, \dots, f_n\}$ be a set of propositions then the proposition $\tilde{f} = \lambda w.\lambda i.f_1(w)(i) \wedge \dots \wedge f_n(w)(i)$ is the conjunction of all $f_i \in f$.

Thus the new proposal for worlds as choice sequences is given in (7). The parameter list of the interpretation function is again reduced to $\tilde{f}$ for readability.

(7) a. $[\![\textbf{must}]\!]^{\tilde{f}} = \lambda p.\forall i \forall n(i^* \in n \wedge \tilde{f}(i^*)(\text{lth}n - 1) \rightarrow p(i^*)(\text{lth}n - 1))$
b. $[\![\textbf{can}]\!]^{\tilde{f}} = \lambda p.\exists i \exists n(i^* \in n \wedge \tilde{f}(i^*)(\text{lth}n - 1) \wedge p(i^*)(\text{lth}n - 1))$

Finally we look at Kratzer's treatment of **must** and **can** in her *Graded Modality* analysis. Kratzer (1991) adds a second parameter g (not to be confused with the variable assigment which assigns to the salient world another set of propositions. The parameter f is used again as a filter on the entire set of worlds and determines the *modal base.* The set of propositions g, the *ordering source*, then induces an order on the worlds picked by f.[16] Via the order induced by g a set of optimal worlds $O(f, g, w)$ is defined that determines the final set of accessible worlds. Thus the meanings of **must** and **can** in Kratzer's analysis is as repeated in (8).

(8) a. $[\![$ **must** $\alpha]\!]^{f,g} = \{w \in W : w' \in [\![\alpha]\!]^{f,g}$ for all $w' \in O(f, g, w)\}$
b. $[\![$ **can** $\alpha]\!]^{f,g} = \{w \in W : w' \in [\![\alpha]\!]^{f,g}$ for some $w' \in O(f, g, w)\}$

We will not remodel the ordering to induce the set of optimal worlds, since the definition would require quantification over an infinite set. Instead we will define a set of maximal consistent subsets of g. The ordering source is, amongst other reasons, used to account for situations in which the elements of the conversational background f of the second treatment (example (6)) are inconsistent. The optimal worlds with respect to the ordering source are those that satisfy a maximal simultaneously satisfiable subset of g, i.e. a maximal consistent subset.

[16] g is also a conversational background relative to the world of evaluation. Like for f we will identify g with its value.

Since g is a finite set of propositions, we can define a new parameter $\tilde{g}$ which is a disjunction of propositions $\tilde{g}_i^c$ which are conjunctions over all propositions contained in a maximal consistent subset of g.

Definition 3.22. For a finite set of propositions $g = \{g_1, \ldots, g_m\}$, the parameter $\tilde{g} = \bigvee_{g_i^c \in \Gamma} \tilde{g}_i^c$ is a disjunction of propositions $\tilde{g}_i^c$, which is a conjunction of all propositions in g_i^c. A set g_i^c is an element of Γ, the set of maximal consistent subsets of g.

$$\Gamma = \bigcup_{\substack{g' \subseteq g \\ g' consistent \\ g' maximal}} g'$$

To give an example, let $g = g_1, g_2, g_3, g_4$ be a set of propositions and let $\{g_1, g_2\}$ be inconsistent. Then:

$$\tilde{g} = \bigvee_{g_i^c \in \Gamma} \tilde{g}_i^c \text{ where } \Gamma = \{\{g_1, g_3, g_4\}, \{g_2, g_3, g_4\}\}$$

Thus:

$$\begin{aligned} \tilde{g} &= \tilde{g}_1^c \vee \tilde{g}_2^c \\ &= \lambda w.\lambda i.(g_1(w)(i) \wedge g_3(w)(i) \wedge g_4(w)(i)) \vee (g_2(w)(i) \wedge g_3(w)(i) \wedge g_4(w)(i)) \end{aligned}$$

The definitions of the meanings of **must** and **can** in the spirit of *graded modality* can hence be formulated as in (9).

(9) a. $[\![\textbf{must}]\!]^{\tilde{f},\tilde{g}} = \lambda p.\forall i \forall n(i^* \in n \wedge \tilde{f}(i^*)(\text{lth}n-1) \wedge \tilde{g}(i^*)(\text{lth}n) \rightarrow p(i^*)(\text{lth}n-1))$

b. $[\![\textbf{can}]\!]^{\tilde{f},\tilde{g}} = \lambda p.\exists i \exists n(i^* \in n \wedge \tilde{f}(i^*)(\text{lth}n - 1) \wedge \tilde{g}(i^*)(\text{lth}n) \wedge p(i^*)(\text{lth}n - 1))$

To fully illustrate the workings of the reformulation of the two meaning postulates, we will give a model for the sentence **Peter must read a book**. The *modal base* f is circumstantial, e.g. "the professor wants his students to read any book to get a good grade", and the **ordering source** g is bouletic, e.g. Peter's wishes to have free time and to get a good grade.

The model $M = (W, U, F, g)$ is made up of the following parts:

$M = \langle w_0, w_1, w_2 \rangle$

$U = \{$Peter, Sense and Sensibility, Vanity Fair$\ldots\}$

$F = \langle F_0, F_1, F_2 \rangle$

$g = \emptyset$

The three worlds have the form

$$w_0 = \langle 0,0\rangle, \langle 1,1\rangle, \langle 2,0\rangle$$
$$w_1 = \langle 1,0\rangle, \langle 1,1\rangle, \langle 0,1\rangle$$
$$w_2 = \langle 2,1\rangle, \langle 2,1\rangle, \langle 1,1\rangle$$

the conversational backgrounds f and g relative to $w_0 0$ are

$$f = \{f_0\} \text{ and } g = \{g_0, g_1, g_2\}$$

with
$f_0 = \lambda w.\lambda i.$ the professor wants $\ldots \pi_1^2(w(i))$,
$g_0 = \lambda w.\lambda i.$ Peter has free time in $\pi_1^2(w(i))$ and
$g_1 = \lambda w.\lambda i.$ Peter gets a good grade in $\pi_1^2(w(i))$.
The three extensional models $\mu_i = (U, F_i),\ i = 0,1,2$ are

μ_0 :
$[\![\textbf{Peter}]\!] = \{\text{Peter}\}$
$[\![\textbf{book}]\!] = \{\text{Sense and Sensibility, Vanity Fair}\}$
$[\![\textbf{read}]\!] = \{\langle \text{Sense and Sensibility}, \langle \text{Peter}, 0\rangle\rangle, \langle \text{Vanity Fair}, \langle \text{Peter}, 1\rangle\rangle \ldots\}$
$[\![\mathrm{f}_0]\!] = 1, [\![\mathrm{g}_0]\!] = 0, [\![\mathrm{g}_1]\!] = 1$

μ_1 :
$[\![\textbf{Peter}]\!] = \{\text{Peter}\}$
$[\![\textbf{book}]\!] = \{\text{Sense and Sensibility}\}$
$[\![\textbf{read}]\!] = \{\langle \text{Sense and Sensibility}, \langle \text{Peter}, 0\rangle\rangle\}$
$[\![\mathrm{f}_0]\!] = 0, [\![\mathrm{g}_0]\!] = 1, [\![\mathrm{g}_1]\!] = 0$

μ_2 :
$[\![\textbf{Peter}]\!] = \{\text{Peter}\}$
$[\![\textbf{book}]\!] = \{\text{Sense and Sensibility, Vanity Fair}\}$
$[\![\textbf{read}]\!] = \{\langle \text{Sense and Sensibility}, \langle \text{Peter}, 1\rangle\rangle, \langle \text{Vanity Fair}, \langle \text{Peter}, 1\rangle\rangle\}$
$[\![\mathrm{f}_0]\!] = 1, [\![\mathrm{g}_0]\!] = 1, [\![\mathrm{g}_1]\!] = 0$

The postulated meaning for **Peter must read a book** is

$$
\begin{aligned}
&[\![\textbf{Peter must read a book}]\!]^{\tilde{f},\tilde{g}} = 1 \text{ iff}\\
&\forall i.\forall n.(i^* \in n \wedge \tilde{f}(i^*)(\mathrm{lth}n-1) \wedge \tilde{g}(i^*)(\mathrm{lth}n-1) \rightarrow\\
&\ \text{Peter reads a book in } \pi_1^2(i^*(\mathrm{lth}n-1)))
\end{aligned}
$$

The address slices i^* for $i = 0, 1, 2$ are

$$
\begin{aligned}
0^* &= \langle 0, 0\rangle, \langle 1, 0\rangle, \langle 2, 1\rangle\\
1^* &= \langle 1, 1\rangle, \langle 1, 1\rangle, \langle 2, 1\rangle\\
2^* &= \langle 2, 0\rangle, \langle 0, 1\rangle, \langle 1, 1\rangle
\end{aligned}
$$

and $\tilde{f} = f_0$ and $\tilde{g} = g_0 \vee g_1$.
Since the last entry of each finite address slice is looked at, in sum all values of the address slices are considered. So which are the world-values that satisfy $\tilde{f}$ and $\tilde{g}$? Only the extensional models μ_0 and μ_2 satisfy $\tilde{f}$. Thus all values with 1 as the first component are filtered by f from the modal base. Checking g_0 and g_1, we see that μ_0 and μ_2 each satisfy one of the two propositions. Thus they satisfy $\tilde{g}$ and

$$\forall i.\forall n.(i^* \in n \wedge \tilde{f}(i^*)(\mathrm{lth}n-1) \wedge \tilde{g}(i^*)(\mathrm{lth}n-1) \rightarrow \text{ Peter reads a book in } \pi_1^2(i^*(\mathrm{lth}n-1)))$$

is true in this model and thus **Peter must read a book** is true.
Peter must read a book becomes false if we change $[\![\mathrm{f}_0]\!] = 0$ to $[\![\mathrm{f}_0]\!] = 1$ in μ_1.

3.3 Summary

In the first part of this chapter Montague's system IL_M was introduced and some more modern modifications of this system were given. Furthermore, we looked at three possible treatments of modality in natural language semantics. Afterwards, in the second part, we reformulated the models of IL'_M, a modernised version of IL_M, in terms of choice sequences. We adapted and discussed the type system of lexical items and then translated the three treatments of modality such that the meaning postulates of the two modals **must** and **can** were compatible with the reformulated model.

Chapter 4

Further Issues

The discussion in the previous three chapters, especially in chapters 2 and 3, brought up four questions for further research that I want to illustrate in this chapter.

Issue 1 concerns other possible types of choice sequences that differ from the original choice sequences in terms of for which element of the domain (usually $\mathbb{N}$) the next value is added to the sequence. For normal choice sequences as well as indexed choice sequences and bundled choice sequences the next value after αn is always added at $\alpha n + 1$. This guarantees that these choice sequences are total mappings on $\mathbb{N}$.

Bilateral choice sequences are more liberal in terms of where the next value is added to the already existing values. It is easy to see that this strategy of adding values means that bilateral choice sequences are not necessarily total on $\mathbb{Z}$. Just take the sequence where $\alpha 0$ is the relativised root and all values are only added to the right or left.

As we have already seen in chapter 2, it is difficult to define the initial segment relation for bilateral sequences and hence operations directly applied to the sequence can not be defined without considerably restricting the strategy of adding values again.

The first question to be further examined is therefore how choice strategies which are not as systematic and rigid as the one for choice sequences, indexed sequences and bundled sequences influence totality of the choice process and moreover the definition of operations on these sequences and their characteristics. The main point to consider here is that if a choice process is not total, some positions of the domain remain "untouched". Thus operations depending on these positions getting assigned a value will remain undefined for such processes.

The second issue regards the application of choice sequences to Montague semantics in chapter 3. We substituted the model formed with the help of choice sequences for the standard Kripke possible-worlds semantics.

Hence it is of interest quite generally to ask about the correlations between choice sequences and Kripke semantics. A finite Kripke semantics is obviously a subpart of a model built from choice sequences, i.e. the state of the model at a certain time. A comparison of Kripke semantics with an infinite set of points or worlds and choice sequences as models seems to us to be the more interesting question.

Particularly end points of Kripke semantics, i.e. points that are accessible for some points but from which no other point is accessible, are interesting in this regard since for such a point $\Box p$ is valid and $\Diamond p$ is unsatisfiable for all formulas p.

The third area of further interest is the investigation of the behaviour of the model built from choice sequences and the sentences that are true in this model before and after a step of information gain. Since the set of worlds and the set of extensional models are also enumerations growing with time, the addition of further worlds and extensional models is bound to have an impact on the validity of intensional propositions.

The main question is thus how the different states of the model are connected to each other (if they indeed are comparable) and the predictions made by such a dynamic model.

The last and fourth issue, on an entirely different note, regards the modal structure of Kratzer's *Graded Modality* (Kratzer 1991). The assumption is that *Graded Modality* satisfies at least the axioms of system **K** but not all axioms of $\mathbf{S}_5$.[1]

Definition 4.1. The system **K** is the classical propositional calculus for the language of modal logic extended by

K-axiom schema: $\Box(\phi \to \psi) \to (\Box\phi \to \Box\psi)$
rule of modal generalisation: if $\vdash \phi$ then $\vdash \Box\phi$

Definition 4.2. The system $\mathbf{S}_5$ is the system **K** extended by

T-axiom schema: $\Box\phi \to \phi$
and either
5:$\Diamond\phi \to \Box\Diamond\phi$
or
4:$\Box\Box\phi \to \Box\phi$
B:$\phi \to \Box\Diamond\phi$

$\mathbf{S}_5$ is in particular characterized by models where the accessibility relation of the modal operators is an equivalence relation, i.e. reflexive, symmetric and transitive. **K** poses no restrictions on its models. Between **K** and $\mathbf{S}_5$ other systems lie with more restrictions on models than **K** and less restrictions than $\mathbf{S}_5$ (Fitting 2007). Thus, the accessibility

[1]M. Baaz p.c.

relation induced by f and g in *Graded Modality* has to be studied in detail to show how the model of *Graded Modality* can be characterized.

We think that all of the above issues are worthwhile questions to investigate. Unfortunately, we have no answers to these questions.

Bibliography

Brouwer, L.E.J. 1992. *Intuitionismus*. BI-Wissenschaftsverlag.

Carnap, Rudolf. 1947. *Meaning and Necessity*. University of Chicago Press.

Chomsky, Noam. 1957. *Syntactic structures*. The Hague: Mouton.

Church, Alonzo. 1951. A formulation of the logic of sense and denotation. In *Structure, Method and Meaning*, ed. P. Henle. The Liberal Arts Press.

Fitting, Melvin. 2007. Modal Proof Theory. In *Handbook of Modal Logic*, ed. Frank Wolter Patrick Blackburn, Johan van Benthem, chapter 2, 86–138. Elsevier.

Frege, Gottlob. 1892. Ueber Sinn und Bedeutung. *Zeitschrift fuer Philosophie und philosophische Kritik* 100:20–50.

Heim, Irene, and Angelika Kratzer. 1998. *Semantics in Generative Grammar*. Oxford: Blackwell.

Heyting, Arend. 1956. *Intuitionism, an introduction*. North-Holland.

Higginbotham, James. 1985. On Semantics. *Linguistic Inquiry* 16:547–593.

Kratzer, Angelika. 1991. Modality. In *Handbuch Semantik*, ed. Arnim von Stechow and Dieter Wunderlich, chapter 23, 639–650. De Gruyter.

Kreisel, Georg. 1968. Lawless sequences of natural numbers. *Compositio mathematica* 20:222–248.

Montague, Richard. 1973. The Proper Treatment of Quantification in Ordinary English. In *Approaches to Natural Language*, ed. P. Suppes K.J.J Hintikka, J.M. Moravcsik, 221–242. Dordrecht.

Moschovakis, Joan Rand. 1986. Relative lawlessness in intuitionistic analysis. *Journal of Symbolic Logic* 52:68–88.

Partee, Barbara. 1973. Some transformational extensions of Montague grammar. *Journal of Philosophical Logic* 2:51–76.

Partee, Barbara. 2005. Montague, Richard (1930-71). In *The encyclopedia of language and linguistics*, ed. Keith Brown. Elsevier, 2nd edition edition.

Schwager, Magdalena. 2008. The Modal Operator Analysis for Imperatives (Handout ESSLLI 2008).

Troelstra, A.S. 1977. *Choice Sequences - a chapter of intuitionistic mathematics*. Oxford University Press.

Troelstra, A.S. 1983. Analysing Choice Sequences. *Journal of Pilosophical Logic* 12:197–260.

Troelstra, A.S., and D. van Dalen. 1988a. *Constructivism in Mathematics*, volume I of *Studies in Logic*. Elsevier.

Troelstra, A.S., and D. van Dalen. 1988b. *Constructivism in Mathematics*, volume II of *Studies in Logic*. Elsevier.