LOGIC, TOPOLOGY, AND INTEGRATION

Karl Schlechta
Laboratoire d'Informatique de Marseille, URA CNRS 1787
Université de Provence, Case 32
3, Pl. Victor Hugo
F-13331 Marseille Cedex 3, France
ks@gyptis.univ-mrs.fr

14 July 1994

Abstract

The central notion will be that of closeness of (or difference between) two theories. In the first part, we give intuitive arguments in favour of considering topologies on the set of theories, continuous logics, and the average difference between two logics, i.e. the integral of their difference. We continue by arguing for the importance of the difference between theories in a wide range of applications and problems. In the second part, we give some basic definitions and results for one such type of topology. In particular, separation properties and compactness will be discussed, and examples given. The techniques employed for constructing the topology will also be used for defining a $\sigma - algebra$ of measurable sets on the set of theories, leading to the usual definition of the Lebesgue integral, and a precise definition of the average difference of two logics.

1 OVERVIEW, MOTIVATION, AND BASIC DEFINITIONS

1.1 Overview

In the first part, we argue for the usefulness and naturalness of introducing topological notions into a logical, and Artificial Intelligence, context. In the second part, we construct and investigate one such topology. We certainly do not claim that the techniques developed there are sufficient to satisfy all of the needs described in the first part, we hope to have elaborated a sound starting point.

In particular, in the first part, we

1. motivate the construction of the second part which uses a (partial) order on the propositional variables
2. show the relevance of the notion of distance, and continuity, for
   (a) an abstract characterization of a logic as "well-behaved" in the sense that it "does not jump"
   (b) the design and testing of expert systems
   (c) theory revision
   (d) approximative reasoning
   (e) automatic generation of abstracts of a text
   (f) the concept of graceful degradation
3. argue for the importance of the average difference of two logics.

In the second part, we develop and investigate one topology on the set of theories of a given (propositional) language $\mathcal{L}$ in detail. (An extension to the first order case is indicated following Corollary 2.4.) We first work semantically and then carry the construction over to the set of theories of the language $\mathcal{L}$. Our starting point is to (partially) order the propositional variables of $\mathcal{L}$ "by importance", and define the distance between two models, or sets of models, through the set of propositional variables on which they differ: The more important the propositional variables on which they disagree, the more they are distant from each other. Following customs in the field of topology, we examine the basic properties of the space so constructed. In particular, separation properties and compactness (as a topological space) are investigated, and it is shown that the topology has enough properties of a metric space to permit, in a straightforward generalization, the definition of a uniformly continuous function on it. A function on the set of theories of a language is a logic, so we have defined continuous and uniformly continuous logics (for $\mathcal{L}$ and that topology). The connections between continuity and compactness (as a logic) are investigated.

We then use the uniformity of the construction to define in a natural way the Lebesgue integral of the difference of two logics for $\mathcal{L}$, making the average difference of two logics precise. We show in conclusion that the average difference, i.e. the integral of the difference, of two continuous logics is defined.

1.2 Motivation to consider continuous logics and the intuition behind our definition of the topology

Initial Motivation: It is often seen as a problem that changing the input of a logic a little may change the output drastically: Two theories T and T' may differ only a little, but the sets of their consequences under a certain logic $\mid\sim$, $\overline{\overline{T}}$ and $\overline{\overline{T'}}$ may differ much. The corresponding notion in analysis is that of a discontinuous function and related concepts. To speak about continuity of logics, one first needs a topology on its domain and range, i.e. the set of theories (seeing a logic as a function from the set of theories to the set of theories). We think here e.g. of medical diagnosis, where we first determine that the patient has liver trouble, and then that it is an infection, and then the type of infection: We have a refinement process. Changing the diagnosis totally, e.g. to heart attack, is something different, and we want to make this precise: the difference between refinement and drastic revision. In our intuitive start, we follow very closely e.g. physical theories, which are most successful when the functions are well-behaved (changing the force a little will change the speed a little, and once we know the order of magnitude of force, we know the order of magnitude of speed). Physical theories and their applications are far less successful in more chaotic systems. So we intend to introduce a corresponding notion for logics, i.e. a new notion of well-behavedness of logics (cf. e.g. cumulativity or rational monotony - for definition and discussion see e.g. [KLM90] and [LM92] - which are purely logical, and need no additional structure, like a topology). A further elaboration of the concept of continuous logics gives us uniform continuity, and in our technical development, we will give a technique to define this concept for logics, even if the space is not metric - a slight generalization will do.

Construction of the topology by an order on the propositional variables:

We think here of the propositional variables as of digits in a physical measurement, they are ordered from the important to the negligeable. In our formal development, two (classical) models will be close if they agree on the important variables, differing maybe on the less relevant ones. They will be far apart, if they disagree on the important variables - irrespective of their behaviour on the less relevant ones. Our notion of distance between models therefore does not "count" the differences as e.g. in [Dal88], but weighs them according to the order. The distance between two theories is then defined via the distance of their models. Two theories T and T' will be close, iff for every model m of T, there is a model m' of T' such that m and m' are close, i.e. agree on the important propositional variables, and, vice versa, i.e. iff for every model m' of T' there is a close model m of T.

Let me motivate this approach by the following two examples:

(a) Suppose two witnesses describe a nighttime traffic accident. Let p_0 stand for "first car is a truck", p'_0 stand for "second car is a passenger car", p_1 stand for "second car is black", p'_1 stand for "second car is dark blue". Given poor visibility, will will not accord so much importance to p_1 and p'_1, i.e. if the two witnesses disagree on p_1/p'_1, we will not necessarily doubt the validity of their testimony. If, however, one says that the first car was a truck, and the other, that it was not, we will seriously question their reliability. Thus, p_0, p'_0 are more important, reliable than p_1, p'_1.

(b) In plant or animal taxonomy, it is often possible to achieve a rough classification by some trait. So, e.g., having 5 petals will allow the classification into the rose family, and the shape of the fruit will tell whether it is a pear or an apple tree. (Hopefully this is all correct.) Thus, the number of petals is more important than the shape of the fruit - botanically more robust so to say.

It is obvious that this gives us - for continuous logics wrt. that order

- a notion of *relevance*, and also of locality: We may not need to decide about p_m, when p_0 already determines the logical order of magnitude, and this order of magnitude is sufficient for our purposes.

In analogy, a physical theory whose results depend on the 10th digit of measurement but neglects the first, is certainly funny, and bad for predictions, or, maybe, the measuring device is just the wrong one. For measuring physical entities, we have amperemeters etc., so what are the "logimeters"? Of course, our categories and the importance we attach to them. Likewise, a "good" logic should first consider the important categories, and refine its conclusions by considering the less relevant information, and not make "jumps" when the less relevant information is changed a little.

Recall that in non-monotonic logics, we may have conflicting information of different degrees of reliability. In the case of "Birds fly", "Penguins do not fly", "Tweety is a penguin", we conclude that Tweety does not fly, as the more specific information (penguins are all birds but not vice versa) is considered to be more reliable.

Our order of importance is not this order of reliability of non-monotonic logics: the 10th digit says *nothing* about the order of magnitude, but weak information says *something*, though this may be overridden by stronger information. It has something to do with qualitative physics, where categories describe ranges, and (the hope is that) they are fine enough for the purpose at hand. But these categories are rigid, and we treat also overlapping sets of various sizes - we are more flexible. As far as I see, our order of importance has nothing to do with variable precision logic ([MW86]).

Where does the additional structure - the order on the variables - come from? We have to put it in, as the epistemic entrenchment relations for theory revision àla Gärdenfors et al. Our independence results (Corollary 2.4/4) - i.e. part 4. of Corollary 4 in Section 2 - and Lemma 2.8/2), will however show that, to a certain extent, the exact order does not matter: In many cases the results are robust when a cofinality property is respected (see Definition 2.1).

1.3 Further arguments for the utility of the notion of (uniform) continuity

(a) It supports conceptual clarification: Designing a logic - or an expert system - to be continuous helps to clarify concepts. It forces to distinguish the important from the less important, and, thus may even help to introduce new and meaningful concepts.

It gives us therefore a strategy to develop a system along clear lines of significance: The important information will be - in a way - treated separately from the only marginal one. It helps to find a modular architecture for such a system, working with several layers of importance.

(b) A continuous system is seen as more reliable by a user: A continuous system or logic is designed to make less "unforeseeable" jumps, and is thus seen as more reliable by a user.

(c) A (uniformly) continuous system is much easier to test: If we know the "steepness" of the system, i.e. the maximal quotient o/i of output difference vs input difference, and if we consider a certain amount of error e still tolerable, it suffices to test the system on a set of representative examples, which can have distance e*i/o.

Moreover, if we have verified that the system behaves properly on this set of examples, we know that the system is well-behaved up to error e everywhere - we do not have to rely on some lucky intuition which led us to choose the right set of examples.

Let us look at some ramifications:

1.4 Further applications of topology and of our construction in logics

Precision of Information: The measurement.11 is more precise than the measurement.1?, but.?1 does not really tell us anything. Likewise, $p_0 \wedge p_1$ is more precise than p_0, but p_1 alone may tell us almost nothing - when p_0 is the most important information, determining the order of magnitude, so to say.

Inconsistency and Theory Revision: Suppose we have an inconsistent database of two theories, $T_0 := \{p_0, p_1, p_2\}$, $T_0 := \{p_0, \neg p_1, p_2\}$, where reliability decreases with increasing index of the $p_i's$. It makes sense to say that we believe in p_0, as it is safe, p_1 is contradicted, but to believe in p_2 makes no sense: It is like being told that your scales err on the kilos, and still believe in the grams. So, $T_0 + T_1 = \{p_0\}$, and not $= \{p_0, p_2\}$, as most Theory Revision approaches will say (cum grano salis). So we avoid here unnecessary (even ridiculous) precision. Moreover, $T'_0 := \{p_0, p_1, p_2\}$, $T'_1 := \{\neg p_0, p_1, p_2\}$ are in a much stronger sense inconsistent than T_0 and T_1 are, so we have also a degree of inconsistency.

Approximation: Given a topology, we can say that a sequence of logics $\mid\sim_n$ approximates a logic $\mid\sim$ (or even that it does so uniformly, given the necessary machinery). Knowing e.g. that in n seconds, we can compute $\mid\sim_n$, and $\mid\sim_n \rightarrow \mid\sim$ (maybe uniformly so, and maybe knowing the amount of difference), may be a very nice and comforting result. After all, classical computing did just that: approximate complex functions by simpler ones, with much success. So why not approximate a complex logic with simpler ones, easy to compute, as long as we know what we are doing - i.e. what the quality of approximation is? And, moreover, if we want to speak about such approximations, we need a notion of proximity, which topology provides. Thus, in automated reasoning, where complexity is often a crucial issue, one would often be content to have a sufficient approximation, e.g. a sufficiently precise diagnosis. And thus we need the topological notions of proximity.

Semantics: Once we have a notion of approximation, we have a new notion of semantics, too: Maybe $\mid\sim$ has no "good" semantics (or none at all), but we know that it is a good approximation to $\mid\sim'$, sufficient for our purposes, which has a good semantics $\models'$. We can say: What I really mean, is $\models'$, and $\mid\sim$ is a good approximation, and $\mid\sim$ has e.g. nice computational properties. The author feels that this is a common ground on which the "logical purists" and the "engineering fraction" in AI (see e.g. the discussion in Artificial Intelligence 47 [Nil91], [Bir91]) could agree. The latter obtain what they need: a working tool, the former a clean theory which speaks in precise notions of differences, semantics etc.

1.5 Average difference between two logics

The topology defined on the set of theories (of a given language) allows us to express continuity (a form of well-behavedness) of a logic, express the distance (e.g. at one point, or maximal, minimal - when defined) between two logics, convergence of a sequence of logics, even uniform convergence etc. But, maybe we are more interested in the *average* behaviour, that is, we feel e.g. that even strong divergence of two logics is acceptable if it extends only over a small area. We might want to speak about the average difference of two logics, say that a sequence of logics approximates another logic on the average etc. The mean of a real-valued function is, of course, its integral. So, given we find some way to assign a real value to the difference of two theories, $T \triangle T'$, we define for two logics $\mid\sim$ and $\mid\sim'$ their difference $\mid\sim \triangle \mid\sim'$ by $(\!\mid\sim \triangle \mid\sim')(T) := \dashv\sim (T) \triangle \mid\sim'(T)$ and ask for $\int (\!\mid\sim \triangle \mid\sim)$ ($\int (\!\mid\sim \triangle \mid\sim) dT$ if you prefer) over the set of theories. To provide the necessary machinery, we have to construct a $\sigma - algebra$ $\mathcal{A}$ of measurable sets on the set of theories of the language, and construct a measure μ on $\mathcal{A}$. (These standard definitions will be repeated below.) The difference between two theories $T \triangle T'$ will be the exact numerical analogue of our topological construction. Moreover, the measurable sets and the measure will also be defined closely following the topology, and we will be able to obtain that for two continuous logics $\mid\sim$ and $\mid\sim'$ $\int (\!\mid\sim \triangle \mid\sim')$ is defined.

Example of the utility of the average distance between two logics: In some applications, the average distance is not a suitable measure to test the relative quality of a system. E.g. for ethical reasons, one grave though rare case of failure might not be acceptable.

In calculating production costs or product quality, where a better quality will give a higher price, however, this may just be the right measure.

In production processes, there are often several strategies possible to achieve a desired result. The author has once programmed a control system for the production of plastic sheets, which offered several possible modifications of the geometry of the production line, to obtain a desired profile of the produced sheet. In addition, temperature and composition of the plastic material could be modified. Some parameters of the output sheet were more important than others, and some modifications of the product could only be achieved by modifying some specific control parameters. Moreover, some modifications in the production process were more costly than others - e.g. expensive additives to the plastic material, wear of the production line, delay of the adjustment to become effective, cost of heating etc.

We have thus a hierarchy of parameters - or propositional variables - resulting in different sales prices or demanding different production investments. As different strategies of control are possible, it is not a simple question of elementary arithmetic to find the best solution.

Given two expert systems, or logics, or experts, which provide control strategies for this process, the adequate comparison is then between the average costs and prices.

1.6 More applications of the distance between theories

A topology on the set of theories (for a given language) gives us a measure of the distance between theories. We discuss some applications of such a notion of difference, and of related concepts.

A) Automatic Generation of Abstracts: Assessing the quality of possible candidates We see an abstract T' of a text T as a theory which is in some sense still close to T - though a simplification. Let us put this a little in perspective and give a rough idea how such an abstract generation might proceed. Like for all difficult and complex problems it is important to cut the problem into well-defined substeps, use at least initially a simple scenario, provide enough background information, thus making at least some small progress possible. A suitable scenario might be a scientific area, with clear concepts and structure, and relatively simple and easy to understand information.

I assume that the original text has somehow been transformed into an internal representation T, say in FOL, so T is a FOL theory, i.e. a set of FOL formulas. The task is to find T' which is, in some sense to be specified, the FOL representation of an abstract of T.

As background information, humans could not do without knowing
- what is important
- which information is a refinement of some other information (e.g. blue/

dark blue), leading maybe to structured sorts and structured predicates. This knowledge has to be given explicitely.

The aim of the abstract has to be defined: On the one extreme, one might be interested in extracting only the "new" information of a text, on the other hand, applying this to a textbook of a known area will give the empty set, and this is clearly not always wanted. (Looking for the "new" might involve looking for violated defaults.)

Modesty might involve not looking for a cognitively adequate result (and, how is this made precise? - see, however, below), but first to a content-adequate abstract.

The abstraction process might be composed of several steps, e.g.:

- Finding (local) generalizations like dark blue → blue. This step seems to have something to do with learning.
- Finding candidates for the abstract, i.e. sets of formulas from the original text T, and the generalizations.
- Assessing the candidates (and choosing the "best").

Assessing the candidates means assessing the compromise between correspondence with the original and complexity (or granularity). A text, which is "close" to natural categories might be considered simple, even if it is relatively long. The use of defaults can reduce the complexity. Exactness or correspondence can be measured by the difference of the two theories T and T' in the suitably chosen topology.

A1) Several other problems are more or less closely related to abstract generation:

- answering why-questions: In principle, anything between "I found a proof" and dumping the database is an answer. A good one is probably something like an abstract of a proof (argument).
- giving reasons for decisions
- giving an overview of the knowledge, or of some part of it, stored in a database
- structuring text (in a "cognitively adequate" way) going e.g. from the more general to the more specific, the former a kind of abstract

- conversely, such structure constitutes part of the coherence of a text
- more complex phenomena of granularity
- text understanding (forming abstracts while processing, which are used for further analysis, when the text as a whole is too long).

A2) Any answer to the abstract generation problem might also shed light on:

- human organization of knowledge (cognitive adequacy of the results)
- human use of defaults
- standard assumptions (domain specific or not) in human texts, e.g. violated in the output, or possible use thereof for simplifications.

B) Graceful degradation The folklore notion of graceful degradation describes the desirable property of e.g. an expert system to lose its competence gradually outside its intended domain of application - as a human expert will recur outside his field of expertise to his less precise and less reliable, but still useful common knowledge.

It seems too restrictive to limit "graceful degradation" only to the transition between expert and common knowledge. Neither experts nor people in general have a homogenous precision over their field of knowledge. Limiting the discussion to the abovementioned case might prevent a natural formulation and solution of the problem. A hierarchy of knowledge, as given e.g. by the order discussed above seems a natural and flexible tool to obtain such grace - where "graceful" could in the end be made precise by a first derivative!

C) Important Lemmata and Partial Proofs We argue on the background of our intuition that even mathematical proofs are usually not $performed$ by a rigorous concatenation of axioms and rules, and mathematical theories do not consist of such in our mind, but by putting somehow manageable more or less good approximations to the true theory together, e.g. "small models".

From this point of view, an important lemma is not a lemma in whose "inferential neighbourhood" - in the author's opinion a somewhat doubtful notion - lie other "important" results, but its translation into intuitive terms adds really something important *new* to the intuitive theory, it is *not* a consequence of the already present intuitive knowledge. Thus an important lemma is only important on the background of some fixed intuition, which seems to be well in accord with the use of the notion of importance.

A partial proof can then be seen e.g. as
- a complete proof in the approximative theory
or, weaker, as
- a sequence $\phi_0 \ldots \phi_n$ where $\mid\sim$ $\phi_{2i} \rightarrow \phi_{2i+1}$ in the approximative theory and ϕ_{2i+1}, ϕ_{2i+2} are "similar" in a suitable topology.

A partial existential proof - e.g. a counterexample - can be seen as a construction
- which has approximately the desired properties
- exists "approximately" (is consistent wrt. approximative logic)

We turn to the technical part.

1.7 Relation between the motivational and the technical part

Let me emphasize that the technical development below is but a modest first step towards the constructions whose utility we hopefully have demonstrated above. Thus, in particular, we refrain ourselves to propositional logic, with, however, an indication how to extend our definitions and results to the first order case - see after Corollary 2.4 below.

We shall define a topology based on a partial order on the propositional variables. We proceed semantically, by first defining the S-neighbourhood of a set of models (Definition 2.1, $U_S(G)$)) and extend this definition in a natural way to theories (Definition 2.6, $U_S(T)$). Fact 2.1 - Lemma 2.7 discuss basic topological properties of our construction. Example 2.2 - Lemma 2.8 show that our construction is not really a metric space, but almost so. Lemma 2.13 - Example 2.4 discuss various logics - continuous or not. In the rest of the paper, we define a $\sigma - algebra$ of measurable sets for ω many propositional variables, define a numerical difference between theories, and show that the average difference, i.e. the integral of the difference, between two continuous logics on that space is defined.

I would like to emphasize that I do not think that the topologies and other constructions given below are "the best" for applications. They are theoretical constructions which show that the things we have been discussing in the introduction can be made to exist, in a hopefully not outright preposterous way. Moreover, they are generic, and allow a lot of manipulation. They are intended to start a discussion, and the author would be happy to revise them in the light of demands from applications.

The pure topological constructions discussed below are more or less known. They are rather straightforward, we give them in full detail to make the text self-contained. Of course, we then emphasize the connections with logic in the further treatment.

1.8 Basic definitions from topology, set and measure theory

As said above, we shall almost exclusively work with propositional logic.

Assume a propositional language $\mathcal{L}$, given by its variables $v(\mathcal{L})$. For a $\mathcal{L}-theory$ (a set of $\mathcal{L}-formulas$) T, $\overline{T}$ shall denote the set of its classical propositional consequences, $\overline{\overline{T}}$ the set of its consequences under some logic$\mid\!\sim$, whose properties - e.g. continuity - are to be examined, so $\overline{\overline{T}} := \{\phi : T \mid\!\sim\ \phi\}$. $Th_{\mathcal{L}}$ shall denote the set of $\mathcal{L}-theories$.

$\mathcal{L}'s$ classical models m are defined as usual: For each $p \in v(\mathcal{L})$, we fix arbitrarily $m \models p$ or $m \not\models p$. The definition is extended inductively to formulas by $m \models \neg\phi$ iff $m \not\models \phi$ and $m \models \phi \wedge \psi$ iff $m \models \phi$ and $m \models \psi$ (and $m \models \phi \vee \psi$ iff $m \models \phi$ or $m \models \psi$ unless $\vee$ is introduced via $\neg$ and $\wedge$ etc.). Thus, each classical model is fully determined by its values on $v(\mathcal{L})$, in other words, classical models can be identified with functions from $v(\mathcal{L})$ to {true,false} or 2={0,1}. We denote by $M_{\mathcal{L}}$ the set of all classical $\mathcal{L}-models$, and by M(T) the set of the models of some $\mathcal{L}-theory$ T, i.e. the set of those $m \in M_{\mathcal{L}}$ such that $m \models \phi$ for all $\phi \in T$. $\boldsymbol{D}$ shall denote the sets of $\mathcal{L}-models$, which are definable by some $\mathcal{L}-theory$, i.e. $X \in \boldsymbol{D}$ iff there is a $\mathcal{L}-theory$ T such that X is exactly the set of its models. E.g. in [Sch92], an example is given for a set of models which is not definable by a theory ($v(\mathcal{L})$ is then infinite).

A logic is a function from theories to theories $f(T) := \overline{\overline{T}}$. To avoid problems with (classically) equivalent reformulations, we work on the set of models of theories, so we assume for all logics to be considered that $\overline{\overline{T}}$ is closed under classical logic, and that for classically equivalent T, T' $\overline{\overline{T}} = \overline{\overline{T'}}$ for all theories T and T', we may thus consider a logic equivalently as a function $f : \boldsymbol{D} \to \boldsymbol{D}$. So, to speak about continuous logics, we need a topology on $\boldsymbol{D}$. As said above, we can identify classical models with functions $f : v(\mathcal{L}) \to 2$, so we have to look for topologies on $\boldsymbol{D} \subseteq \mathcal{P}(2^{v(\mathcal{L})})$ - $\mathcal{P}$ denotes the power set operator. Of course, one may look at non-classical models, or describe models by some other set of properties than the propositional variables, so we start a little more general on $\mathcal{F} \subseteq \mathcal{P}(D^X)$, were D and X are just sets. We examine in detail one method of constructing topologies for $\mathcal{F}$, first in the abstract setting, and return to logic only later on. The reader will see that already the abstract frame gives some strong results for our construction.

For the reader's convenience, we shall now repeat the basic definitions needed from general (point set) topology, set theory (ordinals, more precisely), and the theory of measure and integration. Standard references are e.g. [Kel75] or [Eng77] for topology, [Jec78] for set theory, and [Hal50] for the theory of measure and integration.

A topological space is a pair $< X, \mathcal{T} >$, where $\mathcal{T} \subseteq \mathcal{P}(X)$. The elements of $\mathcal{T}$ are called open sets. $\mathcal{T}$ has to satisfy (a) $\emptyset, X \in \mathcal{T}$, (b) $\mathcal{A} \subseteq \mathcal{T}$ finite $\to \bigcap \mathcal{A} \in \mathcal{T}$, (c) $\mathcal{A} \subseteq \mathcal{T} \to \bigcup \mathcal{A} \in \mathcal{T}$. So $\mathcal{T}$ is closed under finite intersections and arbitrary unions. Fix now $< X, \mathcal{T} >$. $A \subseteq X$ is called closed iff $X - A \in \mathcal{T}$, and clopen iff A is closed and open. Thus, $\emptyset$ and X are always clopen. $\boldsymbol{B} \subseteq \mathcal{T}$ is called a basis for $\mathcal{T}$ iff (a) all $A \in \mathcal{T}$ are the union of some elements from $\boldsymbol{B}$, (b) $\boldsymbol{B}$ is closed under finite intersection. $\boldsymbol{B} \subseteq \mathcal{P}(X)$ is said to generate a topology $\mathcal{T}$ iff $\mathcal{T}$ is the least system satisfying (a)-(c) above and containing $\boldsymbol{B}$. $\boldsymbol{B} \subseteq \mathcal{T}$ is called an open cover of X, iff $\bigcup \boldsymbol{B} = X$. $A \subseteq X$ is called a neighbourhood of $x \in X$ iff there is $B \in \mathcal{T}$ such that $x \in B \subseteq A$. Obviously, the system of neighbourhoods of x forms a filter. $< X, \mathcal{T} >$ is said to be compact iff any open cover $\boldsymbol{B} \subseteq \mathcal{T}$ of X contains a finite subset $\boldsymbol{B}' \subseteq \boldsymbol{B}$, which is still an open cover of X. (Sometimes, T_2 - see below - is also demanded.) If $\mathcal{A} \subseteq \boldsymbol{B} \subseteq \mathcal{P}(X)$ such that for all $B \in \boldsymbol{B}$ there is $A \in \mathcal{A}$ with $A \subseteq B$, then $\mathcal{A}$ is said to be dense in $\boldsymbol{B}$. A function $d : X \times X \to \Re$ - where $\Re$ is the set of reals - is called a metric on X iff (a) $d(x, y) \geq 0$, (b) d(x,y)=0 iff x=y, (c) d(x,y)=d(y,x), (d) $d(x, z) \leq d(x, y) + d(y, z)$ for all $x, y, z \in X$. d defines a topology on X by taking the topology generated by the system of open circles $U_r(x) := \{y \in X : d(x, y) < r\}$, $r \in \Re$, $x \in X$. $< X, \mathcal{T} >$ is called a metric space iff there is a metric which generates its topology. We now introduce the separation properties:

1. $< X, \mathcal{T} >$ is said to satisfy T_0 iff for all $x, x' \in X$ there is $A \in \mathcal{T}$ such that $x \in A$ and $x' \notin A$ or $x' \in A$ and $x \notin A$

2. $< X, \mathcal{T} >$ is said to satisfy T_1 iff for all $x, x' \in X$ there are $A, A' \in \mathcal{T}$ such that $x \in A$ and $x' \notin A$ and $x' \in A'$ and $x \notin A'$

3. $< X, \mathcal{T} >$ is said to satisfy T_2 iff for all $x, x' \in X$ there are $A, A' \in \mathcal{T}$ such that $x \in A$ and $x' \in A'$ and $A \cap A' = \emptyset$

4. $< X, \mathcal{T} >$ is said to satisfy T_3 iff for all $x \in X$, $B \subseteq X$ closed, $x \notin B$, there are $A, A' \in \mathcal{T}$ such that $x \in A$ and $B \subseteq A'$ and $A \cap A' = \emptyset$

5. $< X, \mathcal{T} >$ is said to satisfy T_4 iff for all $B, B' \subseteq X$ closed, $B \cap B' = \emptyset$, there are $A, A' \in \mathcal{T}$ such that $B \subseteq A$ and $B' \subseteq A'$ and $A \cap A' = \emptyset$.

Given two topological spaces $< X, \mathcal{T} >$, $< X', \mathcal{T}' >$, a function $f : X \rightarrow X'$ is called continuous, iff for all $A' \in \mathcal{T}'$ $f^{-1}[A'] \in \mathcal{T}$ ($f^{-1}[A'] := \{x : f(x) \in A'\}$). In such situations, we will also write $f :< X, \mathcal{T} > \rightarrow < X', \mathcal{T}' >$ to note dependency from the topologies. Given two metric spaces $< X, \mathcal{T} >$, $< X', \mathcal{T}' >$ defined by d and d', a function $f :< X, \mathcal{T} > \rightarrow < X', \mathcal{T}' >$ will be called uniformly continuous iff for all $r' > 0$ there is $r > 0$ such that for all $x \in X$ $f[U_r(x)] \subseteq U_{r'}(f(x))$ ($f[A] := \{f(x) : x \in A\}$).

A set α is called an ordinal, iff (a) α is transitive, i.e. $x \in y \in \alpha$ implies $x \in \alpha$, (b) α is well-ordered by the $\in -relation$. $0 := \emptyset$ is the least (ordered by $\in$) ordinal, 1:={0}, 2:={0,1}. Each ordinal is the set of its predecessors, and we also write $\alpha < \beta$ for $\alpha \in \beta$, and $\alpha \leq \beta$ for $\alpha = \beta$ or $\alpha < \beta$ when α, β are ordinals. ω is the least infinite, and ω_1 the least uncountable ordinal (uncountable as a set).

$\mathcal{A} \subseteq \mathcal{P}(X)$ is called a $\sigma - algebra$ over X iff (a) $X \in \mathcal{A}$, (b) $A \in \mathcal{A} \rightarrow X - A \in \mathcal{A}$, (c) $\{A_i : i \in \omega\} \subseteq \mathcal{A} \rightarrow \bigcup\{A_i : i \in \omega\} \in \mathcal{A}$. The definition of a $\sigma - algebra$ generated by some $\boldsymbol{B} \subseteq \mathcal{P}(X)$ parallels that of a generated topology. $\mathcal{A}$ is called an algebra iff it satisfies (a) and (b) above, but is closed under finite unions only. If $\mathcal{A}$ is a $\sigma - algebra$, a function $\mu : \mathcal{A} \rightarrow \Re$ will be called a measure iff (a) $\mu(\emptyset) = 0$, (b) $\mu(A) \geq 0$ for all $A \in \mathcal{A}$, (c) If $\{A_i : i \in \omega\} \subseteq \mathcal{A}$ is a set of pairwise disjoint sets, then $\mu(\bigcup\{A_i : i \in \omega\}) = \Sigma\{\mu(A_i) : i \in \omega\}$. Given two spaces $< X, \mathcal{A} >$, $< X', \mathcal{A}' >$ where

$\mathcal{A}, \mathcal{A}'$ are $\sigma - algebras$, and a function $f : X \rightarrow X'$, we call f measurable, iff for all $A' \in \mathcal{A}'$ $f^{-1}[A'] \in \mathcal{A}$. The standard $\sigma - algebra$ $\boldsymbol{B}$ on $\Re$ is that generated by the open intervals (its elements are called Borel sets), and its standard measure is defined by taking the lengths of the intervals as start, and extending to disjoint unions. A function $f :< X, \mathcal{A} > \rightarrow < \Re, \boldsymbol{B} >$ is called simple iff it is measurable and has only finitely many different values. $\mathcal{E}^*$ will then be the set of those $f : X \rightarrow \Re$ such that there is a sequence of simple $f_i, i < \omega$ with $f_i(x) \leq f_j(x)$ for all $x \in X$ and all $i \leq j$ and $f(x) = sup\{f_i(x) : i < \omega\}$ for all $x \in X$. $\mathcal{E}^*$ is the set of measurable functions from $< X, \mathcal{A} >$ to $< \Re, \boldsymbol{B} >$. Given a measure μ on $\mathcal{A}$, the integral of a simple function is defined as $\Sigma\{\mu(f^{-1}[y]) * y : y \in ran(f)\}$ (ran(f) its range). This is extended to $f \in \mathcal{E}^*$ by $\int f d\mu := sup\{\int f_i d\mu : i \in \omega\}$, where the f_i approximate f from below as above, this is the definition of the Lebesgue integral.

2 TECHNICAL DEVELOPMENT

2.1 Outline of the technical part

A said in the introduction, we first define a topology on the (definable) sets of models, which is then carried over to theories. As said there too, we work, in slight generalization, on a space of functions D^X, instead of on the set of models of a given propositional language. X stands for the set of propositional variables, and D generalizes the set {true,false} or 2={0,1}.

Assume for the moment X to be totally ordered by some order < (actually, we can work with a partial order), expressing the degree of relevance or importance of X's elements. For several notational reasons we assume the more important elements to be the smaller ones. For simplicity, let in this outline X be finite, $x_0 < x_1 < \ldots < x_n$. Our basic idea was to say that $f, g : X \rightarrow D$ are the closer to each other, the more they agree on the important elements. If we define e.g. $f(x_0) = g(x_0)$, $f(x_1) \neq g(x_1)$, and $f'(x_0) = g'(x_0)$, $f'(x_1) = g'(x_1)$, f' and g' will be considered closer to each other than f and g - irrespective of their behaviour on the other $x_i's$ - so there is no compensation between the different $x_i's$. If f" and g" disagree already on x_0, they will have maximal difference. We max express intuitively the distance in the above example e.g. by d(f,g):=1/1, and $d(f', g') \leq 1/2$, as f and g agree up to the first x_i, f' and g' at least up to the second x_i. Theories correspond to sets of models, or sets of functions, so we are interested in distances between sets of functions. Our construction generalizes in a straightforward way to sets of functions $F, G \subseteq D^X$: As above, we may say that d(F,G)=1/1 iff for all $f \in F$ there is $g \in G$ such that $f(x_0) = g(x_0)$ and vice versa, so $d(F, G) \leq 1/1$, but there is $f \in F$ such that there is no $g \in G$ with $f(x_1) = g(x_1)$ or vice versa, so $d(F, G) \geq 1/1$. Note that the distance between F and G is largely independent of their cardinalities, one matching element suffices, we consider the restrictions of $f \in F$, $g \in G$ on the sets $\{x_0\}$, $\{x_0, x_1\}$ etc. Obviously, for $f \neq f' \in F$ some restrictions may be identical. Given some order-initial segment $S \subseteq X$, we denote by $F\lceil S$ the set of all $f\lceil S$, $f \in F$ - $f\lceil S$ the usual restriction of a function to a subset of its domain. The S-neighbourhood of F, $U_S(F)$ will then be the set of all $G \subseteq D^X$ such that $F\lceil S = G\lceil S$, i.e. which are identical up to at least S. If S gets larger, $U_S(F)$ will usually become smaller. We take the $U_S(F)$, S an initial segment of X, $F \subseteq D^X$ as base sets of our topology, in fact, if $U_S(F) \cap U_{S'}(F') \neq \emptyset$, then $U_S(F) \subseteq U_{S'}(F')$ or vice versa, all open sets will be unions of such $U_S(F)$. (This is essentially the material up to Corollary 2.4.)

The reader less interested in the topological development may skip the material discussed in this pararaph: It is a standard investigation in topology to examine separation properties and compactness, we do this in Lemma 2.5 - Lemma 2.7. We have spoken above of the "distance" between two functions f,g, this leads naturally to the question whether or not our spaces are metric. It is trivial to show that they are not - our order $<$ can be too deep - but they are almost so, as the initial segments S give a uniform distance measure. In particular, we can define uniform continuity of a function between two such topological spaces and show that any continuous function on a compact space is uniformly continuous. (This is discussed in the material from Example 2.2 to Lemma 2.8.)

The subsequent material up to Corollary 2.11 carries the construction over to logics, to theories more precisely. X is now presumed to be the set of propositional variables of some language $\mathcal{L}$, ordered by importance, the most important ones the smallest. If $S \subseteq X$ is an initial segment, T an $\mathcal{L} - theory$, then $U_S(T)$, the S-neighbourhood of T, consists of all theories T' such that $M(T') \in U_S(M(T))$, where M(T) denotes the set of models of T etc. We can transform T into a classically equivalent theory

T^d, consisting of all classical consequences of T which are pure disjunctions of literals. $T^d \dagger S$ will be the subset of T^d consisting of all such $\phi \in T^d$ in which only propositional variables from S occur. We have $T' \in U_S(T)$ iff $T^d \dagger S = T'^d \dagger S$, moreover $T^d \dagger S \in U_S(T)$, and $T^d \dagger S$ is the most general S-neighbour of T, i.e. $T' \in U_S(T)$ implies $T' \vdash T^d \dagger S$.

Fact 2.12 shows compactness of the theory space in a special case, and Lemma 2.13 - Example 2.4 discuss continuity of logics on our theory spaces and the relation between continuity and compactness (as a logic). The hurried reader may skip this part too.

In the last section, we define a $\sigma - algebra$ $\mathcal{A}$ on $Th_{\mathcal{L}}$ ($\mathcal{L}$ countable), and a measure $\mu : \mathcal{A} \to \Re$. To this end, we construct a sequence of partitions of $Th_{\mathcal{L}}$. We start with the trivial partition $R'_0 := \{Th_{\mathcal{L}}\}$. For $n < \omega$, we set $R'_n := \{Th(G\lceil n) := \{T : M(T) \in U_n(G)\} : G \subseteq 2^{v(\mathcal{L})}\}$. It is easily seen that this is a successive refinement: If $n' > n$ and $Th(G\lceil n) \cap Th(G'\lceil n') \neq \emptyset$, then $Th(G'\lceil n') \subseteq Th(G\lceil n)$. We assign each $Th(G\lceil n)$ inductively an interval on the real line $g(Th(G\lceil n))$, starting with $g(Th_{\mathcal{L}}) := [0,1)$, and, if $g(Th(G\lceil n)) = [x, x')$, we give each $Th(G'\lceil n+1)$ with $Th(G'\lceil n+1) \subseteq Th(G\lceil n)$ an equal share $[y, y') \subseteq [x, x')$. The size of [x,x') determines that of $Th(G\lceil n)$, i.e. $\mu(Th(G\lceil n)) := x' - x$, if we start with the standard measure on $\Re$. The definition extends naturally to the $\sigma - algebra$ $\mathcal{A}$ generated by $\bigcup\{R'_n : n < \omega\}$, defining a measure on $\mathcal{A}$. For two theories T, T', we set their difference $T \triangle T' = 0$ iff $\overline{T} = \overline{T'}$. Otherwise, there will be a least n such that $M(T)\lceil n \neq M(T')\lceil n$, and $T \triangle T'$ will be set $\frac{1}{2^n}$. Given two logics $\mid\!\sim$, $\mid\!\sim'$, we can then define their difference function $\mid\!\sim \triangle \mid\!\sim'$ by $(\mid\!\sim \triangle \mid\!\sim')(T) := (\mid\!\sim(T)) \triangle (\mid\!\sim'(T)) \in \Re$. We conclude by showing that, if $\mid\!\sim$ and $\mid\!\sim'$ are continuous wrt. our standard topology, then $\int(\mid\!\sim \triangle \mid\!\sim')$, i.e. their average difference, is defined. We use here the fact that all open sets of the topology are in the $\sigma - algebra$ $\mathcal{A}$.

2.2 The topological construction

Definition 2.1 *Let X be a set, partially ordered by $\leq$. If X is an ordinal, $\leq$ is the natural order, unless said otherwise. $S \subseteq X$ is called an initial segment of X, iff $x \in S$, $x' \in X$, $x' \leq x$ implies $x' \in S$ (i.e. S is downward closed). To avoid pathological cases, we assume $S \neq \emptyset$ in the sequel - though* *we will not be very exact about this point, it really does not matter. Let* $\Sigma \subseteq \mathcal{P}(X)$ *be a set of initial segments of X, totally ordered by set-inclusion. Let D a set,* $F := D^X =$ *the set of functions* $f : X \to D$. *For* $G \subseteq F$, $S \subseteq X$, *let* $G\lceil S := \{f\lceil S : f \in G\}$, *where* $f\lceil S$ *is the restriction of f to the subset S of its domain X. Let* $\mathcal{F} \subseteq \mathcal{P}(F)$ *be fixed. For* $G \in \mathcal{F}$, $S \subseteq X$ *let* $U_S(G) := \{H \in \mathcal{F}\colon G\lceil S = H\lceil S\}$. *Given F as above,* $\mathcal{F}_1$ *will denote the set of its singletons:* $\mathcal{F}_1 := \{\{f\} : f \in F\}$. *If* $\leq'$ *is a maybe different partial order on X,* Σ' *a set of* $\leq'$ *−initial segments of X totally ordered by set inclusion, we say* $\Sigma \preceq \Sigma'$ (Σ *is finer than* Σ') *iff for all* $S' \in \Sigma'$ *there is* $S \in \Sigma$ *such that* $S' \subseteq S$.

Fact 2.1 *Let* $G, G' \subseteq F$, $S' \subseteq S \subseteq X$, $G\lceil S \subseteq G'\lceil S$, *then* $G\lceil S' \subseteq G'\lceil S'$.

Proof: Straightforward: Let $f \in G\lceil S'$, so there is $g \in G$ such that $f = g\lceil S'$, thus $g\lceil S \in G\lceil S \subseteq G'\lceil S$, so there is $g' \in G'$ such that $g\lceil S = g'\lceil S$, but then f $= g\lceil S' = (g\lceil S)\lceil S' = (g'\lceil S)\lceil S' = g'\lceil S' \in G'\lceil S'$. □

We obtain the basic

Lemma 2.2 *1) Let* $G, G' \in \mathcal{F}$, $S' \subseteq S \subseteq X$, $U_S(G) \cap U_{S'}(G') \neq \emptyset$, *then* $U_S(G) \subseteq U_{S'}(G')$.
2) In particular, $U_S(G) \cap U_S(G') \neq \emptyset \Leftrightarrow U_S(G) = U_S(G') \Leftrightarrow G\lceil S = G'\lceil S \Leftrightarrow G' \in U_S(G)$ *(for* $G, G' \in \mathcal{F}$*).*
3) Let $G \in \mathcal{F}$, $S' \subseteq S \subseteq X$, *then* $U_{S'}(G) = \biguplus\{U_S(H)\colon H \in \mathcal{F}, G\lceil S' = H\lceil S'\}$ *- where* $\biguplus$ *denotes disjoint union -*
4) Let D be finite, $G \subseteq D^X$, $S \subseteq X$, $x \in X - S$, $G\lceil S$ *be finite, and K:=* $\{H\lceil(S\cup\{x\})\colon H \subseteq D^X, H\lceil S = G\lceil S\}$, *then* $card(K) = card(2^D - 1)^{card(G\lceil S)}$.

Proof: 1) Let $H \in U_S(G) \cap U_{S'}(G')$, $K \in U_S(G)$. By $K\lceil S = G\lceil S = H\lceil S$ and Fact 2.1, $K\lceil S' = H\lceil S' = G'\lceil S'$, so $K \in U_{S'}(G')$.
2) Note that always $G \in U_S(G)$, so $U_S(G) \neq \emptyset$, and $U_S(G) = U_S(G')$ implies $G' \in U_S(G)$.

3) Easy: Disjointness follows from 2); " $\supseteq$ " follows from 1), as $G\lceil S' = H\lceil S'$ implies $H \in U_{S'}(G)$; finally, any $H \in U_{S'}(G)$ is in $U_S(H)$.
4) If $G\lceil S = \emptyset$, and $H\lceil S = G\lceil S$, then $H = \emptyset$ (see Fact 3 below), so $K = \{\emptyset\}$. Otherwise, note that any $f\lceil S \in G\lceil S$ may be continued by $\{f\lceil S \cup \{< x, d' > \} : d' \in D'\}$ for any $\emptyset \neq D' \subseteq D$. $\Box$

Fact 2.3 *1)* $G\lceil S = \emptyset \Leftrightarrow G = \emptyset$

2) $U_S(\emptyset) = \{\emptyset\}$

3) $G \neq \emptyset \rightarrow \emptyset \notin U_S(G)$ *for all* S

Proof: 1) trivial.
2) $G \in U_S(\emptyset) \rightarrow G\lceil S = \emptyset\lceil S = \emptyset \rightarrow G = \emptyset$ by 1)
3) $\emptyset \in U_S(G) \rightarrow G\lceil S = \emptyset\lceil S \rightarrow G = \emptyset$ by 1) $\Box$

Corollary 2.4 *Let* $\mathcal{F}$, $\leq$, Σ, $\leq'$, Σ' *be as above, then*

1) $\{U_S(G)$: $G \in \mathcal{F}$, $S \in \Sigma\}$ *is the basis for a topology* $\mathcal{T} = \mathcal{T}_\Sigma$ *on* $\mathcal{F}$

2) Each $U_S(G)$ *is clopen (= closed + open) in* $\mathcal{T}$

3) Let $S \in \Sigma$, $\mathcal{H} \subseteq \mathcal{F}$, *then* $U := \bigcup\{U_S(H) : H \in \mathcal{H}\}$ *is clopen.*

4) If $\Sigma \preceq \Sigma'$ *then* $\mathcal{T}_{\Sigma'} \subseteq \mathcal{T}_\Sigma$.

Proof: 1) We have to show closure under finite intersection. Let $G, G' \in \mathcal{F}$, $S, S' \in \Sigma$, and $U_S(G) \cap U_{S'}(G') \neq \emptyset$. By prerequisite for Σ, $S \subseteq S'$ or $S' \subseteq S$. So by Lemma 2.2/1), $U_S(G) \cap U_{S'}(G') = U_{S'}(G')$ or $U_S(G)$
2) By definition, $V := \bigcup\{U_S(H)$: $H \in \mathcal{F}$, $H\lceil S \neq G\lceil S\}$ is open in $\mathcal{T}$, as its complement $U_S(G)$ is closed. Moreover, $\mathcal{F} = U_S(G) \cup V$. It remains to show $U_S(G) \cap V = \emptyset$. Assume $K \in U_S(G) \cap U_S(H)$ for some H such that $H\lceil S \neq G\lceil S$. Then $G\lceil S = K\lceil S = H\lceil S \neq G\lceil S$, Contradiction.
3) Let $\boldsymbol{G} := \{G \in \mathcal{F}$: $G\lceil S \neq H\lceil S$ for all $H \in \mathcal{H}\}$ and $V := \bigcup\{U_S(G) : G \in \boldsymbol{G}\}$, which is open. If $K \in \mathcal{F}$ is such that there is $H \in \mathcal{H}$ with $K\lceil S = H\lceil S$, then $K \in U$. Otherwise, $K \in V$. Suppose there is $H \in \mathcal{H}$, $G \in \boldsymbol{G}$ such that $U_S(G) \cap U_S(H) \neq \emptyset$, then $H\lceil S = G\lceil S$, Contradiction.
4) Immediate from Lemma 2.2/3), as the base sets wrt. Σ' are also open wrt. Σ. $\Box$

4) shows that, to a certain extent, the order does not matter. What matters here is, how far we approximate X from below. In particular, if

$\Sigma \preceq \Sigma' \preceq \Sigma$, then any function continuous wrt. $\mathcal{T}_\Sigma$ will be continuous wrt. $\mathcal{T}_{\Sigma'}$ too. An even stronger independence result is given in Lemma 2.8/2) where we show that such Σ, Σ' do not change the uniformity of continuous functions. (Of course, this does not always hold when the topology stays the same, as manipulating the usual metric by e.g. $\frac{1}{x}$ shows.)

This may be the place to indicate a possible (and straightforward) extension to the predicate logic case:

For reasons of homogenity, we work on a fixed "superuniverse" W, which we take large enough - a class, if you don't like to use downward Löwenheim/ Skolem. Our real universes will be subsets of W. More precisely: Let $\mathcal{L}$ be a first order language, T an $\mathcal{L}-theory$. We add a new unary universe predicate U to $\mathcal{L}$ and relativize all $\phi \in T$ to U, i.e. for each constant c, we add U(c), functions and predicates will be restricted to U, quantifiers too, i.e. $\forall x \phi(x)$ will be replaced by $\forall x(U(x) \rightarrow \phi(x))$, $\exists x \phi(x)$ by $\exists x(U(x) \wedge \phi(x))$. Without loss of generality, we now assume that $\mathcal{L}$ has only predicate symbols. As in the propositional case, we assume an order $<$ to be given on the predicates - including U - which determines the degree to which models differ. For simplicity, assume them to be totally ordered and finite in number: $p_0 < p_1 < \ldots p_n$. We can now say that two $\mathcal{L}-structures$ m and m' have distance i, iff $p_0 \ldots p_i$ have the same extension in m and m', but p_{i+1} is interpreted differently in m and m'. Note that our treatment of the universe as a new predicate allows us to attribute arbitrary importance to the preservation of the universe - depending on where we place U in our order. Obviously, this development is totally in parallel to the propositional case. We may, of course, add refinement by considering also the size of the area on which the predicates differ (by cardinality or subset-relation). The difference of two theories is then again determined by the difference of the sets of their models.

Corollary 2.4/2) has strong consequences for separation properties:

Lemma 2.5 *1) $\mathcal{T}$ satisfies T_3 and T_4*

2) If $\mathcal{T}$ satisfies T_0, then it satisfies T_2, so T_0, T_1, T_2 coincide for $\mathcal{T}$.

Proof: 1) T_3: Let $A \subseteq \mathcal{F}$ be closed in $\mathcal{T}$, $G \notin A$. $\mathcal{F} - A$ is open, and contains G, so there is some $U_S(H) \subseteq \mathcal{F} - A$, which contains G (and even by the above $U_S(H) = U_S(G)$). As $U_S(H)$ is clopen, $U_S(H)$ and $\mathcal{F} - U_S(H)$ are as desired.

T_4: Let $A, A' \subseteq \mathcal{F}$ be closed in $\mathcal{T}$, $A \cap A' = \emptyset$. So there are I, I' with $\mathcal{F} - A = \bigcup\{U_{S_i}(G_i) : i \in I\} \supseteq A'$ and $\mathcal{F} - A' = \bigcup\{U_{S_j}(G_j) : j \in I'\} \supseteq A$. Let $I_0 := \{i \in I$: $U_{S_i}(G_i) \cap A' \neq \emptyset\}$ and $I'_0 := \{j \in I'$: $U_{S_j}(G_j) \cap A \neq \emptyset\}$. Obviously, $U := \bigcup\{U_{S_i}(G_i) : i \in I_0\}$ and $U' := \bigcup\{U_{S_j}(G_j) : j \in I'_0\}$ are open, $A' \subseteq U$, $A \subseteq U'$, $A \cap U = \emptyset$, $A' \cap U' = \emptyset$. Suppose $U \cap U' \neq \emptyset$. So there is $i \in I_0$, $j \in I'_0$ with $U_{S_i}(G_i) \cap U_{S_j}(G_j) \neq \emptyset$. If $S_i \subseteq S_j$, then $U_{S_j}(G_j) \subseteq U_{S_i}(G_i)$. But $U_{S_j}(G_j) \cap A \neq \emptyset$, so $U \cap A \neq \emptyset$, Contradiction. The case $S_j \subseteq S_i$ is symmetrical.
2) Let $G, G' \in \mathcal{F}$, $G \neq G'$, and suppose e.g. that there is open U with $G \in U$, $G' \notin U$. So there is $U_S(H)$ such that $G \in U_S(H) \subseteq U$. But $U_S(H)$ is clopen.
$\Box$

For the intuition, the reader is referred to Diagrams 1 and 2 of the Appendix. A base set $U_S(G)$ can be identified with $G\lceil S$, so any open set with $G_i\lceil S_i$, $i \in I$. We can add a root and consider $\mathcal{F}_1$ (for simplicity) as a tree T of branching factor card(D). So $G\lceil S$ is a - usually non-cofinal - branch in T, shown by full lines in Diagram 1. The open set $\bigcup\{U_{S_i}(G_i) : i \in I\}$ consists of all continuations of some $G_i\lceil S_i$, speaking "botanically", of all end-growths from some $G_i\lceil S_i$ - dotted lines in Diagram 1. The elements of the complement, i.e. of a closed set, are "side-growths" of the $G_i\lceil S_i$ - shown by dotted lines in Diagram 2, they are not continuations of any $G_i\lceil S_i$.

Definition 2.2 *1) Σ is called cofinal, iff $\bigcup \Sigma = X$.*
2) Σ is called $\mathcal{F} - cofinal$, iff for all $G, G' \in \mathcal{F}$, $G \neq G'$ there is $S \in \Sigma$ such that $G\lceil S \neq G'\lceil S$.

It is not true that Σ cofinal $\rightarrow \Sigma$ $\mathcal{F} - cofinal$ for all $\mathcal{F}$, as the following example shows (see, however, Lemma 2.9 below for $\boldsymbol{D}$).

Example 2.1 *Let $X = \Sigma = \omega$, D=2 - recall that any ordinal is the set of its predecessors - (note also that we are slightly inexact here, as $0 \in \Sigma$, and we assumed above that all $S \neq \emptyset$, but this is not important, so we just neglect it). Obviously, Σ is cofinal. For $n \in \omega + 1$ let $f_n : \omega \rightarrow 2$ be defined by $f_n(m) = 0$ iff $m < n$. Let $G := \{f_n : n < \omega\}$, $G' := \{f_n : n \leq \omega\}$, $\mathcal{F} := \{G, G'\}$. Then, for each $S = n \in \omega = \Sigma$, $f_\omega\lceil S = f_n\lceil S$, and thus for each $S \in \Sigma$ $G\lceil S = G'\lceil S$.*
$\Box$

Lemma 2.6 *1)* Σ *is* $\mathcal{F}-cofinal$ *iff* $\mathcal{T}$ *is* T_2.

2) If Σ *is* $\mathcal{F}-cofinal$, $G \in \mathcal{F}$, *then* $\bigcap\{U_S(G)\colon S \in \Sigma\} = \{G\}$.

Proof: 1) " $\rightarrow$ ": Let $G \neq G' \in \mathcal{F}$, and $S \in \Sigma$ such that $G\lceil S \neq G'\lceil S$. Then $U_S(G) \cap U_S(G') = \emptyset$ by Lemma 2.2.
" $\leftarrow$ ": Assume there is $G, G' \in \mathcal{F}$, $G \neq G'$ such that there is no $S \in \Sigma$, $G\lceil S \neq G'\lceil S$. But then $G' \in U_S(G)$ for all $S \in \Sigma$, so $\mathcal{T}$ is not T_1, so a fortiori not T_2.
2) We can argue either with T_2, or directly, as above: Let $G \neq G' \in \mathcal{F}$, and $S \in \Sigma$ such that $G\lceil S \neq G'\lceil S$. Then $G' \notin U_S(G)$. $\Box$

It is easy to find non-compact examples, e.g. by making D infinite, or by taking $\Sigma = X = \omega + 1$, D=2, and considering $\mathcal{F} := \mathcal{F}_1$. The following positive result is more interesting:

Definition 2.3 $\mathcal{F}$ *has the* $\Sigma-$*intersection property iff for all* $U = \{U_{S_i}(G_i) : i \in I\}$ $(S_i \in \Sigma,\ G_i \in \mathcal{F})$, *totally ordered by set-inclusion,* $\bigcap U \neq \emptyset$.

Lemma 2.7 *Let D finite,* $\Sigma = X = \omega$, $\mathcal{F} \subseteq \mathcal{P}(D^X)$ *such that* $\mathcal{F}$ *has* $\Sigma-intersection$ *property, then* $\mathcal{T}$ *is compact.*

Proof: Suppose not, so there is $W := \{U_{S_i}(G_i) : i \in I\}$, $S_i \in \Sigma$, $G_i \in \mathcal{F}$ with $\bigcup W = \mathcal{F}$, but for no finite $I_0 \subseteq I$ we have $\bigcup\{U_{S_i}(G_i) : i \in I_0\} = \mathcal{F}$. Define a tree on $T := \{< U_S(G), S >\colon S \in \Sigma,\ G \in \mathcal{F}, \forall S' \leq S. U_{S'}(G) \notin W\}$ by $< U_{S'}(G), S' > < < U_S(H), S >$ iff $S' < S$ and $U_S(H) \subseteq U_{S'}(G)$. $< U_S(G), S > \in T$ has height S in T: To demonstrate this, we show that $\{< U_{S'}(G), S' >\colon S' < S\}$ are exactly the predecessors of $< U_S(G), S >$. First, for $S' < S$ $< U_{S'}(G), S' > \in T$: As $< U_S(G), S > \in T$, for all $S" < S$ $U_{S"}(G) \notin W$, so a fortiori by $S' < S$ for all $S" < S'$ $U_{S"}(G) \notin W$. Second, obviously $< U_{S'}(G), S' > < < U_S(G), S >$. On the other hand, if $< U_{S'}(H), S' > < < U_S(G), S >$, then by definition $S' < S$ and $U_S(G) \subseteq U_{S'}(H)$, so $G \in U_{S'}(H)$, but then $U_{S'}(H) = U_{S'}(G)$. Note that for each $S \in \Sigma$ $\{U_S(G) : G \in \mathcal{F}\}$ is finite, as D and S are, so there are only finitely many $G\lceil S$. Thus, each level of T - which is of the form $\{< U_S(G), S >\colon G \in \mathcal{F}\}$ - is finite. But T has also elements of arbitrary height $< \omega$: Suppose not. Then there would be S, such that for all G there is $S' < S$ and $G \in U_{S'}(G) \in W$, giving a finite subcover of W. By König's infinity lemma, T has an infinite branch $B = \{< U_{S_j}(G_j), S_j >: j \in \omega\}$, and by $\Sigma -intersection$ property, there is $G \in \bigcap\{U_{S_j}(G_j) :< U_{S_j}(G_j), S_j >\in B\}$. But $G \notin \bigcup W$: Suppose $G \in U_{S_k}(G_k) \in W$ for some $k \in I$, then $S_k = S_j$ for some $j \in \omega$, $G \in U_{S_j}(G_k) \cap U_{S_j}(G_j)$, and $U_{S_k}(G_k) = U_{S_j}(G_j) \notin W$ (by Lemma 2.2/2)), contradiction. □

If any topological space $(\mathcal{F}, \mathcal{T})$ is a metric space, defined, say by a metric ρ, then $\{U_r(x) : x \in \mathcal{F}, r > 0\}$ with $U_r(x) := \{y : \rho(x, y) < r\}$ forms a basis, so for all $x \in \mathcal{F}$ there is a countable set $\{U_{\frac{1}{n}}(x) : n < \omega\}$ of open neighbourhoods of x dense in the neighbourhood filter of x, i.e. if U(x) is a neighbourhood of x, then there is $n < \omega$ such that $U_{\frac{1}{n}}(x) \subseteq U(x)$.

Thus the following example is not a metric space:

Example 2.2 *Let* $X = \Sigma = \omega_1$, *D=2,* $\mathcal{F} = \mathcal{F}_1$, $G = \{f\} \in \mathcal{F}$. *Obviously, for* $S' < S < \omega_1$ $U_S(G) \overset{\subset}{\neq} U_{S'}(G)$. *Suppose there is a countable set* $U_i(G) : i < \omega$ *dense in the neighbourhood filter V(G), so it is in particular dense in* $\{U_S(G) : S < \omega_1\}$, *and as each* $U_i(G)$ *contains some* $U_{S_i}(G)$, $S_i < \omega_1$, *there is a dense set* $\{U_{S_i}(G) : i < \omega\}$. *But, by regularity of* ω_1, $sup\{S_i : i < \omega\} = \alpha < \omega_1$, *so* $U_{\alpha+1}(G) \overset{\subset}{\neq} U_{S_i}(G)$ *for all* $i < \omega$, *Contradiction.* □

Usually, "uniform continuity" or "uniform approximation" are defined for metric spaces only, but one does not really need all properties of a metric to make such definitions. The uniformity of the construction of the $U_S(G)$ allows us to speak about uniform continuity (and approximation): all we really need is for all $G \in \mathcal{F}$ a dense subset V'(G) of the neighbourhood filter such that all V'(G) are comparable. But, of course, the systems $\{U_S(G) : S \in \Sigma\}$, being a basis of the open sets, provide such dense sets. In particular, we define and obtain:

Definition 2.4 *Let* $\mathcal{T}$ *be defined on* $\mathcal{F}$, $\mathcal{T}'$ *on* $\mathcal{F}'$ *as above.* $f : \mathcal{F} \rightarrow \mathcal{F}'$ *is called uniformly continuous for the spaces* $(\mathcal{F}, \mathcal{T})$, $(\mathcal{F}', \mathcal{T}')$ *iff* $\forall S' \in \Sigma' \exists S \in \Sigma \forall G \in \mathcal{F}. f[U_S(G)] \subseteq U_{S'}(f(G))$. *We shall also use the shorthand notation* $f : (\mathcal{F}, \mathcal{T}) \rightarrow (\mathcal{F}', \mathcal{T}')$ *to fix the topologies concerned. Simple continuity is, of course, defined as usual, i.e. by* $f^{-1}[U'] \in \mathcal{T}$ *for all* $U' \in \mathcal{T}'$, *or, equivalently, by* $\forall S' \in \Sigma' \forall G \in \mathcal{F} \exists S \in \Sigma. f[U_S(G)] \subseteq U_{S'}(f(G))$.

Lemma 2.8 *1) Let $f : (\mathcal{F}, \mathcal{T}) \to (\mathcal{F}', \mathcal{T}')$ be continuous, $\mathcal{T}$ compact, then f is uniformly continuous.*
2) If $\mathcal{T}^+$ is constructed from Σ^+, $\mathcal{T}^*$ from Σ^*, $\mathcal{T}$ from Σ, $\mathcal{T}'$ from Σ', i.e. $\mathcal{T}^+ = \mathcal{T}_{\Sigma^+}$ etc. in the notation of Corollary 2.4, and $\Sigma^+ \preceq \Sigma$, $\Sigma' \preceq \Sigma^*$, and $f : (\mathcal{F}, \mathcal{T}) \to (\mathcal{F}', \mathcal{T}')$ is uniformly continuous, then so is $f : (\mathcal{F}, \mathcal{T}^+) \to (\mathcal{F}', \mathcal{T}^*)$.

Proof: 1) Let f be continuous, fix $S' \in \Sigma'$. Thus, for $G \in \mathcal{F}$, there is $S_G \in \Sigma$ such that $f[U_{S_G}(G)] \subseteq U_{S'}(f(G))$. $\{U_{S_G}(G) : G \in \mathcal{F}\}$ is an open cover of $\mathcal{F}$, so by compactness contains a finite subset $\{U_{S_G}(G) : G \in \mathcal{F}_0\}$ which covers $\mathcal{F}$. Let $S := \bigcup\{S_G : G \in \mathcal{F}_0\}$, then $S \in \Sigma$, as Σ is totally ordered by $\subseteq$, and $\mathcal{F}_0$ is finite, and consider $U_S(G)$ for any $G \in \mathcal{F}$. For some $H \in \mathcal{F}_0$ $G \in U_{S_H}(H)$, so $U_S(G) \subseteq U_{S_H}(H)$ by $S_H \subseteq S$. Thus $f(G) \in f[U_S(G)] \subseteq f[U_{S_H}(H)] \subseteq U_{S'}(f(H))$, so $f(G) \in U_{S'}(f(H))$ and $U_{S'}(f(H)) = U_{S'}(f(G))$, thus $f[U_S(G)] \subseteq U_{S'}(f(G))$ for all G, and we are done.
2) trivial by Lemma 2.2/3. □

2.3 We turn to logics.

Let X be the set of propositional variables for some propositional language $\mathcal{L}$. (We sometimes identify propositional variables with their indices to simplify notation.) Then any (classical) model m is equivalent to a function $f_m : X \to 2$, so we can identify m and f_m. Not all sets of models correspond to theories, so we consider only $\boldsymbol{D} \subseteq \mathcal{P}(2^X)$. Moreover, m (or f_m) shall also denote the set of literals which hold in m, and for $Y \subseteq X$, $m\lceil Y$ or $f_m\lceil Y$ will also denote those which hold in m and are in Y. T etc. shall denote a $\mathcal{L} - theory$, i.e. $T \in Th_{\mathcal{L}}$.

Lemma 2.9 *If Σ is cofinal, then it is $\boldsymbol{D} - cofinal$.*

Proof: Suppose not, so there is $A, A' \in \boldsymbol{D}$, $A \neq A'$, but for all $S \in \Sigma$ $A\lceil S = A'\lceil S$. Let A=M(T) (i.e. A is the set of models of the theory T), A'=M(T'), and $m \in A' - A$ (the case $m \in A - A'$ is symmetrical). Thus, by our above assumption, for all $S \in \Sigma$, there is $m_S \in A$ such that $m\lceil S = m_S\lceil S$.

As $m \not\models T$, there is $\phi \in T$ such that $m \models \neg\phi$, but, by the inductive definition of models and finiteness of ϕ (or by compactness of classical logic), there is $X_0 \subseteq X$ finite such that $m\lceil X_0$ determines ϕ. As Σ is cofinal and X_0 finite, there is $S \in \Sigma$ such that $X_0 \subseteq S$, so $m\lceil S$ determines $\neg\phi$. But, as $m\lceil S = m_S\lceil S$, $m_S \models \neg\phi$, so $m_S \not\models T$, and $m_S \notin A$, Contradiction. $\Box$

Definition 2.5 *1) A formula ϕ will be said to be in strictly disjunctive form iff it is of the form $\pm p_{i_0} \vee \ldots \vee \pm p_{i_n}$ with p_{i_j} propositional variables. A theory T is said to be in strictly disjunctive form, iff all $\phi \in T$ are.*
2) Recall that $\overline{T}$ denotes the classical closure of T, and let $T^d := \{\phi \in \overline{T} : \phi$ is in strictly disjunctive form$\}$. Obviously, $\overline{T}$ and T^d are classically equivalent. ($\overline{T}$ would do, but T^d gives the intuitively clearer picture.)
3) Let $Y \subseteq X$, then $T\dagger Y := \{\phi \in T$: in ϕ occur only propositional variables from $Y\}$.

Note that for any model m, $m\lceil S$ decides all $\phi \in T\dagger S$. We have:

Fact 2.10 *1) $m\lceil S \in M(T)\lceil S \Leftrightarrow Con(T, m\lceil S)$ (where Con stands for classical consistency, i.e. iff $T \cup m\lceil S$ is classically consistent)*
2) $T^d\dagger S \subseteq T'^d\dagger S \Leftrightarrow M(T^d\dagger S)\lceil S \supseteq M(T'^d\dagger S)\lceil S$
3) $M(T^d\dagger S)\lceil S = M(T)\lceil S$, thus $T^d\dagger S = T'^d\dagger S \Leftrightarrow M(T)\lceil S = M(T')\lceil S$
4) $M(T^d\dagger S) = \{m : m\lceil S \in M(T)\lceil S\}$, so, beyond S, $M(T^d\dagger S)$ can take any value.

Proof: 1) $Con(T, m\lceil S) \rightarrow$ there is a model m' such that $m' \models T \cup m\lceil S$ $\rightarrow m\lceil S = m'\lceil S \in M(T)\lceil S$. $m\lceil S \in M(T)\lceil S \rightarrow Con(T, m\lceil S)$ is trivial.
2) " $\rightarrow$ ": trivial
" $\leftarrow$ ": Suppose there is $\phi \in T^d\dagger S - T'^d\dagger S$, so ϕ consists of variables from S only. By $\phi \in T^d\dagger S$, $\neg Con(T^d\dagger S, \neg\phi)$, but $Con(T'^d\dagger S, \neg\phi)$, the latter as we first formed $\overline{T'}$ to construct T'^d. Thus, there is a model $m' \models T'^d\dagger(S \cup \{\neg\phi\})$, so $m'\lceil S \in M(T'^d\dagger S)\lceil S$. If there were $m \in M(T^d\dagger S)$ such that $m\lceil S = m'\lceil S$, then $m \models \neg\phi$, as $m'\lceil S$ determines ϕ, contradiction. So $M(T^d\dagger S)\lceil S \not\supseteq M(T'^d\dagger S)\lceil S$.

3) " $\supseteq$ " is trivial, as $T \vdash T^d \dagger S$. So let $m \models T^d \dagger S$, we have to show $m \lceil S \in M(T) \lceil S$. By 1), it suffices to show $Con(T, m \lceil S)$. But if $T \vdash \neg m \lceil S$, then $\neg(m \lceil S) \in T^d \dagger S$, more precisely its strictly disjunctive form, contradicting $m \models T^d \dagger S$.
4) " $\subseteq$ ": $m \in M(T^d \dagger S) \rightarrow m \lceil S \in M(T^d \dagger S) \lceil S = M(T) \lceil S$ by 3). " $\supseteq$ ": Let m be such that $m \lceil S \in M(T) \lceil S = M(T^d \dagger S) \lceil S$. We have to show $m \models T^d \dagger S$. But $m \lceil S$ decides $T^d \dagger S$, so $m \models T^d \dagger S$. □

Definition 2.6 *We abuse notation and define* $U_S(T) := \{T' \in Th_{\mathcal{L}} : M(T') \in U_S(M(T))\} = \{T' : M(T') \lceil S = M(T) \lceil S\}$. *For* $G \subseteq 2^X$, $S \subseteq X$, *let* $Th(G \lceil S)$*:=* $\{T \in Th_{\mathcal{L}} : M(T) \lceil S = G \lceil S\} = \{T \in Th_{\mathcal{L}} : M(T) \in U_S(G)\}$. *(Thus, for* $T \in Th(G \lceil S)$, $Th(G \lceil S) = U_S(T)$, *in particular* $Th(M(T) \lceil S) = U_S(T)$.*)*

Corollary 2.11 *1)* $T^d \dagger S = T'^d \dagger S \Leftrightarrow$ *(by Fact 2.10/3))* $M(T) \lceil S = M(T') \lceil S \Leftrightarrow M(T') \in U_S(M(T)) \Leftrightarrow T' \in U_S(T)$
2) $T^d \dagger S$ *is the most general S-neighbour of T*
3) A logic f, which we also write in the more "logical" notation $.^=$*, is continuous iff* $\forall S' \in \Sigma \forall T \exists S \in \Sigma \forall T' (T' \in U_S(T) \rightarrow \overline{\overline{T'}} \in U_{S'}(\overline{\overline{T}}))$.
4) If $G, G' \subseteq F = 2^X$, $S \subseteq X$ *and* $G \lceil S \neq G' \lceil S$, *then* $Th(G \lceil S) \cap Th(G' \lceil S) = \emptyset$.
5) Let $S \subseteq S'$, $G \subseteq F$, *then* $Th(G \lceil S) = \biguplus \{Th(H \lceil S') : H \subseteq F, H \lceil S = G \lceil S\}$
6) If Σ *is cofinal, then* $\{T\} = \bigcap \{Th(M(T) \lceil S) : S \in \Sigma\}$ *(up to logical equivalence)*
7) If S is finite, then $Th(G \lceil S) \neq \emptyset$ *for any* $G \subseteq F$
8) Thus, by 4) and 7), for finite S, there is a one-one correspondence $\{G \lceil S : G \subseteq F\} \Leftrightarrow \{Th(G \lceil S) : G \subseteq F\}$ *such that* $G \lceil S \neq G' \lceil S \rightarrow Th(G \lceil S) \cap Th(G' \lceil S) = \emptyset$.

Proof: 2) Obviously, $(T^d \dagger S)^d \dagger S = T^d \dagger S$, so $T^d \dagger S \in U_S(T)$ by 1). On the other hand, if $T'^d \dagger S = T^d \dagger S$, then $T' \vdash T^d \dagger S$.

3) f is continuous iff $\forall S' \in \Sigma \forall G \in \boldsymbol{D} \exists S \in \Sigma \forall G' \in \boldsymbol{D}(G' \in U_S(G) \rightarrow f(G') \in U_{S'}(f(G))) \Leftrightarrow \forall S' \in \Sigma \forall T \exists S \in \Sigma \forall T'(M(T') \in U_S(M(T)) \rightarrow M(\overline{\overline{T'}}) \in U_{S'}(M(\overline{\overline{T}})))$.
4) trivial.
5) (as for Lemma 2.2/3) " $\subseteq$ " If $T \in Th(G\lceil S)$, then $M(T)\lceil S = G\lceil S$, and $T \in Th(M(T)\lceil S')$. " $\supseteq$ " If $T \in Th(H\lceil S')$ with $H\lceil S = G\lceil S$, then $M(T)\lceil S' = H\lceil S'$, so $M(T)\lceil S = H\lceil S = G\lceil S$. Disjointness follows from 4).
6) " $\subseteq$ " is trivial. " $\supseteq$ " Let T' $\in \bigcap\{Th(M(T)\lceil S) : S \in \Sigma\} \rightarrow \forall S \in \Sigma . T' \in Th(M(T)\lceil S) \rightarrow \forall S \in \Sigma . M(T')\lceil S = M(T)\lceil S \rightarrow$ M(T') $\in \bigcap\{U_S(M(T)) : S \in \Sigma\} = \{$M(T)$\}$ by Lemma 2.6/2 and Lemma 2.9, so $\overline{T} = \overline{T'}$.
7) If $G = \emptyset$, consider $\phi := \bot$, otherwise $\phi := \bigvee\{\bigwedge f\lceil S : f \in G\}$ and $T := \{\phi\}$. Obviously, any f such that $f\lceil S \in G\lceil S$ is a model of T, and any f such that $f\lceil S \notin G\lceil S$ makes none of the disjuncts of ϕ true. $\Box$

Fact 2.12 *$\boldsymbol{D}$ has the $\Sigma -$ intersection property for all Σ, so in particular, $\mathcal{T}$ is compact for $\Sigma = X = \omega$.*

Proof: Let $U := \{U_{S_i}(G_i) : i \in I\}$ - without loss of generality without repetitions - totally ordered by set inclusion, and define $<$ on I by $i < j :\Leftrightarrow U_{S_i}(G_i) \subset U_{S_j}(G_j)$. Let T_i be such that $G_i = M(T_i)$.
We have $i < j \rightarrow S_j \subseteq S_i$ and $T_i^d \dagger S_j = T_j^d \dagger S_j$:
By total order of Σ, $S_j \subseteq S_i$ or $S_i \subseteq S_j$, so it suffices to derive a contradiction from $S_i \subseteq S_j$. But $S_i \subseteq S_j \wedge U_{S_i}(G_i) \subset U_{S_j}(G_j) \rightarrow U_{S_j}(G_j) \subseteq U_{S_i}(G_i)$ by Lemma 2.2, Contradiction. On the other hand, $U_{S_i}(G_i) \subseteq U_{S_j}(G_j) \rightarrow G_i \in U_{S_j}(G_j) \rightarrow T_i^d \dagger S_j = T_j^d \dagger S_j$ (by Fact 2.10/3)).
Consider now T:= $\bigcup\{(T_i^d \dagger S_i) : i \in I\}$, we show $M(T) \in \bigcap U$. By Corollary 2.11/1) $M(T) \in U_{S_i}(G_i) \Leftrightarrow T^d \dagger S_i = T_i^d \dagger S_i$, so we have to show the latter for all $i \in I$. Obviously, $T_i^d \dagger S_i \subseteq T^d \dagger S_i$. Let now $\phi \in T^d \dagger S_i$, so there is finite $I_0 \subseteq I$, and $\phi_j \in T_j^d \dagger S_j$, $\bigwedge\{\phi_j : j \in I_0\} \vdash \phi$. Let $k := min(I_0)$ (by above order on I). For $j \in I_0$, $k \leq j$, so $\phi_j \in T_j^d \dagger S_j = T_k^d \dagger S_j \subseteq T_k^d \dagger S_k$, so all $\phi_j \in T_k^d \dagger S_k \subseteq T_k^d$, so $\phi \in T_k^d \dagger S_i$.
Case 1: $i \leq k$. Then $T_i^d \dagger S_k = T_k^d \dagger S_k$, so all $\phi_j \in T_i^d \dagger S_k \subseteq T_i^d$, so $\phi \in T_i^d \dagger S_i$.
Case 2: $k \leq i$. Then $\phi \in T_k^d \dagger S_i = T_i^d \dagger S_i$. $\Box$

Lemma 2.13 *Let $X = \Sigma = \omega$, f (written also $\mid\!\sim$ or $^=$) a continuous logic on X, then f is compact as a logic.*

Proof: Let $T \mid\sim \phi$, so $\phi \in T' := \overline{\overline{T}}$, we have to find $T_0 \subseteq T$ finite such that $T_0 \mid\sim \phi$. As ϕ is finite, there is $S' \in \Sigma$ such that for any model m, already $m \lceil S'$ decides ϕ, thus for $T" \in U_{S'}(T')$, $T" \vdash \phi$. By continuity, there is S such that $f[U_S(T)] \subseteq U_{S'}(T')$. In particular, $f(T^d \dagger S) \in U_{S'}(T')$, so $T^d \dagger S \mid\sim \phi$, but we have to do a little more. By finiteness of S, $T^d \dagger S$ is equivalent to a finite set of formulas, so there is some finite $T_0 \subseteq T$ such that $T_0 \vdash T^d \dagger S$. It remains to show $T_0 \mid\sim \phi$. But $T \vdash T_0 \vdash T^d \dagger S$, so $T^d \dagger S \supseteq T_0^d \dagger S \supseteq (T^d \dagger S)^d \dagger S = T^d \dagger S$, so $T_0 \in U_S(T)$ by Corollary 2.11, so $T_0 \mid\sim \phi$. □

But not vice versa:

Example 2.3 *A logic which is compact, but not continuous:*

Let $X = \Sigma = \omega$. Define

$$\overline{\overline{T}} := \begin{cases} \mathcal{L} \text{ (i.e. all } \mathcal{L}\text{-formulas)} & \text{iff } T \vdash \neg p_i \text{ for some } i < \omega \\ \overline{T} & \text{otherwise} \end{cases}$$

$\mid\sim$ is compact: Let $T \mid\sim \phi$, if $T \vdash \neg p_i$ for some $i < \omega$, then there is $T_0 \subseteq T$ finite such that $T_0 \vdash \neg p_i$, then $T_0 \mid\sim \phi$. If $T \not\vdash \neg p_i$ for any $i < \omega$, then $T \vdash \phi$. By the well-known compactness of classical logic, there is finite $T_0 \subseteq T$ such that $T_0 \vdash \phi$, so also $T_0 \mid\sim \phi$ by definition.
$\mid\sim$ is not continuous: Consider $T_0 := \{p_i : i < \omega\}$, then $\overline{\overline{T_0}} = \overline{T_0}$, and let $S' \in \Sigma$ be given, then $\mathcal{L} \notin U_{S'}(\overline{\overline{T_0}})$ (As $M(\mathcal{L}) = \emptyset$, but $M(\overline{\overline{T_0}})$ not). On the other hand, for any $S \in \Sigma$, there is $T' \in U_S(T_0)$ such that $\overline{\overline{T'}} = \mathcal{L}$, just take $T' := \{p_i : i < S\} \cup \{\neg p_i : S \leq i < \omega\}$, giving a direct contradiction to the definition of continuity. □

Example 2.4 *A logic which is continuous, but not uniformly continuous:*

Let $\Sigma = X = \omega + \omega$. Define

$$\overline{\overline{T}} := \begin{cases} \overline{T + p_\omega} & \text{iff there is } m \in \omega \text{ such that } m = \{n \in \omega : T \vdash p_n\} \\ & \text{and } T \vdash p_{\omega+m} \\ \overline{T} & \text{otherwise} \end{cases}$$

$\mid\sim$ is continuous:
We examine the different cases for T.
Case 1: $\{n \in \omega : T \vdash p_n\} \notin \omega$. Then $\overline{\overline{T}} = \overline{T}$. Let $S' \in \Sigma$ be given, consider

$S := max\{\omega, S'\}$. Let $T' \in U_S(T)$. Then $\{n \in \omega : T' \vdash p_n\} \notin \omega$, so $\overline{\overline{T'}} = \overline{T'}$. We have to show $\overline{\overline{T'}} \in U_{S'}(\overline{\overline{T}})$, but $T' \in U_S(T) \to \overline{\overline{T'}} = \overline{T'} \in U_S(\overline{T}) = U_S(\overline{\overline{T}}) \subseteq U_{S'}(\overline{\overline{T}})$.
Case 2: $\{n \in \omega : T \vdash p_n\} =: m \in \omega$, but $T \not\vdash p_{\omega+m}$. Again, $\overline{\overline{T}} = \overline{T}$. Let $S' \in \Sigma$ be given, consider $S := max\{S', \omega + m + 1\}$. Thus, $T' \in U_S(T) \to$ T' coincides with T up to $p_{\omega+m}$, thus $\overline{\overline{T'}} = \overline{T'}$. Finish as in Case 1.
Case 3: $m := \{n \in \omega : T \vdash p_n\} \in \omega$, and $T \vdash p_{\omega+m}$, so $\overline{\overline{T}} = \overline{T + p_\omega}$. Let S' be given, and consider $S := max\{S', \omega + m + 1\}$. For $T' \in U_S(T)$, $m = \{n \in \omega : T' \vdash p_n\}$, and $T' \vdash p_{\omega+m}$. So $\overline{\overline{T'}} = \overline{T' + p_\omega}$. We have to show $\overline{\overline{T'}} \in U_{S'}(\overline{\overline{T}})$. $T' \in U_S(T) \Rightarrow T'^d \dagger S = T^d \dagger S \Rightarrow_{(*)} (T' + p_\omega)^d \dagger S = (T + p_\omega)^d \dagger S$ $\Rightarrow \overline{\overline{T'}} \in U_S(\overline{\overline{T}}) \subseteq U_{S'}(\overline{\overline{T}})$. For (*), argue like this: $T' + p_\omega \vdash \psi \Rightarrow T' \vdash p_\omega \to \psi$ $\Rightarrow T \vdash p_\omega \to \psi \Rightarrow T + p_\omega \vdash \psi$ for $\psi \in \mathcal{L}_S$, as $p_\omega \to \psi$ is also in $\mathcal{L}_S$.
$\mid\!\sim$ is not uniformly continuous:
Let $S' := \omega + 1$, and assume there is such $S < \omega + \omega$ which proves uniform continuity. Let $S < \omega + m$, $m < \omega$, and consider $I_0 := m$, $I_1 := m \cup \{\omega + m\}$, $T_0 := \{p_i : i \in I_0\} \cup \{\neg p_i : i \in (\omega + \omega) - I_0\}$, $T_1 := \{p_i : i \in I_1\} \cup \{\neg p_i : i \in (\omega + \omega) - I_1\}$. Then $T_1 \in U_S(T_0)$, but $p_\omega \in \overline{\overline{T_1}} - \overline{\overline{T_0}}$, so $\overline{\overline{T_1}} \notin U_{S'}(\overline{\overline{T_0}})$. $\square$

2.4 A measure on $Th_{\mathcal{L}}$,integration of the difference between two logics

Assume in the sequel $\mathcal{L}$ to be countable with propositional variables $\{p_i : i < \omega\}$.
The following construction is by Corollary 2.11/5) a successive refinement of disjoint partitions of $Th_{\mathcal{L}}$:
Set $R'_0 := \{Th_{\mathcal{L}}\}$. Recall that $\mathcal{F} = 2^{v(\mathcal{L})}$ and set $R'_n := \{Th(G\lceil n) : G \subseteq F\}$ for $0 < n < \omega$ (Corollary 2.11/8) and Lemma 2.2/4) will give the cardinalities.) R':= $\bigcup \{R'_n : n < \omega\}$ (so each $\vartheta \in R'$ is a subset of $Th_{\mathcal{L}}$).
We define inductively $g : R' \to \boldsymbol{B}$, $\boldsymbol{B}$ the Borel sets of the real line, by $g(Th_{\mathcal{L}}) := [0, 1)$ Let $g(Th(G\lceil n)) = [x, x')$, $card(G\lceil n) = m$, say $\{H\lceil n+1 : H\lceil n = G\lceil n\} = \{h_1, \ldots, h_{3^m}\}$, then for $1 \leq i \leq 3^m$ let $g(Th(h_i)) := [x + (i-1)\frac{x'-x}{3^m}, x + i\frac{x'-x}{3^m})$.

Let some measure μ on $\boldsymbol{B}$ be given, define for $\vartheta \in R'$ $\mu(\vartheta) := \mu(g(\vartheta))$, let R be the algebra generated by R' (i.e. the closure of R' under complements and finite unions), then by construction of g (preservation of unions and disjointness back and forth), μ can be extended uniquely to a measure on R, and if $\mathcal{A}$ is the $\sigma - algebra$ generated by R (i.e. the closure of R - or R' - under complements and countable unions), then there is a unique extension of μ to $\mathcal{A}$. (This is a standard result of measure theory.)

Assume now Σ to be cofinal in ω, each $S \in \Sigma$ finite, $\Sigma = \{S_i : i < \omega\}$, $S_i \subset S_j$ for $i < j$. By Corollary 2.11/6), and countability of Σ, $\mathcal{A}$ contains all singletons (up to logical equivalence). Moreover, for each S there is n such that $S \subseteq n$. Let now $T \in Th_{\mathcal{L}}$, then $U_S(T) = Th(M(T)\lceil S) = \biguplus\{Th(H\lceil n) : H \subseteq F, H\lceil S = M(T)\lceil S\}$, a finite union of sets from R', so $U_S(T) \in R$. As the set $\{U_S(T) : T \in Th_{\mathcal{L}}, S \in \Sigma\}$ is countable (Σ is, and for $S \in \Sigma$ $\{U_S(T) : T \in Th_{\mathcal{L}}\}$ is finite by finiteness of S), each open set in $\mathcal{T}_\Sigma$ is in $\mathcal{A}$. If $\overline{T} \neq \overline{T'}$, n:= the least m such that $M(T)\lceil S_m \neq M(T')\lceil S_m$ is defined (by Lemma 2.9),

we set

$$T \triangle T' := \begin{cases} 0 & \text{iff } \overline{T} = \overline{T'} \\ \frac{1}{2^n} & \text{otherwise} \end{cases}$$

$(n\ as\ above)$

Fact 2.14 *1)* $T \triangle T" \leq max(T \triangle T', T' \triangle T")$, *and thus also*
$T_0 \triangle T_n \leq max(T_0 \triangle T_1, \ldots, T_{n-1} \triangle T_n)$
2) $T \triangle T' < T' \triangle T" \rightarrow T \triangle T" = T' \triangle T"$.

Proof: Straightforward:
1) Case 1: At least one of the values is 0: trivial. Case 2: all are > 0: Let $T \triangle T' = \frac{1}{2^n}$, $T' \triangle T" = \frac{1}{2^m}$, it suffices to show that M(T) and M(T") agree on all S_k for $k < min(n, m)$. If there is no $k < n, m$, we are done. Otherwise, for all such k $M(T)\lceil S_k = M(T')\ S_k = M(T")\lceil S_k$, as $k < n, m$.
2) Let $T' \triangle T" = \frac{1}{2^n}$, by $T \triangle T' < T' \triangle T"$, we see that for $m \leq n$ $M(T)\lceil S_m = M(T')\lceil S_m$, so $M(T)\lceil S_n = M(T')\lceil S_n \neq M(T")\lceil S_n$ and for $m < n$ $M(T)\lceil S_m = M(T')\lceil S_m = M(T")\lceil S_m$, so $T \triangle T" = \frac{1}{2^n}$. □

For two logics $\mid\sim$ and $\mid\sim'$ $\mid\sim \triangle \mid\sim'(T) :=\mid\sim (T) \triangle \mid\sim'(T)$ defines a function $\mid\sim \triangle \mid\sim' : Th_{\mathcal{L}} \rightarrow [0, 1]$, and we can define as sketched in the introduction the Lebesgue integral $\int(\mid\sim \triangle \mid\sim')$.

We conclude by showing that $\mid\sim \triangle \mid\sim' \in \mathcal{E}^*$, and thus that $\int(\mid\sim \triangle \mid\sim')$ is defined, if $\mid\sim$ and $\mid\sim'$ are continuous.

Lemma 2.15 *If* $\mid\sim$, $\mid\sim'$ *are continuous (wrt.* $\mathcal{T}_\Sigma$*), then* $\int(\mid\sim \triangle \mid\sim')$ *is defined.*

Proof: By the above, it suffices to show that $\mid\sim \triangle \mid\sim'$ is $\mathcal{A}-\boldsymbol{B}-$ *measurable*, i.e. for any Borel set X $(\mid\sim \triangle \mid\sim')^{-1}(X) \in \mathcal{A}$. As the open intervals generate the Borel sets, it suffices to show the condition for all open intervals X=(x,x'). We saw above that $\mathcal{T}_\Sigma \subseteq \mathcal{A}$, so it suffices to show that $(\mid\sim \triangle \mid\sim')^{-1}(x,x')$ is open in $\mathcal{T}_\Sigma$. Let now T be such that $(\mid\sim \triangle \mid\sim')(T) = y \in (x,x')$.
Case 1: y=0, i.e. $\mid\sim (T) = \mid\sim'(T)$, let n be such that [y - $\frac{1}{2^n}$, y + $\frac{1}{2^n}$] $\subseteq$ (x,x'). and n" such that $n \subseteq S_{n"}$. By continuity of $\mid\sim$, $\mid\sim'$, there is n' such that $\mid\sim (U_{S_{n'}}(T)) \subseteq U_{S_{n"}}(\mid\sim (T))$ and $\mid\sim'(U_{S_{n'}}(T)) \subseteq U_{S_{n"}}(\mid\sim'(T))$, thus for $T' \in U_{S_{n'}}(T)$ $\mid\sim (T')\triangle \mid\sim (T) \leq \frac{1}{2^n}$ and $\mid\sim'(T')\triangle \mid\sim'(T) \leq \frac{1}{2^n}$, so $\mid\sim (T')\triangle \mid\sim'(T') \leq$
$max(\mid\sim (T')\triangle \mid\sim (T), \mid\sim (T)\triangle \mid\sim'(T), \mid\sim'(T)\triangle \mid\sim'(T'))$
$\leq \frac{1}{2^n}$, and $\mid\sim \triangle \mid\sim'(U_{S_{n'}}(T)) \subseteq$ (x,x').
Case 2: $y = \frac{1}{2^n}$. Let n" and n' such that $n+1 \subseteq S_{n"}$, $\mid\sim (U_{S_{n'}}(T)) \subseteq U_{S_{n"}}(\mid\sim (T))$ and $\mid\sim'(U_{S_{n'}}(T)) \subseteq U_{S_{n"}}(\mid\sim'(T))$, thus for $T' \in U_{S_{n'}}(T)$ $\mid\sim (T')\triangle \mid\sim (T) \leq \frac{1}{2^{n+1}}$ and $\mid\sim'(T')\triangle \mid\sim'(T) \leq \frac{1}{2^{n+1}}$, so by $\mid\sim (T)\triangle \mid\sim'(T) = \frac{1}{2^n}$ and Fact 2.14/2) $\mid\sim (T')\triangle \mid\sim'(T') = \frac{1}{2^n}$. So again $\mid\sim \triangle \mid\sim'(U_{S_{n'}}(T)) \subseteq$ (x,x'). Thus, $(\mid\sim \triangle \mid\sim')^{-1}(x,x')$ is open in $\mathcal{T}_\Sigma$, and we are done. □

Finally, we give a very brief sketch of another way of defining a numerical difference between two theories, which is motivated by the following: Suppose we have two theories, T and T', consisting of rules and facts, which are used e.g. to take preventive action - medical diagnosis, monitoring a power plant etc. Let T be the correct theory, T' a hopefully good approximation to T. If $T \vdash \phi$, but $T' \not\vdash \phi$, we may not be warned of imminent danger, measured by some cost and probability of occurrence. Conversely, deducing $T' \vdash \phi$, when $T \not\vdash \phi$ may cause unnecessary preventive action, a certain cost, too. Going via models, we may assign each model a cost (separately for both possibilities discussed), and probability, and define the difference between the two theories as the sum (for both ways) of the sum of probability*cost over all models where they diverge.

2.5 A toy application

Suppose we are given two therapeutical expert systems, and one "real" expert, who makes no mistakes and who is taken as a standard. Suppose they differ from the medical point of view only slightly from the real expert, by their recommendation of an additional drug alleviating a minor symptom in some cases.

Describing the situation in logical terms, we may be given a set of diagnostic variables $s_1 \ldots s_n$ representing symptoms, and a number of therapeutical variables $t_1 \ldots t_k$ representing possible drugs. The expert (system) can then be described as a logic, taking as input a conjunction of literals from $s_1 \ldots s_n$, and giving as output a conjunction of literals from $t_1 \ldots t_k$.

Say that t_k stands for the unimportant drug mentioned above. So given a set of literals from $s_1 \ldots s_n, \mid\sim_1, \mid\sim_2, \mid\sim_s$ - representing the two expert systems and the real expert - give identical answers with possible exception of t_k. Order the s_i and t_i according to their index.

(a) We thus know that the maximal difference between any two of the three logics is of size 1/k, which can be tolerated from the medical point of view, so the choice of one system is as good as any other. On the other hand, we may also be interested in the average difference. For that, we multiply the expected probability - i.e. the measure - of each symptom set with the difference between the logics considered, i.e. calculate the integral of their difference. Administering the drug t_k unnecessarily causes unnecessary expenditures, vice versa, failure to administer it when adequate will affect slightly the patient's well-being, and may result in his reluctance to recommend the clinic - causing costs too. Assume this cost can be estimated reasonably. We can then make a commercially well-founded decision for one of our systems.

(b) Suppose, on the other hand, that we are still in the test phase of one expert system as above, but know already from its internal structure that the decision for $t_1 \ldots t_{k-1}$ depends only on $s_1 \ldots s_{n-m}$. In our terminology, we can characterize it as uniformly continuous for 1/k: If the input differs by less than 1/(n-m), then the output differs by less than 1/k, which is tolerable. To test the system, it suffices thus to test all values on $s_1 \ldots s_{n-m}$, instead of testing it on the full set $s_1 \ldots s_n$.

2.6 Conclusion:

We have argued for the utility and naturalness of the topological notion of distance in the context of logics. Our examples ranged from the purely theoretical - well-behavedness of a logic, notions of semantics - to considerations for the design and testing of expert systems.

We have examined one topology on the set of theories of a given language $\mathcal{L}$ in detail. The constructed space is almost metric, and uniform continuity of functions on it, i.e. of logics for $\mathcal{L}$, are definable. Standard separation properties of the space as well as continuity and uniform continuity of logics on it are investigated, so is the connection to compactness (as a logic). The uniformity of the construction allows the definition of the average difference, i.e. the integral of the difference, of two logics. It is shown that the integral of the difference of two continuous logics exists.

2.7 Acknowledgements

This paper was written while the author was at IBM Germany, Stuttgart, the final version was prepared during his stay at GMD, Bonn, Germany. He expresses his gratitude to both - and to a referee who has tried to make this paper more readable.

3 Appendix

Two Diagrams for the Intuition

References

[Bir91] L.Birnbaum, "Rigor mortis", Artificial Intelligence 47 (1991), p.57-77

[Dal88] M.Dalal, "Investigations Into a Theory of Knowledge Base Revisions: Preliminary Report", Proc. 7th Nat. Conf. AI, Minneapolis, p.475-479, 1988

[Eng77] R.Engelking, "General Topology", Warszawa, 1977

[Hal50] P.Halmos, "Measure Theory", New York, 1950

[Jec78] T.Jech, "Set Theory", Academic Press, New York, 1978

[Kel75] J.L.Kelley, "General Topology", Springer, Berlin, 1975 (original edition Princeton 1955)

[KLM90] S.Kraus, D.Lehmann, M.Magidor, "Nonmonotonic reasoning, preferential models and cumulative logics", Artificial Intelligence, 44 (1-2), p.167-207, July 1990

[LM92] D.Lehmann, M.Magidor, "What does a conditional knowledge base entail?", Artificial Intelligence, 55(1), p. 1-60, May 1992

[MW86] R.S.Michalski, P.H.Winston, "Variable Precision Logic", Artificial Intelligence 29 (1986), p.121-146

[Nil91] N.J.Nilsson, "Logic and artificial intelligence", Artificial Intelligence 47 (1991), p.31-56

[Sch92] K.Schlechta, "Some Results on Classical Preferential Models", Journal of Logic and Computation, Vol.2, No.6, Dec. 1992