A Teleological Challenge for Relevant Similarity: The Ugly Duckling of Legal Interpretation

Hans Christian Nordtveit Kvernenes

Abstract

1 Introduction

> Endelig revnede det store æg. "Pip! pip!" sagde ungen og væltede ud; han var så stor og styg. Anden så på ham: "Det er da en forfærdelig stor ælling den!" sagde hun; "ingen af de andre ser sådan ud!"
>
> *At last the big egg did crack. "Peep, peep" said the young one, and out he tumbled, but he was so big and ugly. The duck took a look at him. "That's a frightfully big duckling," she said. "He doesn't look the least like the others."*
>
> (Hans Christian Andersen, The Ugly Duckling (1843))

In Hans Christian Andersen's story 'The Ugly Duckling', the duck mother is surprised to see the strange looking duckling that came out of the big egg. It looked so different from all the other ducklings. But is it really different? By a generalisation of the Goodman's grue/bleen challenge, Watanabe (1965) arrives at what is now known as 'the ugly duckling theorem', stating that without some prior bias, any pair is as similar to each other as any other pair. In other words, any two ducklings is as similar to each other as a duckling and a swan. Despite its rather counterintuitive result, this show that there is no objective way of measure the degree of similarity between two things.

This paper intends to describe how this problem persists in the context of legal interpretation, the search 'relevant similarity' being one of its core features. In the first part it will show how this problem occurs by introducing a notion of uniquely distinguishing features. In the second part it will show that even by imposing a common formal restriction, known as the condition of efficiency or the proportionality requirement, on the similarity relations, we end up with everything being similar to everything in an infinite number of ways as long as one is clever enough to find the proper predicate. In the third part, it will sketch a solution to the given problem by describing the structure of a plausible bias and how the problem might be diminished by the introduction of higher-order interpretations and analogies.[1]

---

[1] In this paper we will understand the word 'analogy' in a very broad sense to include

1

The goal of this paper is to first describe how a theoretical problem of classification actually converts into a practical issue regarding the interpretation of legal concepts. Despite its acknowledgment in the literature of computer science, this problem seems surprisingly absent in the context of law. The notion of bias involved in legal reasoning is frequently discussed in the literature, though it is not obvious that this actually solves the problem of interpretation in itself. We seem to also rely on some notion of higher-order interpretations to make this possible.

2   Problem and uniquely distinguishing feature

In the literature, a common challenge to analogical reasoning is that in the end, everything seems somehow similar to everything. How can we then distinguish the similarities that are actually important from those that are not? To illustrate the point, we can start by identifying some common and uncontroversial similarity claims. My left hand is similar to my right hand. An apple is similar to a pear, as they both are fruit. And being even numbers, '4' is also similar to '2'. So far so good. None of the mentioned examples contradict our general intuition regarding the similarity of things. Leaving out the controversial concept of *overall similarity* for now, we will concentrate on *relative similarity* where the similarity of two (or more) things in question is judged relative to some category or aspect. Different things are in this way similar in some aspect and different in other aspects. They might also be *gradual*. Relative to practical boxing knowledge, any amateur boxer is more similar to Muhammed Ali than I am. For the sake of the argument, we also (for now) leave out any graduality involved in similarity judgments. (Morreau 2010) We are now therefore left what is seemingly the least problematic cases of similarity judgments (also leaving out anyone recognising numerical identity as a kind of similarity), namely where two things are similar (or not) relative to some properly defined category or aspect.

We will now therefore understand similarity as being two instances of the same category. This means that if we have a category A, a is similar to b if both a and b belong to the category A. Similarity is then understood by the notion of a property, aspect or predicate. To say that a is similar to b means that both a and b have some property P. The only requirement for two things to be similar is that they share some characteristic. A challenge to such notion of similarity is, as mentioned, that in logical terms an individual might even have an infinite amount of different characteristics if we understand it as predicates. Because of our understanding of similarity as being two instances of the same category, everything seems somehow similar to everything if we are just clever enough to find a proper predicate.

Think of one thing. Then think of another. Now, does it exist some aspect, property or category whereby on might judge the two things to be similar? I chose think about my car and the number the computer I am currently working on. There are of course many differences, but I only need to find one aspect uniting them. Here, there are many. They both contain metal; have buttons;

---

all kinds of legal interpretation that is dependent on a judgment of similarity. We might therefore speak about 'analogical interpretation', but for simplicity we will only use the notion of 'analogy' to describe this process.

do calculations; having been manufactured in a fabric. So maybe I chose an easy example? After all, they are both physical, technological tools (another similarity) and one swallow does not make a spring. Had we chosen another example, maybe things would have turned differently. I will therefore try again and now choose my colleague in the neighbour office and the number '5'. Might a physical human being and one abstract mathematical object have anything in common? An obvious similarity: they are both mentioned in this paper. And if this is not a convincing uniting property, we also have many more to choose from. They are both contributing to research, their name in English start with the same letter and they are both numerically distinguishable from the number '0'. The point is therefore that no matter which two things one chooses (even non-existent if one admit that they might have properties), one is always able to find some property that is shared between the two.

This might cause some problems for arguments by analogy. If everything can be similar to everything, how can some similarities be somehow 'better' than others? One attempt might be to measure the number of similarities so that an analogy is 'good' only if it does not exist another analogy with a higher number of similarities than the first one. This is a simplified description of the approach taken by Lewis (1968) and used in developing the concept of overall similarity. Though we will see that this turns out to be a rather problematic solution as it is not simply so that everything is similar to everything in some way or another. Everything seems similar to everything in an *infinite* number of ways. (Morreau 2010)

Think for example of the predicate 'being non-identical to the number 1'. We might attribute this predicate to all individuals that are not identical to the number 1, meaning most individuals. In a similar way we might take the predicate: 'being non-identical to the number 2'. This applies to all individuals that are not identical to the number 2. This includes the number 1, since if something is identical with the number 1, it is not identical to the number 2. We might continue in a similar way with 'being non-identical to the number 3' and 'being non-identical to the number 4' and so on. By this way, we might create a potentially infinite amount of predicates that will apply to any individual that we speak about, including the numbers 1, 2, 3, ... and so on. The point is that, for every two individuals, we might find an infinite number of predicates that are shared by both. This makes it impossible to measure the degree of similarity in this way. One way to solve this problem could be to reject some predicates when describing the similarity. An attempt could be to distinguish positive predicates, properties actually present, from negative predicates, the negation of the presence of certain properties. This however, does seem like a difficult task. In the previously mentioned example, we used predicates of the form 'being non-identical to ...' and 'non-identicality' might be said to contain a negation and therefore be somehow a negative predicate. However, this seems to only be an aspect of language. We could also have used a predicate of 'difference' or 'being numerically distinguishable from', where the negation is not clearly present. In general, it does seem difficult to properly define a notion of positivity and negativity for predicates so that it would be sufficient for using them as a foundation for analogies.

In the legal literature on analogy this worry has been taken rather seriously and a common way to solve it is to restrict the choice of predicate to not only point out a certain similarity between two cases, but to require that we have to

accept all cases where this property is present as similar and all cases where this property is absent as different. We can call this a correct sorting requirement, the condition of efficiency or the proportionality principle. If the property unites cases that should be held apart or distinguishes cases that should be united, we have according to this restriction reason to consider the proposed property as wrong for the present purpose. This requirement resembles to a great extent what we might call a consistency requirement of inductive reasoning, requiring us not to have any counterexamples to any good induction. Introducing such requirement seems to deal with the proposed challenge with identity and numbers. All cases will have the property of 'being non-identical with the number 2' and we therefore have to accept that all of them actually are similar. This will of course be an unacceptable result as we will be provided with a reason for following the legal result of all cases and end up with an inconsistency. At least for now, the proportionality requirement seems to have solved the issue of everything being similar to everything in an infinite numbers of ways.

Though even if this gets analogy out of the frying pan, we will see that we are dangerously close of throwing it into the fire. As often happens, when we seem to have solved one problem, another one arises. One common motivation for introducing a proportionality requirement is to restrict the choice of predicate so that there is no more an infinite amount of potential similarities, which could be used to establish a similarity with whatever case one would like. We will show here that it fails to do so. We will argue that even by introducing such restriction, we are able to, no matter the source case show that it is similar to target case in a way that will satisfy the proportionality principle. Not only do we seem able to find one such predicate, but there are seemingly, for whatever source case(s), an infinite amount of such predicates.

To demonstrate the existence of such predicates we will rely on the assumption that every legal case is unique. This assumption of uniqueness means that no case share all its features with some other case, namely that for every case, there is at least one feature or combination of features that is not shared with any other case. At first sight this assumption might seem objectionable. Though by a closer look, we will see that it should not be considered very controversial, but that rather reasonable and almost self-evident. Notice here that we are not restricting our claim to be about *legally relevant* features, as establishing legal relevancy is precisely the goal and object of study for this given investigation and such restriction would seem to be a clear instance of begging the question. The assumption is therefore that amongst all features of a case, there is at least one feature or combination of features that uniquely identify the case in question (we leave the question whether a case has a finite or infinite amount of features open; this argument holds in both situations). We will call such feature or combination of features a *uniquely distinguishing feature*. We will argue for the existence of such uniquely distinguishing features by first highlighting some particular categories of features that seem to play this role and then continue to show that this actually follows from Leibniz's (somewhat controversial) law regarding the identity of indiscernibles.

We have earlier highlighted that this question is concerned with not only legally relevant features, but with all features of a case. Typically, a legal case would be described by some certain facts that are considered to be important for the precise treatment of this case. The legal agent(s) would then investigate the legal facts of the case and apply some legal principles and rules and thereby

end up with a decision that establishes its legal treatment. A crucial part of this treatment, however, is to be able to justify the choice of precisely these facts and not all others as legally important. An event or action[2] that forms the foundation of a legal case can be described in a significantly larger number of ways than what would usually be included in a legal description, and some of these seem to be what we have called uniquely distinguishing features. Think for example of the precise time and place the event took place or of the precise individual(s) that performed the action. That an individual i performed a certain action of type A or that a certain event e took place at time t in place p seem to both be features that for most situations would be uniquely distinguishing. Though, what if i actually performed A a second time, or that there were several individuals involved so that there actually were several events taking place in p at time t, and thereby yielding distinct cases? A conjunction of the two seems to solve this problem, so we characterise the action as i performing A at time t in place p. One obstacle remain here though.

So far we have only been speaking about actions and events, but legally it seems clear that a single action described in this way might be ground multiple cases. The most obvious way this might happen is treatment by several court instances. Normally in a legal system either party have the possibility (under some conditions) to appeal a decision in a lower court to some higher court (for example a court of appeal, supreme court or some international court). We will then be in a situation where we have several distinct treatments of a particular event. Though it is unclear whether this actually is a problem for our argument. By appealing to some higher court, one appeals a decision of a particular case in some lower court, whose decision usually overrules any decision in the lower court. The case seems thereby to be the same throughout the appeal, it is only its *treatment* that is different. This point is not essential to the point made here, as we might simply consider it a further specification for the distinguishing feature, namely being treated by court c.

A further challenge might be that the same event where the same individual is involved might also be treated multiple times by the same court. An example of this might be if an individual i performed some illegal action towards individual j. Here, we might have one legal treatment for the sanctioning of i's performed action and a separate treatment for the monetary claim that j might have towards i. In both treatments it is the same event that grounds it. Our approach to this challenge will be to say that they indeed are different cases, but that there will be some systematic difference between them. First, a case does not contain only a particular event, but also any legal demands in relation to that event. In our example, in the first the demand was for the defendant to be sanctioned and in the second that the defendant compensates the plaintiff. One therefore demands a sanction and the other a compensation. Second, in many legal systems also the plaintiff seems to be a different agent in the two cases. For a claim of compensation, the plaintiff is typically the particular individual, here j, while in a criminal case of sanctions the plaintiff is not the individual, but the state or the people in a general sense on behalf of the particular individual. We have now only pointed out some important differences for the particular situation of criminal law, not shown something general about all areas of the

---

[2]We here consider 'action' to be a particular kind of event that was performed by some agent(s). We do not wish to take any standpoint on the question whether also events, and not only actions can found a basis for a legal case.

law. To do this, we need a stronger argument and we will provide this argument by showing that rejecting the proposal will amount to an, if not absurdity, a very problematic legal situation.

Say that a legal case s and a case s' are identical in respect to all requirements we have listed so far. Both are based on that an agent i performed a certain action of type A at time t in place p and both cases are to be treated by court c and the plaintiff has precisely the same demand towards i. It is certainly (conceptually) conceivable that both s and s' might be treated independently as distinct cases, though this can hardly be seen as anything else than a mistake. If the court should allow treating s and s' as independent cases, it must be precisely because there is a difference (or at least a suspicion of a difference) between the two cases.[3] If the court would admit of no difference between s and s' and still treat them separately, we would be in a situation where the same situation might receive multiple treatments in the same legal system and any departing decision between the two cases would show to be very problematic as will not know which one to follow. Such situation would clearly end up with inconsistencies.[4] We would also have to admit of having a legal system where any plaintiff or defendant could ask for a new treatment of the event whenever they are dissatisfied with the result. Based on this we can conclude that such characteristic is highly unwanted and that we might not accept treating identical cases multiple times.

In the given arguments, we have based our argument on some categories that seem important for distinguishing cases. These were agent, event, time and place, court and plaintiff/defendant. This is not to say that these should be considered an exhaustive list and we might have situations where these categories are not sufficient (anything else would be surprising) for distinguishing cases from each other. Our claim however, is just that for two cases to be distinct, there must be some difference between them. If no such difference can be found, we must consider the cases to actually be one and the same. This claim actually seems to be an instance of Leibniz' law regarding the identity of indiscernibles:

> qu'il n'est pas vrai que deux substances se ressemblent entièrement et soient différentes solo numero
>
> *that it is not true that two substances resemble each other entirely and being different only in number*

<div align="right">(Leibniz, Discours de métaphysique, 9)</div>

---

[3]Another plausible alternative is that the court finds that s actually was wrongly treated and thereby will treat s' again. This however does not seem to be a problem for us as the court, in making such decision also admits the two cases of actually being the same. See the argument regarding higher and lower courts. If for some reason this should not be convincing, we might also further specify our notion of court to not only be its level, but to describe the precise court assembly for this particular court decision.

[4]Also note here that we do not speak about multiple treatments in terms of appeals, where we contrary to this example have a hierarchy of decisions. Another particular situation might also be hybrid legal systems where we have multiple distinct legal systems co-existing, such as the legal system of Indonesia. However, here a great effort has been placed into analysing how these systems can co-exist, precisely by utilising analogies and parallel reasoning. See Iqbal (2022) for an analysis of the Indonesian hybrid model.

We now assume having established that no two cases can be treated distinct, without there being some feature that distinguishes them. This then means that for each case, we should be able to find a uniquely distinguishing feature or combination of features so that this feature holds only for one particular case. If for some case no such feature can be found, we have argued that it should be considered a mistake to consider it as a distinct case.

3   A challenge for analogy

After having established our ability to distinguish each case by a unique feature, we will proceed by showing how this actually ends up being a large problem to the common restrictions imposed on analogical reasoning. We started our discussion by claiming that if everything is similar to everything, analogical reasoning seems problematic. The common solution to this problem is to restrict what kind of similarities that should be considered relevant and this has been done by what we have called the proportionality principle, playing a comparable role for analogy as consistency plays for induction. This is often thought of as solving this problem of similarity, though by looking a bit closer we will now see that it actually just moves the problem a bit further ahead, but does not in any way actually solve it.

Imagine that we have a certain amount of cases. For simplicity let us assume that we speak about 4, named s_1, s_2, s_3 and s_4. Since we have shown earlier that every case has at least one uniquely distinguishing feature. For our cases, let us call them F_1, F_2, F_3 and F_4, so that F_1 is a uniquely distinguishing feature of s_1, F_2 for s_2 and so on. We are now able, no matter what source case, to produce a feature that unites it with whatever source case(s) we like and still satisfies the proportionality requirement. Say that we want to claim that s_1 and s_4 are similar to each other, but to no other. One feature that would provide this is simply the disjunction of F_1 and F_4. In this sense, two cases are relevantly similar if they both satisfy $F_1 \vee F_4$. Amongst our cases, this feature unites the cases that should be united and distinguishes the cases that should be distinguished. This in itself is of course no problem, as we simply made an analogy. The challenge is that no matter what cases we choose, we can find something that will unite them and distinguish them from all others. Say that we want to argue for the similarity between s_2 and s_4. We can then simply make a disjunction of the uniquely distinguishing features of the cases, namely $F_2 \vee F_4$. Do you prefer s_1, s_2 and s_3? Simply use $F_1 \vee F_2 \vee F_3$. As long as we accept that no case is exactly like another (which we should), analogical reasoning seems to be in trouble.

At first sight, one might think that this can be handled. After all, we have only shown that there is a possibility of finding one similarity between any cases that satisfies the proportionality requirement. Can we not simply count the amount of these similarities and require from our analogy that there should be one more than the possible combinations of the uniquely distinguishing features? This would seem to work on our introduced example, but by slightly modifying it, it does indeed seem to escape also this solution.

Think again on the initial example given regarding the problem of analogy, that all cases seem similar to each other by the predicate 'being non-identical to the number n'. Let us call this predicate N_n, so for any natural number n.

Even though this kind of feature was ruled out based on the proportionality requirement, it now seems to come back in full force. Say that our initial feature that united s_1 and s_4 was not simply $F_1 \vee F_4$, but $(F_1 \vee F_4) \wedge N_1$. The notion of similarity is then based upon having either the property F_1 or F_4 in conjunction with being non-identical to the number 1, N_1. Since the original proposal $F_1 \vee F_4$ satisfied proportionality, also the revised $(F_1 \vee F_4) \wedge N_1$ does (since N_1 holds for all cases). We are then in a situation where not only did we have one feature that satisfied proportionality, but two. In the same way as presented earlier, this argument is not restricted to the number 1. It also holds for the number $2, 3, \ldots$ so that we actually once more end up with not one uniting feature, but a potentially infinite amount of uniting features that does satisfy the proportionality requirement. And now, we are of course not able to count the number of similarities anymore. The proportionality requirement is not strong enough to solve this and the problem of an infinite number of similarities arise again.

Before trying to sketch some solutions to this problem, we will consider some potential counterarguments against the point we have given here. First, one might object regarding the way we have introduced the comparative predicate. It is not a single predicate, but a disjunction (and eventually a conjunction) of different predicates. Furthermore one might say that a proper notion of similarity must be based on a single predicate, not a combination of predicates and the previously given argument therefore does not hold. There are several problems with this argument. First, as demonstrated by Goodman (1983), what might in a language seem like two predicates is not necessarily two in any objective way. That we represent this particular feature as a disjunction or conjunction of several predicates is very dependent on the particular concepts and logically there would seem to be no reason that we should not be able to consider what we have now represented as multiple predicates as a single predicate. Such predicates are indeed also present in our natural language (here English). Think for example of the predicate 'being a parent'. Should we consider this a single predicate or as a disjunction of predicates, for example 'being a mother' or 'being a father'? When even our common concepts might have such structure, there should be no reason to think that our legal concepts should be any less complex. This leads us into a second way of rejecting this objection and this is simply that disjunctive and conjunctive properties are actually used in the legal literature and this even seems essential for any legal interpretative question.

A clear example of this can be found in German law, by their treatment of the notion of free speech. The example is then *everyone has the right freely to impart his opinion orally, in writing and in images* from the German Basic Law (the German constitution), Art. 5 para. 1 sent. 2. In order for this to be applicable to particular cases one would have to make a certain interpretation, particularly related to the notion of *opinion*. The expression is considered to capture the right of free speech. In order to apply this right to particular cases, the German Court has to settle the scope, contents and limits for this right. Here this happened stepwise, so that there was a gradually increasing degree of precision by including new disjuncts, defining the complex property. First *opinion* was interpreted as the combination of facts, personal views and ideas. It was then understood to include merely facts that were independent of opinion. Further interpretation provided an inclusion of facts that precondition the making of an

opinion. Afterwards it was also understood to include questions.(Langenbucher 1998, pp. 496-497) This shows that rejecting connectives to be used for a legal concept is hardly the way to go.

This leads us into a second potential counterexample. Maybe we cannot restrict the use of connectives as such, but we might restrict the kind of predicate we connect. Since the property of 'being non-identical to the number 1' not only holds for the particular cases, but it holds for all cases. Can we not then prevent the use of predicates that all or no case actually satisfies? Even with such restriction, we might still claim that everything is similar to everything, though not anymore in an infinite numbers of ways. However, we will show that even such restriction is problematic, in the same way as described for the previous counterargument. The first thing we want to point out is that also this argument is prone to objections in the style of grue and bleen. Furthermore, we seem to be able to describe predicates that would be able to avoid this problem by introducing negations of the identity or uniquely distinct features of other cases. Instead of 'being 'non-identical to the number 1' we can describe the conjunct as 'being non-identical to case c' or 'lacking the uniquely distinct feature of case c'. These predicates will be satisfied for all cases, except case c.

However, we might object again that this only provide us with a list of many predicates, but the list will not be infinitely long. As long as the set of cases is finite, there will also be a finite amount of such predicates. If we also assume that the amount of uniquely distinct features of all cases also is a finite amount, this point might be true. Though, it does not help us to solve the problem, as the length of the conjunction will be exactly the same for all cases. No matter what pair of cases we have, they will share the same number of similarities as any other pair of cases. This point has been named the *Ugly duckling theorem* as it shows that logically, a duck and a swan are just as similar to each other as two swans are. This was first demonstrated by Watanabe (1965) in "Une explication mathématique du classement d'objets"[5]. This should, which was also the goal for the demonstration of the theorem, be considered problematic, as we do want to say that some things are more similar than other things. One plausible way out of this (as suggested by Watanabe (1965)) is to introduce a weighting system, so that some parts should be given more weight than others. This is also the approach taken by Lewis (1968) in "Counterpart theory and quantified modal logic" for establishing the relation individuals have to their counterparts across possible worlds. However, as shown by Morreau (2010), such ranking approach is also problematic, unless we also assume a *dictorship*, that determines a critical feature for us to judge whether two things are similar. But this critical feature is precisely what we tried to establish in the first place! Unless we assume this critical feature to be externally imposed, we seem to do nothing more than begging the question.

A third way to object could be to simply refuse the use of identity statements. Since there should be equality for the law, one might claim (and reasonably so) that we should not be allowed to use identity statements, as this applies a legal rule to an individual simply for being the individual it is and that we thereby violates the legal equality. As the previous objection, this is an objection only towards the point of infinity, as we defined the property by a uniquely distinguishing feature, not by means of being identical to a particular case. This

---

[5]An English chapter describing the proof is found in Watanabe (1969)

however, can simply be avoided by slightly modifying our previous example. Instead of considering 'being non-identical to number n', we can base our argument on 'not being reducible to a number higher than n'. In this property, there is no mentioning of any identity and the objection does not arise.

There might also be a fourth way to reject this argument. And our answer to this fourth objection will actually lead us into a natural extension of the given argument. The objection goes as follows. Say that we are not limited to only actual cases, but also to variations of these case or purely hypothetical cases. If we admit of such cases we seem to not anymore only speak about a finite number of cases. By creating new hypothetical cases or further variations of our cases, we might have to admit that there are a potentially infinite amount of cases. How then is this a problem for our given argument? We assume that also for these hypothetical cases, we might find a uniquely distinguishing feature, as a lack of this would mean that they would be hypothetical in the first place (or distinct from another hypothetical). However, we used this uniquely distinguishing feature to create a long disjunction, and if our amount of cases is actually infinite, it seems problematic to create such disjunction. A disjunction cannot normally be infinitely long, and would therefore not necessarily provide a satisfiable explanation of the infinite amount of cases. However, the existence of such uniting property seems to be an instance of the axiom of choice, stating that even for an infinite number of sets, one might construct a new set containing exactly one element from each of these sets. This new set is then precisely the uniting property that we searched for. Even when admitting an infinite number of hypothetical cases, we seem inclined to accept that they might be united in one way or another. The problem therefore seem to persist, both in the finite and in the infinite situation.

4   Hierarchy and teleology

As promised, we will now sketch a possible solution to the given problem. The solution consists of essentially two parts. First, we will recognise that it seems impossible to speak about similarities and analogies without having some kind of purpose or teleological justification. Second, we will describe how to move towards a telos by describing a hierarchical structure of interpretations.

Watanabe's (1965) point with the ugly duckling theorem is to show that some bias is necessary for any claim of similarity. And stating the obvious, it is essential that the chosen bias is a good one. Our first task is therefore to sketch the outline of what such bias could look like.

Admitting such bias is not new in the literature on legal reasoning, and an important area of the contemporary discussion relates to precisely the identification of the content of this bias. In the light of Brewer's (1996) influential paper, an important debate regarding the status of analogical rules arises.

In 2005, Weinreb publishes a book, *Legal reason: The use of analogy in legal argument*, where he argues against the formulation of any rule, policy or purpose for deciding cases by analogy. The argument for this is that no rule or purpose can fully describe the sorting of similarities and dissimilarities in a satisfiable way. Furthermore, he points out that such rule seems absent in the practical performance of analogical reasoning. (Weinreb 2005, pp. 111-112) This point is not only directed to Brewer, but he also explicitly rejects the similar views by

for example Westen (1981) and Posner (2003). Despite his reluctancy towards rules and purposes, Weinreb (2005, pp. 1-13) establishes the importance of analogies very clearly. Contrary to the other mentioned theorists, he claims that analogical reasoning must be based on similarities, both to accommodate the way he claims we actually do use analogical reasoning and to avoid the mentioned problem with creating acceptable rules.

As a direct reaction, Posner (2006) published a book review (of the very same book), bringing up multiple problems with Weinreb's approach. The core of his objection is however based on the impossibility of even describing such similarities without in the same time providing some reason or general understanding. According to Posner similarities cannot be considered as relevant similarities unless there is some general understanding or reason for claiming their relevancy, and such understanding is in a legal case based on rules, principles, doctrines and policies. In addition, he points out that it seems to be Weinreb that does not understand the practical use of analogies, by not distinguishing between legal rhetoric and legal thought. Such reasons are, according to Posner, often not explicitly articulated, though this does not mean that they do not occur as justification in the substance of the law. So therefore, for an analogy to be guiding for any decision, there must be reasons to determine whether the similarities should be considered grounding for action. (Posner 2006, pp. 765, 768)

At first sight, one might consider the two authors to explore different aspects of legal arguments. From this point of view, Weinreb speaks about analogy from the perspective of arguments legal scholars *give*, while Posner speaks about arguments legal scholars *use*. It seems therefore to be no conflict, as they simply speak about different things, which furthermore seems to give Posner at least partially right when categorising Weinreb's points in the area of legal rhetoric. However, such clear distinction between the two seems potentially problematic. The arguments legal scholars give must clearly be somehow connected with the arguments they actually use. Legal rhetoric surely cannot be entirely independent of the substance of the law. Intuitively, there seems to be two ways of solving this tension. One might consider rhetoric and substance to be the very same or that the substance reflects the rhetoric (as seems to be Posner's view). Or, that rhetoric reflect the substance (the view of Weinreb). This point seems closely connected to the debate about whether we consider law as created or discovered.

What both approaches have in common, and what they also share with Brewer, is their search for *relevant similarity*. The disagreement can be reduced to a question of the precise content of this concept. All approaches describe a logical requirement that should be imposed on legal analogical reasoning. This can be called an *acceptable sorting* (for Brewer) or *consistency and coherence* (for Weinreb). The disagreement lies in their understanding of *relevance*. Posner (2006, p. 773) defends Brewer in stating that the relevance should be viewed as a rule or policy, while Weinreb (2005, p. 126) argues for understanding relevancy in a psychological or epistemological manner.

There is of course a possibility of also recognising other positions regarding the content of the notion of relevancy. The goal of this paper is not to take any standpoint in this discussion, but rather to highlight a part of the present literature on the content of such bias, recognised as a discussion on relevancy. Understood in this way, any judgment of similarity must be accompanied by a judgment of relevancy; *no similarity without relevancy*.

Even if we accept that there exists some constitution for explaining a necessary bias for similarity judgments and analogies, this only seem to partially explain the structure of analogical arguments. Some analogies might admittedly be sufficiently explained by such notion of relevancy, namely in cases where one has a clear psychological/epistemological or rule/policy justifying the given argument. This, however, does not seem to hold for most legal analogies of a certain complexity.

To illustrate this distinctions, we might use a simple example of the question whether a certain object in the target case is a stand-up transporter (Segway). Say that we had a limited amount of source cases that connects to the definition of a stand-up transporter, so that one potential interpretation that would satisfy the proportionality requirement in relation to the source cases would be that it is an object with precisely two motorised parallel wheels. Let us call this the MPW-interpretation. Compare this with an alternative interpretation that defines stand-up transporters in terms of maximal speed, maximal weight and its self-balancing steering mechanism, the SWS-interpretation. Since both satisfy the proportionality requirement, both interpretations would seem well-grounded. Can we reasonably speak about any psychological or epistemological state or a rule or policy that would decide between the two alternative interpretations? Psychologically, both seem to be plausible interpretations of stand-up transporters and we do are likely not to have any direct policy precise enough to govern an interpretation of such precision. As discussed earlier this point seems particularly problematic for any rational reconstruction of legal arguments, such as computer-assisted reasoning and other automatised learning, as we in these situations are required to make all underlying assumptions explicit. Developing a precise enough conception of rationality to decide such questions would at least seem practically and possibly also conceptually impossible.

However, what we might know from other definitions of vehicles is that the amount and position of its wheels should not be the defining characteristic for a category of vehicles. An e-bike might for example be an e-bike even if it has three rather than two wheels. Based on this knowledge, it is also indicated that the interpretation of a stand-up transporter, defined by the amount and position of wheels is incorrect or imprecise. There is therefore an analogy between the e-bike and the standup transporter, as both are small motorised vehicles. This is however an interpretation of an interpretation we thereby extend our notion of interpretation to a higher-order one.

This forces us to take a little detour and introduce a notion of order on interpretations and analogical arguments. We call an analogical interpretation that describes some relevantly similar properties between some source and some target for a *first-order* interpretation. This should be understood as an analogy directly over particular properties or characteristics for some source and target cases. An analogy that describes relevantly similar interpretations[6] (or an interpretation of interpretations) between some cases will be described as a *second-order* interpretation. An interpretation of interpretations of interpretations will be called a *third-order* interpretation and so on for any *n-order* interpretation.

By introducing a more complex notion of interpretation, we also rely on

---

[6] Here 'interpretation' is understood as particular application or result of an analogical argument.

a more complex notion of analogies, since we essentially describe analogies of analogies, or higher-order analogies. By the introduction of this more complex notions of interpretation and analogy we are able to more precisely describe how one might approach the psychological or policy-based goal described earlier. If we continue our line of interpretations regarding stand-up transporters, we see that it MPW-interpretation must be wrong, as the amount and position of wheels should not be the defining property of vehicles. The alternative SWS-interpretation would on the other hand seem to handle this context better. We can then describe how similar criteria would apply to the definition of e-bikes, as both are small motorised vehicles. This interpretation would again be correct as it pertains to an interpretation of safety. Contrary to the particular definition of certain types of vehicles, 'safety' is a notion that might be described in terms of a particular policy or psychology. It can be recognised as a general value in the society to have safe vehicles. We can thereby know that this interpretation is reasonable, at least in relation to some recognised value of safety.

We have in this section argued for reducing the question of interpretations to some rule, policy or psychological state (or bias in Watanabe's (1965) terms), but we claim that this in fact is only possible by relying on a certain 'hierarchy' of interpretations, where interpretations of higher-order plays an essential role. However, two challenges remain. First, how can we know that the solution we have found actually is the most correct, and that there are no better alternative interpretations? Second, how can we be certain that we actually will find a reasonable interpretation, one that is reducible to some bias? The answer to both of these questions is that we cannot know either.

The first problem seems actually to be a variant of an optimisation problem in mathematical analysis. We have preciously argued that even under the restriction of proportionality, there seems to be an infinite number of potential similarities and therefore also interpretations. By reformulating the problem we can see that what it really asks for is a way to ensure that precisely this interpretation is the most precise or correct, and that by none of the other (infinite number of) interpretations are better corresponding to our bias than the one we have found. In fact, this is the problem that the French lawyer and mathematician Pierre de Fermat tries to solve in his treatise *Methodus ad disquirendam maximam et minimam* (1679)[7] It is both a profound theoretical problem throughout the history of mathematics, while in the same time being a very practical challenge, particularly in modern computer science and machine learning. To illustrate the point, we will use an example of "finding the highest mountain peak".[8] Say that you wander around in the dark and your goal is to find the highest mountain peak in some area. How should you proceed? An obvious attempt is to search for a hill that goes upward and continue searching new hills, gradually continuing upwards until there are no more upward-going hills. You have now found the peak of the mountain (the local maxima). Though, you cannot know if you had chosen another path, if the other peak you might have found would be even higher than the first one. You can of course go down and try another path, considering whether it is the first or the second peak that is highest (global maxima) without going through every peak. This is the essential

---

[7]This treatise was published posthumously despite the authors own reluctancy, possibly a result of other contemporary mathematicians critique, whose most prominent character was René Descartes. Fermat's work can be found in Fermat (1894, pp. 133-179).

[8]The example is taken from Minsky (1982).

problem in developing search algorithms, and there are of course significantly more efficient methods than the one described here that might include things like planing, mapping, probability and so on. But there is no general method to always find the absolute highest peak. With Fermat's method we are able to find a local maxima on open sets or continuous function, namely to find a mountain peak.

As mentioned, we are also able to find the absolute highest peak within some area by visiting very peak. In the same way can we find the global maxima within a a particular range of a continuous function. The more general problem occurs if our goal is to find the global maxima over the *entire range* of the function. Here, the mountain area we investigate is not limited, but *infinitely* large.[9] Unfortunately, and as discussed earlier, interpretations seem to be of this last infinite kind. There is therefore no general way to ensure that the interpretation we have found will be the most precise one, as there will be an infinite number of alternative interpretations. And we have no way of knowing whether the interpretation we have found actually is our *global maxima*.

From answering the first question, we are also very close in answering the second. We have shown that it is not enough to arrive at a certain interpretation, but this interpretation must justifiable with respect to some psychological state or policy. Even if we can guarantee that we find some interpretation that satisfies the proportionality principle, we have no way of knowing that this might be justified. Say, for the sake of simplicity that there is one hierarchy of interpretations that will be justified by some psychological state or policy. Since this occurs in an infinite set of interpretations, there is no certainty that these particular interpretations will be found.

We have described this problem in terms of mathematics and search algorithms, but we should not think that it is limited to computational or theoretical descriptions of legal interpretation. The problem of correct interpretations is indeed also a practical one. This is most clearly acknowledged by most legal systems ability to depart from previous decisions. Decisions in civil law frameworks are typically not bound by the principle of stare decisis and this thereby opens up for deciding in a certain way in one case, even considering a similar case that has been decided otherwise (though there is usually also reasons to avoid to this to the extent possible). In the common law systems where the principle of stare decisis holds, we also have the possibility of overruling when an interpretation is fundamentally problematic and the more commonly used approach of distinguishing that introduces a nuance or slight change in a previous decision. We might here of course object and claim that these situations occur simply because the legal agents are human and that it is human to fail. Though as we have pointed out in this paper, the standard of rationality in case-based legal reasoning cannot be entirely neutral and unbiased. And assuming this is right, it does indeed seem difficult to make sense of human error as long as we do not have any neutral rationality to compare it too.

---

[9]To keep the example more plausible, the same problem occurs if we only allow the area to be too large for us to have the time and resources to investigate.

5 Conclusion

We have in this paper shown how the theoretical challenge of the ugly duckling regarding the notion of similarity is not only an abstract problem of computer classification, but also occurs as a practical challenge in relation to the interpretation of legal terms. The first part of the paper described the general problem of similarity and introduced a notion of uniquely distinguishing features. The second part applied this problem to the area of legal interpretation, arguing that even by imposing a requirement of proportionality, we still end up with a potentially infinite number of similarities and dissimilarities. The third part sketched how a solution to the given problem can look like. This solution can be used to describe how multiple challenging interpretations might be distinguished, but there seems to be no general method, describing such procedure.

We have here highlighted a general challenge for analogies and legal interpretation, that it cannot be without bias. Our suggested solution to this challenge actually consisted of two parts. The first part is the assumption of a bias. This has already been an important part of the contemporary discussion, but it does not seem to be sufficient in itself. The second part of the solution consisted in the development of a hierarchy of interpretations, to also include interpretations of interpretations. The combination of a hierarchy of interpretations and a certain bias, here understood as either a policy or a psychological state, seem therefore to be essential if we want to achieve an effective and precise description of the rationality behind a legal argument.

Based on the conclusion of this paper, any rational reconstruction of legal arguments should in addition to formal requirements incorporate some well-defined and acknowledged bias together with the possibility of developing interpretations of higher-order. A bias can be understood as a goal to achieve, possibly in terms of constitutional, political or commonly accepted values that justifies some interpretation of at least a higher-order. The inclusion of a hierarchy of interpretations would seem to indicate a higher-order language that is able to deal with interpretations of higher order. Though if we want to avoid the development of a gradual multiplicity of higher-order languages, such higher-order feature seems also describable the modern mathematical and computational theories dependent types.

References

Brewer, Scott (Mar. 1996). "Exemplary Reasoning: Semantics, Pragmatics, and the Rational Force of Legal Argument by Analogy". In: *Harvard Law Review* 109.5, p. 923.

Fermat, Pierre de (1894). *Oeuvres de Fermat*. Vol. 1. Gauthier-Villars.

Goodman, Nelson (1983). *Fact, fiction, and forecast*. Harvard University Press.

Iqbal, Muhammad (2022). *Arsyad al-Banjari's Insights on Parallel Reasoning and Dialectic in Law: The Development of Islamic Argumentation Theory in the 18th Century in Southeast Asia*. Vol. 25. Springer Nature.

Langenbucher, Katja (1998). "Argument by Analogy in European Law". In: *Cambridge Law Journal* 57.3, pp. 481–521.

Lewis, David K (1968). "Counterpart theory and quantified modal logic". In: *the Journal of Philosophy* 65.5, pp. 113–126.

Minsky, Marvin L (1982). "Why people think computers can't". In: *AI magazine* 3.4, pp. 3–3.

Morreau, Michael (2010). "It simply does not add up: trouble with overall similarity". In: *The Journal of Philosophy* 107.9, pp. 469–490.

Posner, Richard A (2003). *Law, pragmatism, and democracy*. Harvard University Press.

— (2006). "Reasoning by analogy (reviewing Lloyd L. Weinreb, Legal Reason: The Use of Analogy in Legal Argument (2005)". In: *91 Cornell Law Review 76*.

Watanabe, Satosi (1965). "Une explication mathématique du classement d'objets". In: *Information and prediction in science*. New York: Academic Press.

— (1969). *Knowing and Guessing a Quantitative Study of Inference and Information*. Wiley.

Weinreb, Lloyd L (2005). *Legal reason: The use of analogy in legal argument*. Cambridge University Press.

Westen, Peter (1981). "On Confusing Ideas: Reply". In: *Yale LJ* 91, p. 1153.