



From Case Law to Ratio Decidendi

Josef Valvoda^(✉) and Oliver Ray

Department of Computer Science, University of Bristol, Bristol BS8 1UB, UK
{jv16618, csxor}@bristol.ac.uk

Abstract. This paper is concerned with the task of automatically identifying legally binding principles, known as *ratio decidendi* or just *ratio*, from transcripts of court judgements, also called *case law* or just *cases*. After briefly reviewing the relevant definitions and previous work in the area, we present a novel system for automatically extracting ratio from cases using a combination of natural language processing and machine learning. Our approach is based on the hypothesis that the ratio of a given case can be reliably obtained by identifying statements of legal principles in paragraphs that are cited by subsequent cases. Our method differs from related recent work by extracting principles from the text of the *cited* paragraphs (in the given case) as opposed to the text of the *citing* paragraphs (in a subsequent case). We conduct our own independent small-scale annotation study which reveals that this seemingly subtle shift of focus substantially increases reliability of finding the ratio. Then, by building on previous work in the automatic detection of legal principles and cross citations, we present a fully automated system that successfully identifies the ratio (in our study) with an accuracy of 72%.

Keywords: Ratio decidendi · Case law · Natural language processing
Machine learning · Principle detection · Cross reference resolution

1 Introduction

In common law *ratio decidendi*, or just *ratio* [2], are the key principles used to decide the outcome of a legal case and the doctrine of *stare decisis*, or simply *precedent* [5], requires that ratio is applied on subsequent cases with similar facts to decide their outcome. This is different from the law used in civil law countries, where the legal principles are spelled out directly in statutes. Thus the identification of ratio is crucial to the work of lawyers and judges in common law countries such as the UK.

To find the ratio of a case of interest, lawyers must typically perform a detailed analysis of the text of the given case along with the text of key citing and cited cases and other cases related by topic [9]. Although legal search engines, such as Westlaw¹, are typically used to find the related cases relatively efficiently, there is currently a lack of technological support for the task of actually identifying the ratio contained within the transcripts of those cases. While

¹ Westlaw UK, Online legal research from Sweet & Maxwell, <http://westlaw.co.uk>.

there are human generated summaries of a case available, these are not focused on the ratio and provide only a simplified overview. Since a single case can consist of more than 150 paragraphs spread over 30 pages of text, the task of finding a ratio is an extremely time-consuming and error-prone process for humans to perform.

To automatically identify ratio, two issues must be addressed. The first issue is to distinguish statements of legal principles from discussion of specific facts to which those principles are applied in a particular case. The second issue is to determine which of those principles are pivotal to the outcome of the case (and therefore constitute the legally binding ratio), as opposed to those principles which are merely incidental and which are formally known as *obiter dicta*, or just *obiter* [2]. Automating these processes would be of great value to lawyers.

This paper presents a novel system for automatically extracting ratio from cases using a combination of natural language processing and machine learning to solve the two issues noted above. To achieve this we essentially build upon and integrate the recent work of Shulayeva et al. on the detection of principles and facts in case law [8] and the earlier work of Adedjouma et al. on cross reference resolution in statutory law [1].

The contribution of our work stems from several differences from Shulayeva et al. [8]. Whereas Shulayeva et al. specifically emphasise that they are *not* attempting to identify ratio, we specifically emphasise that we *are*. Whereas Shulayeva et al. are primarily concerned with the distinction between cited *facts* and *principles*, we are primarily concerned with the distinction between *orbiter* and *ratio*. Whereas the method of Shulayeva et al. is mainly concerned with analysing *citing* paragraphs, our method is mainly concerned with analysing *cited* paragraphs.

To illustrate these differences and demonstrate the effectiveness of our approach, we conducted our own independent small-scale annotation study which supports our hypothesis that the ratio of a given case can be reliably obtained by identifying the statements of legal principles in paragraphs that are cited by subsequent cases. In fact, our investigation shows that the principles in cited paragraphs correspond to ratio (as manually determined by a human expert using Wambaugh’s Inversion Test [2]) with a precision of 76%, whereas the principles extracted by Shulayeva et al. from the citing paragraphs correspond to ratio with an accuracy of only 68%.

In order to automate our new approach we do three things. First we build upon Shulayeva et al.’s methodology for principle identification in order to achieve a 96% accuracy on this individual task. Second, inspired by Adedjouma’s research on cross reference identification and resolution, we develop our own legal text schema to achieve an accuracy of 94% on the individual task of cited paragraph identification. Third, by combining the two classifiers above, we demonstrate a fully automatic system that successfully identifies ratio with 72% accuracy.

We conclude the paper by discussing how the bar we have set can be potentially raised in future work.

2 Background

In this section we explain our terminology and the distinction between ratio and obiter as discussed in the legal informatics literature [2]. We further describe Shulayeva et al.’s work concerned with automated fact and principle identification [8], Saravanan et al.’s work on case summarisation [7] and Adedjouma et al.’s work on cross reference resolution [1].

2.1 Citing vs. Cited: Cases, Paragraphs and Principles

When referring to paragraphs in legal cases connected by a citation we will use the terminology illustrated in Fig. 1, which is a natural extension of the nomenclature used in Westlaw. Suppose a paragraph (A) in a citing case (X) contains a citation to some paragraph (B) in cited case (Y). Then the former is called the *citing paragraph* (A) while the latter is called the *cited paragraph* (B).

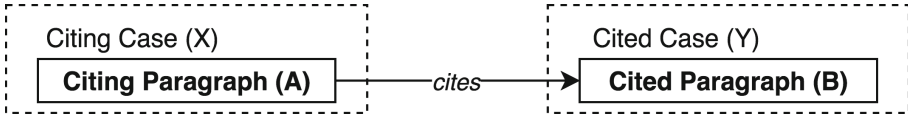


Fig. 1. The distinction between citing cases and paragraphs.

2.2 Defining Ratio via Wambaugh’s Inversion Test

Defining ratio decidendi and establishing a test for finding it in case law is essential for ratio identification. There are many different legal interpretations on what ratio actually is [5,6]. Branting’s research elegantly summarises the differing points of view on the matter [2]. He identifies two general areas of focus. One is on the material facts, the other on the deciding principles. These further translate into five different viewpoints [2]:

1. *The ratio decidendi of a precedent consists of propositions of law explicit or implicit in the opinion that are necessary to the decision.*
2. *A unique proposition of law necessary to a decision can seldom be determined. Instead a gradation of propositions ranging in abstraction from the specific facts of the case to abstract rules can satisfy this condition.*
3. *The ratio decidendi of a precedent must be grounded in the specific facts of the case.*
4. *The ratio decidendi of a precedent includes not only the precedent’s material facts and decisions but also the theory under which the material facts lead to the decision.*
5. *Subsequent decisions can limit, extend, overturn earlier precedents.*

In this paper we choose the first definition from above, since it can be tested for using Wambaugh’s Inversion Test [2]. Under this test, a principle in the case is inverted in its meaning; and if such inversion would affect the outcome of the case, then the principle is deemed a ratio. On the other hand, if the inversion would not affect the outcome, the principle is deemed an obiter. This test is widely used by legal practitioners and lends our research a solid grounding. Note that the outcome of the case is not only the immediate application of the law used in the case (i.e. which party won), but also the points of law established by reaching the conclusion (i.e. the legal precedent set by the case).

We now illustrate the Inversion Test by applying it to the landmark case of *Stack v Dowden* (commonly referred to as just *Stack*) concerned with division of family property after the breakup of a (unmarried) cohabiting couple. The question was whether the parties held the property as beneficial joint tenants, entitling them each to half of the property, or whether their conduct suggests they held the property otherwise, entitling the court to divide the property unequally. To decide on this the court needed to establish what should be considered to determine the parties intention. Thus, applying the Inversion Test on a sentence: *“The search is to ascertain the parties’ shared intentions, actual, inferred or imputed, with respect to the property in the light of their whole course of conduct in relation to it.”* identifies this sentence as a ratio. This is because Lady Hale employs the principle in this sentence, considering the parties conduct as a whole, to ultimately divide the shares in the property 65 to 35%, departing from the default legal position. Reversing the logic of this sentence would therefore disrupt the outcome of the case.

2.3 Automatic Case Summarisation

As part of their work on text summarisation Saravanan et al. consider the identification of ratio as one of seven different categories which they use to classify sentences in judgements: *“1. Identifying the case, 2. Establishing the facts of the case, 3. Arguing the case, 4. Arguments, 5. Ratio of the decision, 6. History of the case, 7. Final decision”* [7]. Upon first look, it might be tempting to compare our respective work on ratio identification.

However, Saravanan et al. are working with a very different case law to us. Where we focus on long English precedent setting cases from higher courts, they focus on shorter applicational cases from the Indian law. Perhaps because the cases they work with are much shorter, the example they provide only contains six paragraphs and a single judgement while our cases contain upwards of hundred and fifty paragraphs and five judges², they classify individual sentences whereas we classify paragraphs.

Because we are looking at such radically different case law to Saravanan et al., the search for the ratio is also a completely different task. This is illustrated

² Unfortunately, we could not access the data corpus used by Saravanan et al. since the links in their paper are out of service and we have not received a reply to our email asking for additional information. Thus, we could only inspect the examples in the paper itself.

by Saravanan et al. reliance on cue phrases to identify ratio. Phrases such as “*We are of the view*”, would not work in English law for ratio identification as judges do not use this or any other phrase to point out where they are dictating the ratio and where an obiter.

Therefore, Saravanan et al.’s work on ratio identification method does not capture the complexity we are dealing with in English Supreme Court cases.

2.4 Detection of Legal Principles

Shulayeva et al. address the task of automatically identifying (re)statements of facts and principles that are being cited from an earlier case. They refer to these as *cited* facts and principles (though it is important to note that they actually appear in the *citing* paragraph (A) in Fig. 1).

Their research demonstrates an agreement³ between two human annotators on annotating cited facts and principles of $\kappa = 0.60$. They have automated their task with 85% accuracy using supervised machine learning framework based on linguistic features. The features they use are: part of speech tags, unigrams, dependency pairs, length of sentence, position in the text and an indicator if/whether the sentence contains a full case citation. Their method is described below [8]:

1. Feature counts were normalised by term frequency and inverse document frequency.
2. Attribute selection (InfoGainAttributeEval in combination with Ranker (threshold = 0) search method) was performed over the entire dataset.
3. The Naive Bayes Multinomial classifier was used for the classification task.
4. Results are reported for tenfold cross-validation. The 2659 sentences in the dataset were randomly partitioned into 10 subsamples. In each fold one of the subsamples was used for testing after training on the remaining 9 subsamples. Results are reported over the 10 testing subsamples, which constitute the entire dataset.

Shulayeva et al. have applied their framework on their so-called Gold Standard corpus comprising of 2659 sentences selected from 50 common law reports that had been taken from the British and Irish Legal Institute (BAILII) website in RTF format. The corpus contained human annotated sentences labeled 60% as neutral, 30% as principles and 10% as facts. Their complete results are in Table 1 of the Appendix.

2.5 Cited Paragraph Detection

Adedjouna et al. focused on cross reference identification (CRI) and resolution (CRR) in legislature [1]. They demonstrate that by developing a legal text

³ κ is the predominant agreement measure that corrects raw agreement $P(A)$ for agreement by chance $P(E)$ [3, 8]:

$$\kappa = \frac{P(A) - P(E)}{1 - P(E)}.$$

schema, it is possible to define a structure capable of CRI, but also CRR, without the need for understanding the semantics of a sentence [1]. Reporting that using natural language processing techniques, a schema can be automatically applied on Luxembourg’s legislation to identify references with 99.7% precision and resolve them with 99.9% precision. They evaluate their method on 1223 references selected from Luxembourg’s income tax law.

3 From Case Law to Ratio Decidendi

According to the doctrine of stare decisis, cases are decided based on the precedent set out in the case law preceding them. Judges cite previous cases to apply the ratio from preceding cases on the facts before them [4]. Therefore, when judges cite specific paragraphs, it is reasonable to expect they are citing the paragraphs mostly for the ratio contained in them. To find the ratio we would thus like to identify cited paragraphs and identify the principles in them.

Shulayeva et al.’s method cannot be used to identify ratio directly because, as they themselves point out, it tends to confuse the ratio of the earlier case with that of the subsequent case: *“The main cause of error for the automatic annotation of principles was that the Gold Standard only annotated principles from cited cases, but often these were linguistically indistinguishable (in our machine learning approach) from discussions of principles by the current judge”* [8]. Perhaps for this reason, Shulayeva et al. explicitly say that *“the term ratio will be avoided”* [8].

Instead, we hypothesise it is better to try and identify ratio by identifying cited paragraphs (B) and extracting principles directly from them.

4 Testing Our Approach

To test if cited paragraphs correlate with ratio more than citing paragraphs, we conducted an empirical study on what we have called the Stack corpus. We created the Stack corpus by collecting and analysing cases from the Westlaw legal search engine, which contains a list of key citing cases for every major case in its database. As the case of interest, Stack v Dowden [2007] UKHL 17 was selected. This had 51 citing cases, of which Westlaw contained 41 in its database. Our Stack corpus is based on the 798 specific citations to Stack contained in these 41 citing cases.

We selected Stack for two reasons. First, it is both cited and cites frequently as a major Supreme Court judgement and therefore gives us plenty of data to work with. Second, because it contains a famous dissenting judgement of Lord Neuberger, containing principles reaching conclusion disagreeing with the rest of the judgement, a good example of an obiter. It is therefore a very challenging and rich case for the purposes of this task.

Our study essentially compares the occurrence of ratio in cited paragraphs containing a principle, with the occurrence of ratio in citing paragraphs containing a principle. We therefore manually identified the citing and cited paragraphs, paragraphs containing ratio and paragraphs with principles.

To determine whether a paragraph contains ratio or obiter, we manually applied the Inversion Test. Under the Inversion Test, the meaning of the principle is inverted and if this disrupts the logic for arriving at the outcome of the case, the principle is ratio, otherwise it is an obiter. Paragraphs without a principle were labeled as an obiter. Employing the Inversion Test we have found out that out of 158 paragraphs in Stack, only 34 contain ratio.

Manually analysing all the 41 citing cases in Stack corpus we found all the paragraphs in Stack which have been cited. Out of 798 citations in the cases, we identified 72 distinct cited paragraphs in Stack. Out of these 62 contain principles. Going through Stack paragraph by paragraph we identified 46 citing paragraphs, out of these 34 contain principles.

The confusion matrix in Table 3 of the Appendix shows the results for cited paragraphs. The data suggest correlation between cited paragraphs and the ratio. Applying this method achieves 76% accuracy in identifying paragraphs containing ratio and $\kappa = 0.45$. We have further tested the opposite assumption looking at citing paragraphs. As per Table 4 of the Appendix there is a drop in accuracy to 68% and most importantly the κ coefficient falls down to mere 0.06. These results support our hypothesis that cited paragraphs are a better indicator of ratio than citing paragraphs. Cited paragraphs are almost twice as precise and have more than three times the recall in identifying ratio than citing paragraphs.

Analysing the mislabeled paragraphs, we discovered two common reasons why judges cite paragraphs without the ratio.

1. The judge might refer to obiter of the cited case. There are numerous reasons why a judge might do this. For example he might highlight a shortcoming of current law:

“27 There is obvious force in the claimant’s contention that, as she and the defendant took out a mortgage in joint names for 43,000, for which they were jointly and separately liable, in respect of a property which they jointly owned, this should be treated in effect as representing equal contributions of 21,500 by each party to the acquisition of the property. It is right to mention that I pointed out in paras 118–119 in the Stack case that, although simple and clear, such a treatment of a mortgage liability might be questionable in terms of principle and authority.” (Lord Neuberger citing himself in *Laskar v Laskar* to suggest he would prefer a different (his own) interpretation of the law.)

Interestingly, a judge might also be using parts of the logic of the dissenting judge, in our case Lord Neuberger, to explain the reasoning of the majority of judges in the case:

*“125 While an intention may be inferred as well as express, it may not, at least in my opinion, be imputed. That appears to me to be consistent both with normal principles and with the majority view of this House in *Pettitt v Pettitt* [1970] AC 777, as accepted by all but Lord Reid in *Gissing v Gissing* [1971] AC 886, 897 h, 898 b - d, 900 e - g, 901 b - d, 904 e - f, and reiterated by the Court of Appeal in*

Grant v Edwards [1986] Ch 638, 651 f - 653 a. The distinction between inference and imputation may appear a fine one (and in Gissing v Gissing [1971] AC 886, 902 g - h, Lord Pearson, who, on a fair reading I think rejected imputation, seems to have equated it with inference), but it is important." (Lord Neuberger's obiter in *Stack v Dowden*, often cited as an explanation of the ratio.)

2. The judge might refer to the facts of the cited case. This happens when the judge is linking two cases on their facts to establish the applicability of the precedent:

"77 A great deal of work was done on the Purves Road property, some of it redecoration and repairs, some of it alterations and improvements. There is no doubt that the parties worked on this together, although there was a dispute as to exactly how much work each did and the judge found that Mr Stack probably did 'more than Ms Dowden gave him credit for' and eventually concluded that 'he had been responsible for making most of these improvements'. But he could not put a figure on their value to the sale price." (Facts cited in Paragraph 77 of *Stack v Dowden*.)

5 Automating Our Approach

To automate the methodology from Sect. 3 we had to resolve two problems. First, how to automatically identify principles and second, how to automatically identify cited paragraphs. Each is explored in a subsection below. Finally we combine the two solutions to automatically identify the ratio.

5.1 Principle Identification

Shulayeva et al. report two problems with their method of cited facts and principles annotation. First, cited principles and facts are often *"linguistically indistinguishable (in [their] machine learning approach) from discussions of principles by the current judge"* and second, sentences can contain both facts and principles at the same time, making it difficult to classify them as only one or the other. Under the definition of a ratio we employ we need only to identify the principles, as opposed to both facts and principles. Further, because we identify cited paragraphs via a separate method (see Sect. 5.2), we are interested in all principles, as opposed to only the restatements of facts and principles that are being cited from an earlier case.

We have therefore relabelled 96 sentences originally labelled neutral (because they are discussions of principles of a current judge) or fact (because they contain both fact and principle) as principles, 4% of the Gold Standard corpus, and created the New corpus. We implemented Shulayeva et al.'s framework in Weka according to their instructions and trained it on the New corpus. Below are some examples of the types of sentences we have re-annotated.

First is a sentence introducing a new principle, instead of restating a cited one, labelled as neither in Gold Standard corpus. We have re-labelled this sentence as a principle in the New corpus, because we are interested in identifying all the principles.

“He is not obliged to comply with any request that may be made to him by the borrower let alone by a surety if he judges it to be in his own interests not to do so.”

Second is a sentence classified in Gold Standard corpus as a fact. This sentence however contains both a fact and a principle, the latter is highlighted by us in bold. For our purposes such sentence still carries a principle and we therefore relabel it from a fact to a principle.

*“Thus for example if the manager explains either when making the request for payment to a third party or on it being questioned by the customer that the third party is a supplier and that the object is to obtain necessary materials for the work more quickly or that the third party is an associated company carrying on the same business **such an explanation might well bring the request for the payment to the third party within the usual authority of a person in his position and therefore within his apparent authority.**”*

The complete results are reported in Table 2 of the Appendix. They show a high accuracy for the task of principle identification, with over 10% increase in accuracy and a jump from $\kappa = 0.72$ to $\kappa = 0.90$ when applying Shulayeva’s model trained on our New corpus on the task of identifying principles compared to using the same model, trained on their Gold Standard corpus to identify restatements of facts and principles.

5.2 Cited Paragraphs Identification

There are noticeable similarities between cross references in legal statutes and paragraph citations in case law. For example in Fig. 2 Adedjouma’s cross reference identification (CRI) would identify “Art. 156.” using regular expression. In case law a paragraph would be cited as “para. 156”. For cross reference resolution (CRR), statutes seem to be comparable to case law, with implicit references such as “the next article” or explicit “the law of April 13th, 1995” which would be “above” or “in *Stack v Dowden*” in case law.

Art. 2. [...] Individuals are considered non-resident taxpayers if they do not reside in Luxembourg but have a local income as per the definition of Art. 156.

Fig. 2. CRI in Dutch legislation.

To identify cited paragraphs we have analysed 798 citations in Stack corpus. The 41 cases in the corpus are concerned with different issues, some citing to distinguish their problem, some to criticise it, and some to apply the law. They also come from a variety of courts and judges. Together, this gives us a representative sample of the variety of approaches judges and transcript writers could use to cite another case.

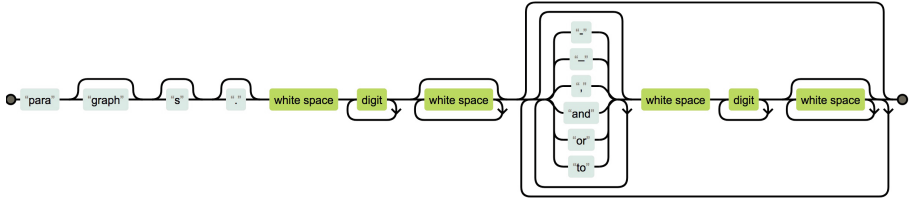


Fig. 3. Regular expression for the first type of paragraph citation.

Just like Adedjouma et al. we start by identifying all the patterns possible for recognizing the citation itself (CRI). There seem to be two general approaches. Under the first approach, the paragraph is either cited as *paragraph* or its abbreviation such as *para* or *paras*, see Fig. 3. The second approach, on the other hand, begins the citation with *at* followed by the paragraph numbers in square brackets, see Fig. 4.

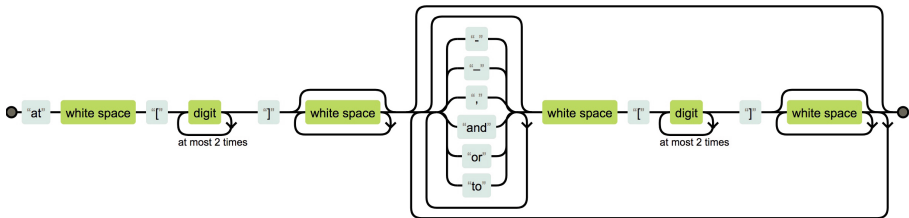


Fig. 4. Regular expression for the second type of paragraph citation.

The paragraph numbers themselves are expressed either as a single *number*, which correspond to Adedjouma's simple cross reference expressions (CRE), or as a *list of numbers* or a *range of numbers*, that correspond to Adedjouma's multivalued CRE's [1]. First we extract these citations using regular expression, then we interpret them to get all the implicit numbers in a range.

With the numbers of paragraphs cited identified the task of CRR remains. Because Adedjouma et al. are focused on statutes, they can simply extract the "Article 156", and assume it's the Article 156 of the same document the citation has been extracted from. Our task isn't as simple. Unlike Adedjouma et al., we

need to distinguish if the paragraph citations is pointing to the case we are interested in or not, since many cases are cited by a single case.

On top of that, we need to do this for case law which is by its very nature less structured than statutes. Consider the example sentences below. While we can relatively easily identify paragraph citation pointing to *Stack v Dowden* in example 1 and 2 below, where the name of the case or its abbreviation is included in the sentence, the same can't be said about examples 3 and 4.

1. *“Indeed, this would be rare in a domestic context, but might perhaps arise where domestic partners were also business partners: see Stack v Dowden, para 32.”*
2. *“First, as in the Stack case (see paras 90–92), the two parties in this case kept their financial affairs separate.”*
3. *“Fourthly, however, if the task is embarked upon, it is to ascertain the parties’ common intentions as to what their shares in the property would be, in the light of their whole course of conduct in relation to it: Lady Hale, at para 60.”*
4. *“In paragraph 42, Lady Hale rejected this approach in Jones v Kernott.”*

Carefully analysing the sentences citing *Stack v Dowden*, we have come up with a schema capable of identifying citations of our case of interest. Our schema takes an advantage of the knowledge of the names of the judges and the name of the case of interest, as both are written at the top of each case, and can be easily identified in the text. We also count the number of paragraphs in the case of interest to be able to reject any citation of a paragraph larger than this as impossible. The case abbreviation can be easily identified, since it is the first name in the case (e.g. for *Stack v Dowden*, it would be *Stack*). Employing these features, we constructed and implemented the Schema in Fig. 5 and applied it on the *Stack* corpus.

Under this schema we can resolve the examples above, including the problematic examples 3 and 4. Example 4 can be rejected as it contains a citation of another case (eg. *Jones v Kernott*, which we identify using regular expression). On the other hand example 3 would be classified as citing *Stack v Dowden*, because it contains the name of Baroness Hale, a judge in *Stack v Dowden*.

Comparing 3 and 4, one might notice there is no smoking gun evidence suggesting *Lady Hale* of example 3 is referencing *Lady Hale’s* judgement in *Stack v Dowden* instead of any other case she has judged, such as *Jones v Kernott* in example 4. However in our approach we work with the knowledge the case we analyse is citing the case of interest at some point. We know this since the cases we are analysing are selected from the list of cases citing *Stack v Dowden*, Westlaw provides. Therefore assuming that the case cited with a judge of the case of interest is indeed the case of interest, is a reasonable assumption to make. And as we report below, while a weakness of our approach, it still allows us to identify cited paragraphs with very high accuracy. Moreover, this is the best we can do without engaging with full analysis of the paragraph, or indeed the full case, which would be necessary to fully resolve this problem.

From 798 citations in the *Stack* corpus, the schema identifies 176 citations of *Stack*, 175 out of these are true positives. The single false positive is of a

sentence where the same Judge, who gave the judgement in the case of interest, is reported citing a paragraph but of a different case that is not mentioned in the same sentence. A schema analysing semantics of a sentence would be required to resolve this issue. There are also 8 citations that the program fails to recognise as citing Stack. These false negatives have not been extracted because the citing entity is contained in a different sentence or paragraph from the citing expression. Despite these shortcomings of our schema we achieve 98.7% precision, comparable to Adedjouma’s 99.9%.

However, since we are focused on identifying paragraphs citing Stack, it’s better to evaluate on how accurate the classifier is at identifying cited paragraphs. Out of 72 paragraphs we have manually identified as cited from 158 paragraphs in Stack, the classifier identifies 64 true positives 85 true negatives, giving it a decent accuracy of 94%. The full results are in Table 5 of the Appendix.

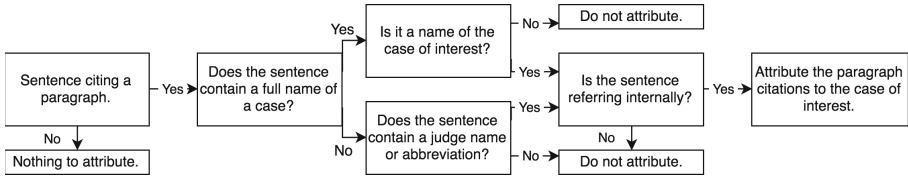


Fig. 5. Schema for attributing a citation to the case of interest from a present case (i.e. the case we are in).

5.3 Automatically Identifying Cited Principles

Finally, simply combining the principle and cited paragraph classifier described above, we evaluate how accurate our method is in identifying the ratio. As per Table 6 the new classifier identifies ratio with 72% accuracy on our Stack corpus.

Combining the principle and cited paragraph classifier therefore not only pin-points the position of the ratio in the paragraph, but also filters the cited paragraphs that do not contain principles, removing the instances where only facts are cited, further improving the performance.

Our automated approach therefore nearly matches it’s theoretical ceiling performance of 76% in the task of ratio identification (as established by our manual annotation study in Sect. 4), proving our hypothesis that focus on principles in cited paragraphs is a possible way of tackling the difficult task of automatic ratio identification.

6 Conclusion and Future Work

In this paper we have presented a novel approach to ratio identification. We demonstrate that identifying ratio can be automated by looking at cited paragraphs with 72% accuracy. We have also improved upon Shulayeva et al.’s work,

adapting their model for the sole identification of principles. Applying similar ideas to Adedjouma et al. we demonstrate that cross reference identification and resolution and cited paragraph identification can be achieved with almost equally high precision. Given the time-consuming nature of manually identifying ratio, our approach is a step forward in helping lawyers and judges spend their time on applying the law rather than looking for it. However, our work is only a preliminary study. A larger dataset, with more cases and human annotators, would be necessary for a full evaluation of the accuracy of our approach as well as the scope of it’s applicability. After all, only heavily cited cases can take benefit of our method. A very recent case, or a case without citations, will naturally be impossible to analyse. This is a limitation of our approach.

In our research, we have focused on all cited paragraphs without discrimination. However, not all citations are of equal importance. They are cited with different frequency, they are cited by cases from courts of different importance (Supreme court might cite differently than the Court of Appeal), the citing cases themselves might be reported immediately after the case comes out, as well as several years or even decades later and the citing case might approve as well as disapprove of a paragraph. All of the above could be used as features discriminating between cited paragraphs. Further, it is not only paragraphs that are cited, a sentence can be directly quoted and paragraphs might be cited individually or in a range. Discriminating between the precision with which text is referred to could again help distinguish between ratio and obiter. In the future, we would like to explore how the features above could further reduce the misclassification of obiter as ratio in our method.

Appendix

Table 1. Per category and aggregated statistics for the original Shulayeva et al.’s principle and fact classifier trained on Gold Standard corpus.

Classified as →	Principle	Fact	Neither
Principle	646	5	160
Fact	4	198	41
Neither	135	38	1432
Type	Precision	Recall	F-measure
Principle	0.823	0.797	0.810
Facts	0.822	0.815	0.818
Neither	0.877	0.892	0.884
Accuracy	0.85	κ	0.72

Table 2. Per category and aggregated statistics for Shulayeva et al.’s classifier trained on New corpus for extraction of principles only.

Classified as →	Principle	Neither	
Principle	837	70	
Neither	48	1769	
Type	Precision	Recall	F-measure
Principle	0.946	0.923	0.934
Neither	0.962	0.974	0.968
Accuracy	0.96	κ	0.90

Table 3. Distribution of ratio and obiter between cited and not cited paragraphs containing principle.

Contained in →	Cited	Not-cited	
Ratio	31	3	
Obiter	41	83	
Type	Precision	Recall	F-measure
Cited	0.468	0.853	0.604
Not-cited	0.948	0.734	0.827
Accuracy	0.76	κ	0.45

Table 4. Distribution of ratio and obiter between citing and not citing paragraphs containing principle.

Classified as →	Citing	Not-citing	
Ratio	9	25	
Obiter	25	99	
Type	Precision	Recall	F-measure
Citing	0.265	0.265	0.265
Not-citing	0.798	0.798	0.798
Accuracy	0.68	κ	0.06

Table 5. Per category and aggregated statistics for cited paragraph classifier.

Classified as →	Cited	Not-cited	
Cited	64	8	
Not-cited	1	85	
Type	Precision	Recall	F-measure
Principle	0.985	0.889	0.935
Neither	0.914	0.988	0.950
Accuracy	0.94	κ	0.88

Table 6. Per category and aggregated statistics for Ratio Decidendi classifier.

Classified as →	Ratio	Obiter	
Ratio	22	12	
Obiter	33	91	
Type	Precision	Recall	F-measure
Principle	0.400	0.647	0.494
Neither	0.884	0.734	0.802
Accuracy	0.72	κ	0.31

References

1. Adedjouma, M., Sabetzadeh, M., Briand, L.C.: Automated detection and resolution of legal cross references: approach and a study of Luxembourg’s legislation. In: 2014 IEEE 22nd International Requirements Engineering Conference (RE), pp. 63–72, August 2014
2. Branting, K.: Four challenges for a computational model of legal precedent. *THINK (J. Inst. Lang. Technol. Artif. Intell.)* **3**, 62–69 (1994)
3. Carletta, J.: Assessing agreement on classification tasks: the kappa statistic. *Comput. Linguist.* **22**(2), 249–254 (1996)
4. Elliott, C., Quinn, F.: *English Legal System*, 1st edn. Pearson Education, New York (2012)
5. Greenawalt, K.: Interpretation and judgment. *Yale J. Law* **9**(2), 415 (2013)
6. Raz, M.: Inside precedents: the ratio decidendi and the obiter dicta. *Common L. Rev.* **3**, 21 (2002)
7. Saravanan, M., Ravindran, B.: Identification of rhetorical roles for segmentation and summarization of a legal judgment. *Artif. Intell. Law* **18**(1), 45–76 (2010). <https://doi.org/10.1007/s10506-010-9087-7>
8. Shulayeva, O., Siddharthan, A., Wyner, A.: Recognizing cited facts and principles in legal judgements. *Artif. Intell. Law* **25**(1), 107–126 (2017). <https://doi.org/10.1007/s10506-017-9197-6>
9. Zhang, P., Koppaka, L.: Semantics-based legal citation network. In: Proceedings of the 11th International Conference on Artificial Intelligence and Law, ICAIL 2007, pp. 123–130. ACM, New York (2007). <http://doi.acm.org/10.1145/1276318.1276342>