The Split-up project: induction, context and knowledge discovery in law

JOHN ZELEZNIKOW†
Joseph Bell Centre for Forensic Statistics and Legal Reasoning, Faculty of Law, University of Edinburgh, Old College, South Bridge, Edinburgh EH8 9YL, UK

[Received on 28 March 2003; revised on 25 January 2004; accepted on 2 February 2004]

Most legal decision support systems have generally operated in domains with well-understood norms. Hence reasoning has been represented by a combination of rule-based and case-based reasoning. However, we analyse legal domains in which decision makers are allowed a significant amount of discretion. We argue that if the domain is bounded, and a sufficient number of commonplace cases exist, then the domain can be modelled using Knowledge Discovery from Databases techniques.

Whilst we focus upon legal principles for decision making in discretionary legal domains, our goal is to develop theory for constructing legal decision support systems. Our jurisprudential theory is hence applied to a practical legal domain—namely the distribution of marital property following divorce in Australia.

We conclude by discussing how we can maintain, update and evaluate the quality of the advice offered by our legal decision support systems.

Keywords: induction; knowledge; discovery; discretion; legal decision; support systems.

1. Introduction

The Split-Up project aims to examine how to model the exercise of discretion in legal decision-making. In doing so, the author and others have developed jurisprudential theories which suggest we may wish to apply knowledge discovery from database (KDD) processes to law.

According to Fayyad et al. (1996) knowledge discovery from databases is the ‘non trivial extraction of implicit, previously unknown and potentially useful information from data’. Knowledge discovery techniques have not been applied extensively in the legal domain despite potential benefits in the automated generation of legal knowledge from data. The absence of data in quantities collected in other fields such as astronomy accounts, in part, for this trend. However, for the most part, KDD has not been extensively performed with legal data because of a lack of clarity about how this can be achieved.

Theories of jurisprudence have proved indispensable for the analysis and development of computational models of legal reasoning. For example, the rule positivism of Hart (1961) underpins the application of logic programming in law exemplified by Sergot et al. (1986). The identification of jurisprudential theories that are particularly applicable to improve KDD in law, and how they can be applied, is the primary objective of this research project.

KDD techniques in general can be grouped into four categories:

† E-mail: john.zeleznikow@ed.ac.uk

© Oxford University Press 2004, all rights reserved
1. **Classification.** The aim of classification techniques is to group data into predefined categories. For example, data representing important case facts from many cases may be used to classify a new case into one of the pre-defined categories, ‘pro-plaintiff’ or ‘pro-defendant’.

2. **Clustering.** The aim of clustering techniques is to analyse the data in order to group the data into groups of similar data. For example, a clustering technique may group cases into six main clusters that an analyst would interpret in order to learn something about the cases.

3. **Series analysis.** The aim of series analysis is to discover sequences within the data. Sequences typically sought are time series. For example, past cases over a time period may be analysed in order to discover important changes in the way a core concept is interpreted by courts.

4. **Association.** The objective of association techniques is to discover ways in which data elements are associated with other data elements. For example, an association between the gender of litigants and the outcome of their cases may surprise analysts and stimulate hypotheses to explain the phenomenon.

Zeleznikow & Hunter (1994) in (IKBALIII) used rule induction to generate indices into cases. Rule induction was used by Rissland & Friedman (1995) to analyse a domain in order to detect a change in the way a legal concept is used by courts. Large numbers of cases were examined by Wilkins & Pillaiapakkammatt (1997), who used the ID3 algorithm in order to estimate the number of days that are likely to elapse between the arrest of an offender and the final disposition of the case.

A judge in Brazil (V. Feu Rosa Pedro) has initiated a programme for the resolution of traffic accident disputes (FeuRosa, 2000). His ‘Judges on Wheels’ programme involves the transportation of a judge, police officer, insurance assessor, mechanical and support staff to the scene of minor motor vehicle accidents. The team collects evidence, the mechanics assess the damage, and the judge makes a decision and drafts a judgement with the help of a program called the Electronic Judge before leaving the scene of the accident.

Although KDD with data from law is not prevalent, important examples of classification, clustering, series analysis and association have been performed. See Stranieri & Zeleznikow (2004) for further details.

In practice, a knowledge discovery from database process involves the incorporation of some domain expertise at each of the following KDD phases: data selection, pre-processing, transformation, mining and evaluation. According to argumentation theorists, domain expertise can conveniently be represented as arguments for or against assertions. Therefore, I surmised that argumentation may provide a convenient framework for the representation of domain expertise when performing KDDs.

Technical details regarding the Split-Up system have been discussed in detail in Stranieri (1998) and Stranieri et al. (1999). This paper extends legal theory initially explored in Zeleznikow et al. (1997) to justify the use of KDD to predict the outcome of decision-making in discretionary legal domains, and claims that provided the data that reflect the decision-making processes are collected, and a sufficient number of commonplace (as opposed to landmark) cases exist, then KDD techniques can be used to discover patterns of judicial decision-making in discretionary fields of law. We demonstrate the use of this legal theory in the development of the Split-Up system.

The next section elaborates on the theoretical foundations that underpin these points.

Currently, the Split-Up system is a prototype used only in the offices of Victoria Legal Aid. We had, in conjunction with LBC (a legal publisher), a venture to market Split-Up. This failed for
non-technical reasons. With Victoria Legal Aid, we are currently developing a web-based version of Split-Up, which we hope to commercialize.

2. Decision making in discretionary legal domains

2.1 Open textured legal domains

Berman & Hafner (1988) indicate that legal reasoning is essentially indeterminate because it is open-textured. Bench-Capon & Sergot (1988) view the indeterminacy in law as a specific consequence of the prevalence of open textured terms. They define an open textured term as one whose extension or use cannot be determined in advance of its application. Prakken (1997) collates and analyses the substantial artificial intelligence literature on open texture. He points out that situations that characterize law as open textured include reasoning that involves defeasible rules, vague terms or classification ambiguities. This analysis of open texture is central to our discussion because we argue that the existence of judicial discretion is a form of open texture that is distinct from the situations considered by Prakken (1997). Prakken (1997) notes that the following are difficulties arising from the open textured nature of law:

(a) Classification difficulties. Hart (1958) presents a local government ordinance that prohibits vehicles from entering a municipal park. He argues that there can be expected to be little disagreement that the statute applies to automobiles. However, there are number of situations for which the application of the statute is debatable. What of roller blades, for instance? Fuller (1958), in a response to Hart posed the situation of a military truck mounted in the park as a statute. Considerable open texture surrounds the use of the term ‘vehicle’ in this case even though there is no question that the truck is a vehicle.

(b) Defeasible rules. Another type of open texture arises from the defeasibility of legal concepts and rules. Any concept or rule, no matter how well defined, is always open to rebuke. Rarely do premises or consequents exist in law that are universally accepted. Whilst a Victorian statute definitively prohibits drink-driving, few courts would convict a person who was forced to drive drunk at gunpoint. In this case the rule is defeated in the context of exceptional circumstances.

(c) Vague terms. Legal tasks are often open textured because some terms or the connection between terms are vague. A judge finds the various interpretations of terms such as reasonable or sufficient stems from the vagueness of these terms and not from classification dilemmas or defeasibility requirements. Brkic (1985) labels this a gradation of totality of terms that he claims is one reason that deduction is an inappropriate inferencing procedure for many problems in law.

The existence of judicial discretion contributes to the open textured nature of law. Yet there are discretionary situations that cannot be described as instances of classification difficulties, defeasible rules or the presence of vague terms. Hence, we argue that the existence of discretion is a distinct form of open texture.

Consider a hypothetical panel of Family Court judges who agree on all the facts of a family law property dispute. Members of the panel can conceivably arrive at different percentages of the assets that ought to be awarded to the wife. The different outcomes may partly be due to the presence of vague terms that are interpreted differently by various judges. In part, the different outcomes
may be due to classification type anomalies. One judge classifies a lottery win as a contribution to the marriage whereas another does not. Different outcomes may even be the result of defeasible rules. One judge applies the principle of an asset-by-asset approach whereas another considers that principle irrelevant and adopts the global approach.

While these scenarios describe situations that are open textured, there is another situation, common in family law cases that are not captured by these instances of open texture. A panel can be imagined where vague terms are interpreted in much the same way by all judges. There are no classification anomalies and all judges have used the same principles. In this scenario, the outcomes may still be different because judges apply different weights to each relevant factor. No judge is wrong at law, because the statute clearly affords the decision-maker precisely this sort of discretion. Thus, an additional situation is apparent; one where the decision-maker is free to assign weights to relevant factors, or combine relevant factors in a manner of his own choosing. This discretion will certainly contribute to the open textured nature of law and to indeterminacy.

Flick (1979) defines discretionary domains as those in which a judicial decision-maker has the freedom to select one interpretation or outcome from a number of permissible options. This definition can be seen to apply to family law property proceedings in the following way.

Dworkin (1977) presents a systematic account of discretion by proposing two basic types of discretion, which he called strong and weak discretion. Weak discretion describes situations where a decision-maker must interpret standards in his own way whereas strong discretion characterizes those decisions where the decision-maker is not bound by any standards and is required to create his or her own standards. McCormick (1978) does not dispute this conceptualization but contends that Dworkin’s distinction between typologies is one of degree and not of type. The discretion apparent in Australian family law exemplifies the weak discretion Dworkin. The vast majority of decisions made by the Family Court of Australia does not introduce new standards, set new precedents nor invoke a new factor that has not previously been considered. Consequently, the majority of such decisions cannot be seen to involve strong discretion. Most cases are those that Zeleznikow et al. (1997) call commonplace cases.

We claim that there are levels of discretion depending on the domain. There are many domains in which the exercise of discretion cannot be explained by the application of rules and principles. We hold this view because there exist domains such as property division in Australian family law, in which two decision makers may be applying identical rules and principles to facts interpreted in the same way, yet both arrive at different, yet legally valid outcomes. Typically, the statute that underlies these domains presents a list of factors to be considered by the decision-maker, but does not indicate the relative weighting of each factor. Christie (1986) calls these types of statutes shopping list acts. Judges in such domains exercise discretion by assigning a relative importance to each factor. The principle statute governing Australian family law, The Family Law Act (1975), is an example of a shopping list act.

Australia has a federal system of government. The Australian Constitution divides authority between the States and the Commonwealth. s.51 of the Commonwealth of Australia Constitution Act 1900 (Cth) gives the Federal Parliament the power to make laws about:

(xx) Marriage; and

(xxi) Divorce and matrimonial causes; and in relation thereto, parental rights, and the custody and guardianship of infants.

Prior to 1959 there were varying state laws about divorce. The Matrimonial Causes Act 1959
(Cth) introduced the first uniform divorce laws for Australia. The principal aim of the Family Law Act (1975) was to reform the law governing the dissolution of marriage. The new Act replaced the Matrimonial Causes Act 1959 (Cth) and superseding State and Territory laws about ‘guardianship, custody, access and maintenance’ of children of a marriage.

The Family Law Act 1975 (Cth), as well as making significant changes to the law relating to divorce in Australia, created the Family Court of Australia to interpret and apply that law to individual cases:

S 21(1) A court, to be known as the Family Court of Australia is created by this Act.
(2) The Court is a superior court of record
(3) The Court consists of:
(a) A Chief Judge who shall be called the Chief Justice of the Court;
(b) a Deputy Chief Judge who shall be called the Deputy Chief Justice of the Court; and
(c) Judge Administrators, Senior Judges and other Judges, not exceeding, in total, such number as is prescribed.

Appeals from a first instance decision of a Family Court judge are ordinarily heard by a Full Court of the Family Court, which must be composed of at least three Family Court Judges (Dickey, 1990). There are currently 52 judges of the Family Court of Australia. For administrative reasons only, the Family Court of Australia is divided into four regions. Each region has a Judge Administrator. Each region has a number of registries. Whilst each of the judges of the Family Court of Australia is assigned to a registry, he/she regularly hears cases in other registries. An appeal from a decision of the Full Court of the Family Court can be made to the High Court of Australia.

Under this Act, the sole grounds for dissolution of marriage are an irretrievable breakdown of the marriage. Matrimonial fault was not deemed relevant as the basis for the distribution of property interests by the framers of the Family Law Act (1975). However, the notion of contributions was regarded as suitable. The principle that a party to a marriage should be rewarded for his or her past efforts introduces a retrospective element into the determination of suitable property orders. The retrospective element is counter-balanced with a prospective element. The prospective element has been called ‘future needs’. However, this terminology is not indicated in the statute. Section 79[4] and Section 75[2] of the Family Law Act (1975) list a number of factors that must be considered when devising a property order. Section 79[4] refers to the prospective element included in Section 75[2].

The statute presents a ‘shopping list’ of factors to be taken into account in arriving at a property order. The relative importance of each factor remains unspecified and many crucial terms are not defined. For example, the nature of a contribution is left unspecified. What weight the retrospective element assumes relative to the prospective element is similarly left unspecified. The age, state of health and financial resources of the litigants are explicitly mentioned in the statute as relevant factors, yet their relative weightings are also unspecified. The Act clearly allows the decision-maker a great deal of discretion in interpreting and weighing factors.

Christie (1986) describes different situations that involve discretion in order to claim that its exercise inevitably involves power relationships within a political system. His approach is particularly useful for us, not because of the socio-political conclusions he draws, but because he
specifically identifies statutes that provide a decision-maker with a shopping list of factors as fields of law that necessitate a kind of Dworkian weak discretion. His main example is reproduced in Stranieri & Zeleznikow (2004) to draw a parallel between the discretion that Australian family law mandates and the discretion given to decision-makers regarding US hazardous wastes. The relevant legislation is Section 520 of Second Restatement of Torts (1977).

2.2  *Levels of stare decisis*

According to Kovacs (1992) family law in Australia differs from other legal domains in that the principle of *stare decisis*, that like cases should be treated alike, is only superficially applied in family law. For example, the Full Bench of the High Court of Australia, determined in Mallet vs. Mallet (1984) 156 CLR 185, that trial judges cannot base their assessment of property matters by assuming a 50/50 split between husband and wife and deviating from this starting point on the basis of contributions and needs. Trial judges are encouraged by the High Court to take all factors indicated in the statute into account.

Kovacs (1992) contends that the High Court in Mallet vs. Mallet failed to take the opportunity to place specific constraints on the way in which trial judges determine property matters. High Court Justice Gibbs, C. J. made this point succinctly in *Mallet*. He said,

> It is proper, and indeed often necessary for the Family Court, in dealing with the circumstances of a particular case, to discuss the weight which it considers should be given, in that case, to one factor rather than another. It is understandable that practitioners, desirous of finding rules or even formulae, which may assist them in advising their clients as to the possible outcomes of litigation, should treat the remarks of the court in such cases as expressing ‘binding principles’. Likewise, Family Court judges who seek certainty or consistency look for guidelines. Decisions in particular cases of that kind can, however, do no more than provide a guide; they cannot put fetters on the discretionary power which the Parliament has left largely unfettered.

Ingleby (1993) largely concurs with the views of Kovacs (1992). He notes that an appeal to the first appellate Court, the Full Bench of the Family Court of Australia, is not permitted if the only ground for appeal is that the appeal court would have arrived at a different result had it heard the case. Permissible grounds for appeal include:

(a) the first instance judge did not include reasons for a discretionary decision,
(b) the trial judge acted on a wrong principle,
(c) the trial judge allowed irrelevant matters to guide him or her,
(d) the trial judge did not take relevant matters into account or did not afford them appropriate weight,
(e) the trial judge mistook the facts.

Stare decisis is a fundamental principle in common law legal systems. The principle dictates that the reasoning, loosely, *ratio decidendi*, used in new cases must follow the reasoning used by decision-makers in courts at the same or higher level in the hierarchy.

Thus, if fields of law such as property division in Australian family law are so discretionary that leading commentators convincingly argue that stare decisis does not apply, can case outcomes
be accurately predicted? If outcomes cannot be accurately predicted, any attempt at modelling techniques is futile.

Perhaps outcomes in discretionary fields cannot be predicted because the discretion that is placed in the hands of the judge encourages so much uncertainty that predictions can only be educated guesses. However, if this were the case, we would expect practitioners in Australian family law to be consistently inaccurate with their own predictions. On the contrary, practitioners are very accurate in predicting outcomes, despite the discretion available to judges. This apparent paradox is resolved by looking more closely at the concept of stare decisis.

Wassestrom (1961) identifies three types of stare decisis. Under traditional stare decisis, a court is bound by prior decisions of courts of equal or higher level. It is this kind of stare decisis that Kovacs (1992) and Ingleby (1993) claim has not occurred fully in family law, because the High Court has failed to lay down specific constraints for trial judges to follow.

Another type of stare decisis, called personal stare decisis, is used to describe the observation that most judges attempt to be consistent amongst themselves. This manifests itself in the Family Court, as the tendency that an individual judge has to be consistent with the way he or she exercised discretion in past, similar cases.

The third type of stare decisis, local stare decisis, represents the tendency of a group of judges that make up a current court to follow its own decisions. Local stare decisis manifests itself in property division in Australian family law, as a desire for judges of the Family Court to exercise discretion in a manner that is consistent with other judges of the same registry of the Court, at the same time. This may occur because the decision makers all share the same values.

Lawler (1964) claims that predicting the outcome of a case is impossible without the concept of stare decisis. Furthermore, the ability to predict an outcome with some accuracy is important if the law is to be both transparent and consistent.

Despite constant controversy about Australian family law property division, by and large, experienced practitioners can predict outcomes with some degree of accuracy. As Kovacs (1992) and Ingleby (1993) point out, this level of predictability is not due to traditional stare decisis. We take the view that the predictability must be the result of the remaining two forms of stare decisis, local and personal stare decisis.

This has ramifications for the data selection, data pre-processing and evaluation phases of KDD. Some case outcomes in discretionary domains are so far removed from other similar cases that it is reasonable to assume the judge has erred. In domains characterized by traditional stare decisis, a judge can err by failing to follow the constraints laid down by superior or equal Courts. In domains characterized by personal and local stare decisis, judges err by failing to be consistent with other judges currently in the same registry of the Court or with themselves.

Another ramification of local and personal stare decisis relates to the types of cases suitable for the data selection phase. Ingleby (1993) argues that the vast majority of cases that come before the Family Court are not extraordinary. They do not involve extraordinary facts, do not have outcomes that are unexpected and are, consequently, rarely reported by court reporting services. Zeleznikow et al. (1997) calls such cases commonplace cases and distinguishes them from landmark or leading cases. In fields where traditional stare decisis is emphasized, any case that is currently viewed as commonplace could be used in the future as a landmark case. This blurs the distinction between landmark and commonplace cases. However, in domains where traditional stare decisis is not strongly followed, if a case is regarded as commonplace at the time of decision, it is extremely
unlikely to be invoked in the future as a landmark case. An ordinary case impacts by adding to the body of cases for personal and local stare decisis.

Our conceptualization of traditional, local and personal stare decisis also has ramifications relating to the way in which we evaluate explanations generated by computer systems that use knowledge from a KDD process. In domains characterized with traditional stare decisis, reasons for a first instance decision often involve principles laid down by appellate courts. In the absence of traditional stare decisis, explanations cannot be rigidly derived from principles, because appellate courts have specifically failed to lay down such principles. Explanations must necessarily be further removed from the sequence of reasoning steps used to infer an outcome.

In building the Split-Up system, we do not use legal principles or rules to model the way judges actually combine factors to arrive at a decision. Rather, what judges have actually done in deciding real cases is assimilated by machine learning techniques so that a sub-symbolic representation of the exercise of discretion is established. The starting point for the process is an identification of factors that are currently, or have in the past been, relevant in the determination of a property outcome. Once relevant factors have been identified, data mining algorithms can learn to weight factors. Christie (1986) and Bayles (1990) analyse jurisprudential assumptions that must be made regarding the concept of discretion in order to adopt the approach used here.

3. The use of cases in legal decision making

3.1 The inadequacies of modelling law as a series of rules

As Berman & Hafner (1988) note, rule- and logic-based systems handle open texture poorly, and generally rely on the user to resolve the open-textured predicate without assistance. To understand reasoning in open-textured and discretionary domains, we investigate the use of cases. One of the real benefits of using cases, whether it is via inductive or analogical reasoning or KDD, is the ability of these techniques to assist in resolving open texture. Nonetheless, deduction is a powerful form of reasoning, particularly in areas which are based on legislation and which have little case law, customary law, implied doctrines, discretionary provisions and so on.

Case-based reasoning is the catch-all term for a number of techniques of representing and reasoning with prior experience to analyse or solve a new problem. It may include explanations of why previous experiences are or are not similar to the present problem and includes techniques of adapting past solutions to meet the requirements of the present problem. Case-based reasoning is very useful in resolving open texture. Zeleznikow et al. (1997) introduced the distinction between landmark and commonplace cases to help differentiate between those cases that are used in case-based reasoners and those that are used in a machine learning environment. Although KDD uses cases, it does not reason directly with cases.

3.2 Landmark and commonplace cases

Kolodner (1993) incorporates context in her definition of a case for case-based reasoning systems. She states that ‘a case is a contextualized piece of knowledge representing an experience that teaches a lesson fundamental to achieving the goals of the reasoner’. Zeleznikow et al. (1997) notes that even in non-contentious areas, Kolodner’s definition provides scope for considerable problems. They disagree with Kolodner that a case necessarily ‘... teaches a lesson fundamental to... the
reasoner’. Certainly some cases do fit this description. Most notably within law, those decisions from appellate courts which form the basis of later decision and provide guidance to lower courts do provide a fundamental lesson, or normative structure for subsequent reasoning. Pound (1908) considers such cases to be formal, binding and inviolate prescriptions for future decision-making, whilst McCormick (1978) sees them as beacons from which inferior or merely subsequent courts navigate their way through new fact situations. The common name for such cases is landmark cases.

However, most decisions in any jurisdiction are not landmark cases. Most decisions are commonplace, and deal with relatively minor matters such as vehicle accidents, small civil actions, petty crime, divorce and the like. These cases are rarely, if ever, reported upon by court reporting services, nor are they often made the subject of learned comment or analysis. More importantly, each case does not have the same consequences as the landmark cases.

Landmark cases are therefore of a fundamentally different character from commonplace cases. Landmark cases will individually have a profound effect on the subsequent disposition of all cases in that domain, whereas commonplace cases will only have a cumulative effect, and that effect will only be apparent over time.

Take, for example, the case of *Mabo v Queensland (No.2)*. Prior to *Mabo* the indigenous people of Australia, the Aborigines, had few, if any, proprietary rights to Australian land. Under British colonial rule, their laws were held to be inchoate and Australia itself was held to be *terra nullius*, ‘empty land’ at the time of white settlement. Hence, the only property laws applicable were those stemming from the introduction of white rule, laws which were less than generous in their grant of land to Aborigines. In *Mabo*, the High Court held that previous decisions holding that Australia was *terra nullius* at settlement, and decisions holding that Aborigines had no property laws affecting land, were simply wrong at law. Hence, the High Court said, Aborigines had sovereignty over parts of Australia under certain conditions. Whether one agrees with the High Court’s interpretive technique, it is indisputable that this is the landmark case in the area, and has formed the basis of future decisions in the area. Indeed, *Mabo*, like many other leading cases, was the spur for political action and we soon saw the introduction of the Federal *Native Title Act*. Thus, landmark cases have the dual effect of determining (to some degree) the interpretation of subsequent fact situations as well as influencing the invocation of normative legislative processes.

Two leading United States landmark cases *Plessy v. Ferguson* and *Brown v. Board of Education of Topeka*, deal with the issue of segregated schools. Section 1 of the Fourteenth Amendment to the US constitution states,

> All persons born or naturalized in the United States, and subject to the jurisdiction thereof, are citizens of the United States and of the state wherein they reside. No state shall make or enforce any law which shall abridge the privileges or immunities of citizens of the United States; nor shall any state deprive any person of life, liberty, or property, without due process of law; nor deny to any person within its jurisdiction the equal protection of the laws.

In 1896, in the case of *Plessy v. Ferguson* the United States Supreme Court ruled that the demands of the Fourteenth Amendment were satisfied if the states provided separate but equal
facilities and the fact of segregation alone did not make facilities automatically unequal. In 1954, in *Brown v. Board of Education of Topeka* the Supreme Court seemingly overturned the decision made in *Plessy v. Ferguson*. In *Brown v. Board of Education of Topeka*, in the opinion of Black (1990), the Supreme Court declared racial segregation in public schools to be in violation of the equal protection clause of the Fourteenth Amendment. They did so, not by overturning *Plessy v. Ferguson* but by using sociological evidence to show that racially segregated schools could not be equal (Dworkin, 1986).

To further indicate the similarity between landmark cases and rules we note that in *Miranda v Arizona* the United States Supreme Court ruled that prior to any custodial interrogation the accused must be warned:

1. That he has a right to remain silent;
2. That any statement he does make may be used in evidence against him;
3. That he has the right to the presence of an attorney;
4. That if he cannot afford an attorney, one will be appointed for him prior to any questioning if he so desires.

Unless and until these warnings or a waiver of these rights are demonstrated at the trial, no evidence obtained in the interrogation may be used against the accused. *Miranda v Arizona* is a landmark case with regards to the rights of the accused in a United States criminal trial. This case has assumed such significance that its findings are known as the Miranda rule.

Landmark cases rarely occur in common practice and are reported and discussed widely. These cases set a precedent that alters the way in which subsequent cases are decided. In the last two decades, the number of landmark cases in the Family Court of Australia is in the order of hundreds while the number of commonplace cases is in the order of multiple tens of thousands.

It should be noted that the notion of a landmark or commonplace case is relevant to the system in which it is used. HYPO (Ashley, 1991) has a case knowledge base consisting of 30 legal cases in the domain of trade secrets. None of these cases have a particular legal significance, and thus legal scholars may consider these cases as commonplace cases. However, because HYPO reasons from these cases, for the purpose of the HYPO system, these cases can be considered as landmark cases.

Some critics believe the use of legal case-based reasoners is limited. Berman (1991) believed legal case-based systems must by necessity simulate rule-based systems and that factors emulate rules. He stated: ‘For developers, as contrasted to researchers, the issue is not whether the resulting rule base is complete or even accurate or self-modifying—but whether the rule base is sufficiently accurate to be useful’. We believe that jurispruders and developers of legal decision support systems use landmark cases as norms or rules. Commonplace cases can be used to learn how judges exercise discretion.

Given that we have a domain with an abundance of commonplace cases, how can we understand the manner in which judges exercise discretion? In Section 4 we investigate this question by examining how we have used induction and knowledge discovery from databases to model decision-making in discretionary legal domains. Our examples come from the domain of property distribution in Australian Family Law.

---

5 384 U.S. 436 (1966)
4. Learning from cases—the use of induction and knowledge discovery from databases in law

4.1 The Split-Up system

In the Split-Up project (Stranieri et al., 1999) we wished to model how Australian Family Court judges exercise discretion in distributing marital property following divorce. Section 79(1) of the Family Law Act (1975) empowers judges of the Family Court to make orders altering the property interests of parties to the marriage but does not lay down procedural guidelines for judicial decision makers. In practice, judges of the Family Court follow a five-step process in order to arrive at a property order:

1. Ascertain the property of the parties.
2. Value all property of both parties.
3. Determine which assets will be paramount in property considerations. This is referred to as common pool property distribution.
4. Determine a percentage of the property to be awarded to each party.
5. Create an order altering property interest to realize the percentage.

The Split-Up system implements steps 3 and 4 above, the common pool determination and the prediction of a percentage split. According to domain experts, the common pool determination task (Step 3) does not greatly involve the exercise of discretion, in stark contrast to the percentage split task (Step 4). Consequently, Split-Up implements the common pool determination by eliciting heuristics as directed graphs from domain experts using a methodology we have called sequenced transition networks. A detailed description of how to model legal decision-making using sequenced transition networks, and the use of sequenced transition networks in building web-based decision support systems, can be found in Stranieri & Zeleznikow (2001).

Domain expertise in family law is represented in the Split-Up system as arguments. This enables an informed data transformation phase and also constrains the data mining. For the philosopher Toulmin (1958), practical reasoning, as distinct from analytical reasoning, involves the construction of an argument. Arguments, regardless of the domain, have a structure, which consists of six basic invariants: claim, data, modality, rebuttal, warrant and backing. Every argument makes an assertion based on some data. The assertion of an argument stands as the claim of that argument. A warrant justifies why the claim follows from the data. The backing supports the warrant and in a legal argument is typically a reference to a statute or a precedent case. The rebuttal component specifies an exception or condition that obviates the claim. The Toulmin argument structure has been used by a number of researchers in various fields to model reasoning. However, a survey by Stranieri et al. (2001) illustrates that the majority of researchers vary the structure to suit their particular use. The variation that we used aimed to facilitate KDD. The structure is illustrated in Fig. 1.

Our variation on Toulmin’s theory of argumentation does not include either modality or rebuttal. Whilst these are important components of legal disputation, it was not felt worthwhile to include these invariants, given the programming difficulties involved in representing them.

The problem of deciding whether a legal task can be modelled by any existing paradigm, and which one is a problem currently tackled in an ad hoc manner by developers of legal reasoning systems. In an attempt to instill some method within our decision-making, we developed a simple classification scheme to classify sub-tasks in Split-Up (Stranieri et al., 1999). The classification scheme is based on two dimensions. These are our estimation of the extent to which a task is open-textured, and our estimation of the extent to which a task displays a feature that we call
boundedness. The scheme illustrated in Fig. 2 has two dimensions: open-texture–well-defined and bounded–unbounded.

The open-textured–well-defined axis reflects the extent to which experts believed factors known to be relevant in a prediction were open textured. Predicting a percentage split of marital assets was considered open-textured, because of the high degree of discretion given to judges. The bounded dimension refers to experts’ beliefs about the completeness of their knowledge of relevant factors. In Split-Up, 94 variables were identified as relevant for predicting a percentage split of assets. Experts were of the view that few factors useful for a prediction were omitted from this list and therefore considered the task to be quite bounded.

![Fig. 2. Classification of percentage split task.](image-url)
Tasks that fall in the narrow-bounded quadrant are well suited to implementation with heuristics elicited as rules because all terms are well defined and all variables relevant for the prediction are known. Discretionary tasks that fall in the wide-bounded quadrant (top left in Fig. 2) can be modelled using the KDD process.

Unbounded tasks, whether or not they contain open-textured terms, cannot be modelled adequately using KDD, since sufficient relevant factors cannot be determined. Zeleznikow (2000) describes such domains as unfettered discretionary domains. Such domains have no norms and judges are not even told what factors must be taken into account in reaching a decision. We do not believe it is wise to model such domains.

An example of unfettered discretion is the determination of the custody of children in Australian family law. According to the Family Law Act (1975) the only factor to be taken into account is the paramount interest of the child. Following considerable litigation and uncertainty the Australian Federal Parliament made minimal attempts to define what are the paramount interests of a child. They did this by including in the legislation factors such as education, health, the child’s relationship with both parents, and the need to keep siblings together. But there is no clear list of factors. Indeed it is much easier to describe what is not in a child’s best interests (for example sexual abuse, violence) than what is in a child’s best interests. The granting of refugee status can also be considered to be an example of unfettered discretion and is an unbounded domain.

The Family Law Act (1975) directs a decision maker to take into account the past contributions of each party to a failed marriage in addition to their resources for coping with life into the future. Rather than offering one definition for contributions and one for needs, the statute presents a ‘shopping list’ of factors to be taken into account in arriving at a property order. For example, the age, state of health and financial resources of each partner are explicitly mentioned in the statute as relevant factors, yet their relative levels of importance are unspecified.

Although the statute presents a flat list of relevant factors without specifying how these factors relate to each other, we realized that the factors could be placed in a hierarchy. The development
of the hierarchy required specific knowledge supplied by domain experts. A sophisticated hierarchy of 94 factors presented in Fig. 4 was elicited. Figure 4 demonstrates that the factors relevant for a percentage split determination (extreme right of figure) are past contributions of a husband relative to those of the wife, the husband’s future needs relative to those of the wife, and the wealth of the
marriage. The factors relevant for a determination of past contributions are the relative direct and indirect contributions of both parties, the length of the marriage and the relative contributions of both parties to the homemaking role. No attempt is made in Fig. 4 to represent the way in which relevant factors combine to infer factors higher in the hierarchy. The hierarchy of Fig. 4 provides a structure that was used to decompose the task of predicting an outcome into 35 sub-tasks. Outputs of sub-tasks further down the hierarchy are used as inputs into sub-tasks higher in the hierarchy. Solid arcs in Fig. 4 represent inferences performed with the use of rule sets whereas dashed arcs depict inferences performed using neural networks (or indeed any other KDD technique).

Cases that set a precedent and change future decision-making (landmark cases) were discerned from commonplace cases. This distinction helps us to select cases that are most appropriately used to discover patterns of discretion in typical cases and not those that result in a change in law.

To collect data for the Split-Up system, we read family court judgements. Values for relevant factors were extracted from each case. Many examples were contradictory. Contradictory examples are those that display different outcomes given the same or very similar inputs. In non-legal domains these are often attributed to noise as erroneously recorded data. However, in discretionary domains of law, we expect some contradictions because individual judges have some latitude to weigh the relevant factors in their own way. There are a number of different ways to deal with contradictions. Most simply, the contradictions can be ignored. Wang & Gedeon (1995) note that a small proportion of noisy examples will not dramatically affect the performance of a neural network. In law, outcomes that contradict others may reflect judicial error and warrant removal from the database. Although this is subjective, we adopt the strategy of removing cases that domain experts consider erroneous. A metric to gauge the extent of similarity in inputs of multiple examples with the same outputs was developed in order to facilitate this.

According to data mining rules of thumb, the number of examples needed to identify useful patterns from 94 variables is in the many tens of thousands. Data from this number of cases are rarely available in the legal domain. Furthermore, few cases involve all 94 variables (e.g. childless marriages have no values for all variables associated with children) so a training set would be replete with missing values. We used the Toulmin structure depicted above to decompose the task into smaller tasks each of which involved a sufficiently small number of variables in order to facilitate KDD with the small number (103) of examples we had. Furthermore, the structure enabled the collation of training sets with no missing values. Figure 3 illustrates the claim and data item of three arguments. The claim of argument B was one of the data items for argument A. In total, the 94 variables were dispersed in 35 arguments. Twenty of these were classified ‘wide bounded’ so training sets were assembled for KDD. Heuristics for the remaining 15 (classified narrow bounded) were sourced from experts for rule sets.

The classification scheme has been used to classify tasks in the domain of family law Zeleznikow & Stranieri (1995) (35 arguments), refugee law (Yearwood & Stranieri, 1999) (200 arguments), copyright law (Stranieri & Zeleznikow, 2000) (50 arguments), eligibility for legal aid (Hall & Zeleznikow, 2002) (8 arguments) and the evaluation of eye-witness evidence (Bromby & Hall, 2002).

### 4.2 Evaluating the Split-Up system

One way to ensure consistency in a legal decision support system is to conduct a thorough evaluation of the system. Split-Up (Stranieri, 1998) has been evaluated in five distinct ways:
1. Domain expert assessment of both the content and structure of the Split-Up knowledge base and the problem solving strategy employed in Split-Up. The factor tree and argument structure used in the percentage split task were viewed positively by both domain experts associated with the project and four independent family law practitioners.

2. Comparison of predictions made by Split-Up with those made by eight lawyers on the facts from the same three cases (Zeleznikow & Stranieri, 1997). In two of the three cases all eight lawyers agreed with each other (deviations of 5% either way from the Split-Up determination were deemed acceptable) and with the system. The third case presented significant controversy. Split-Up awarded the husband 55% of the assets. The lawyer’s predictions varied from 20 to 60%. The four lawyers that produced outcomes that varied with the other lawyers and Split-Up assumed that the wife had contributed significantly more than the husband to the homemaker-role. The case facts indicated that hired helpers performed the household duties and child rearing. The developers of Split-Up and some lawyers interpreted these facts as leading to an equal contribution to the homemaker role. On the other hand, four others lawyers assumed the wife made the major contribution because the husband was fully occupied with his medical practice and was therefore unlikely to have the time to supervise household staff. This illustrates an important problem with the use of legal decision support systems—users need to interpret data. Many disputes are about interpreting data (or facts): for such problems, human input is vital.

3. The use of Split-Up on a new trial case recently concluded in the Family Court of Australia, namely Opie v Opie. The case is an unreported 1996 case tried by Justice Brown in the Melbourne registry of the Family Court of Australia (the cases used in the Split-Up system were taken from the Melbourne registry of the Family Court of Australia in the period 1992 through 1994). The marriage lasted 17 years and resulted in two children—of ages 14 and 16 at the time of the trial. The husband ran a business in the automotive industry, which rarely returned large, profits and no longer exists. The wife primarily worked as the homemaker but often worked part-time in the business. The common pool system determined that the total assets for consideration were $108,800. Both are in the mid-40s and of good health. The wife is to have custody of the children.

Split-Up determines the percentage split in terms of needs, contributions and the level of wealth of the marriage. For the case of Opie v. Opie Split-Up determined

(a) the marriage is considered to be less than average in wealth;
(b) overall the husband has contributed the same as the wife during the course of the marriage;
(c) in the future the husband’s needs are less than those of the wife.

From these three determinations, through the use of a neural network, Split-Up determined Mr Opie should receive 35% of the common pool. In her decision, Justice Brown granted Mr Opie 34.7% of the common pool.

(a) was inferred through the use of a rule-based system given the value of the Common Pool. Domain experts claim the wealth of a marriage is important as future needs are significant for impoverished marriages but far less important for wealthy marriages, where each partner’s needs will be met save for exceptional cases. With regard to contributions, as in (b), Split-Up suggested that the husband and wife contributed equally to the marriage.
Justice Brown said that given the length of the marriage, the parties should be taken to have contributed equally.

With regards to (c), Split-Up suggested that the wife had greater future needs than the husband. The system came to this conclusion because it inferred that the wife’s prospects for the future are not so fair—as she has poor future employment prospects and few resources. The husband, on the other hand, has fair future prospects, because he has good work prospects and some resources for the future. Justice Brown thought likewise.

4. Current research involves feedback from users in four different categories using Split-Up predictions and explanations. Our research is based on the work of Buchanan et al. (1995), which claim that empirical validation with the use of a properly constructed questionnaire is a very useful quantitative indicator of user acceptance. We have used seven lawyers, four registrars, three judges and five lay people to evaluate the system using the quantitative assessment evaluation framework of Reich (1995).

Split-Up is currently being examined by judges, registrars, mediators and lawyers. When first proposed, it was expected that judges and lawyers would primarily use the system. Our subsequent research has shown our initial expectations as to who would be the main beneficiaries of the Split-Up system to be inaccurate.

- **How mediators use Split-Up.** Mediators in family law input both parties’ facts, peruse the resultant prediction and then explore the hierarchy of relevant data, warrant and backing factors with the parties in order to inform and educate them. Points of convergence between the two parties become obvious and the scale and loci of compromise are more easily identified.

- **How lawyers use Split-Up.** A lawyer uses the system a number of times with each client to explore hypothetical scenarios. A typical question is, what difference in outcome is there if I argue that my client performed an equal share of the home maker duties as opposed to arguing that she did most of those duties? A consultation with the system offers a prediction in both scenarios and assists a lawyer in determining which argument to proceed with. Lawyers are less interested in exploring warrants and backings unless these relate precedents that will be used to substantiate an argument chosen.

- **How judges distribute marital property and might use Split-Up.** Judges are required to arrive at an equitable outcome in the shortest amount of time possible. They have no need to educate litigants nor do they particularly need (or want) to evaluate their own judgements. However, they need to reach interim conclusions leading to a final judgement. They often need to interrupt a case for hours or days and then succinctly and quickly remind themselves of the facts and their own interim conclusions. Hence, the only benefit judges will reap from Split-Up will be as a useful tool for structuring their decisions.

- **How divorcees use Split-Up.** Divorcees with little knowledge of family law have often been surprised at predictions provided by the system. They tend to explore all warrants and backings in order to understand the prediction. Ultimately, it is not wise for systems such as Split-Up to be used by those with little family law knowledge, since such users cannot identify unusual (or hard) cases. The distinction between easy and hard cases may be jurisprudentially questionable, in that a case that seems perfectly commonplace today may be subsequently used to fundamentally alter a legal principle (hence becoming
a landmark case). However, in practice, the Family Court, on a daily basis, uses the distinction between commonplace and landmark cases, in order to decide which cases are to be published by Court reporting services.

5. Comparing Split-Up outputs with five written judgements of the Family Court of Australia. These cases were heard in 1995 and 1996 (the cases used in both the Split-Up training and test sets were decided in the three years between 1992 and 1994). This comparison showed that Split-Up inferences were similar to those decided by a judge. Many factors were left implicit in some judgements, which Split-Up currently makes explicit. Some departures displayed by Split-Up from conclusions made in judgements can readily be made by small sample size.

The majority of the evaluation studies of Split-Up focus upon the system’s quality. In evaluating knowledge-based systems it is common to distinguish between the quality and the usefulness of the system. A system’s quality concerns such aspects as the quality of the system’s knowledge base and reasoning mechanism and (in particular) the quality of the system’s output when applied to a problem. The system’s usefulness instead concerns the effects of using the system in practice. Apart from a few interviews, no vigorous field studies have been conducted. However, anecdotal experience has led us to believe that Split-Up is of assistance in advising mediators and divorcees about possible negotiation stances in family law disputes (Zeleznikow & Bellucci, 2003), but provides limited support to lawyers and judges.

Current research is focusing on showing that the system is useful. We are also currently investigating developing Split-Up in an on-line dispute resolution environment.

It is essential to regularly update the Split-Up training sets. When new factors are introduced, we need to redesign the Hierarchy of relevant factors for percentage split determination described in Fig. 4.

4.3 Maintaining and updating the Split-Up system

Currently, following the advice of domain experts, the Split-Up system uses 94 different attributes. The Split-Up architecture provides no mechanism for determining whether the factors are relevant in empirical terms. It is possible that many of the factors declared relevant by our experts do not, in practice, contribute to a prediction. Thus, a family law prediction could possibly be made with only a subset of the factors regarded as relevant by experts.

We have applied feature selection techniques using genetic search to the data used to determine percentage split in the Split-Up system (Skabar et al., 1997). We have used genetic algorithms to determine which attributes are essential to model when distributing marital property. Our research shows a more accurate prediction can be made when using 16 of the 94 variables. An interpretation of this result is that Family Court judges when distributing property rarely use the other 78 attributes.

A major problem with the use of rule-based legal decision support systems is the issue of maintaining and updating the knowledge base. Changes in norms—through the introduction of new legislation or decisions in landmark cases, can lead to a total re-writing of the rule base.

When using cases to build legal decision support systems, if we use landmark cases then we need to give the new landmark cases weights. These weights will depend on the level of the court and how recent decision was made. The estimation of such weights can cause major problems. On the other hand, all commonplace cases carry an equal weight. It is the number of new cases that make a given argument, which proves significant.
So for the top-level Toulmin Argument (part A in Fig. 3) in the Split-Up hierarchy, landmark cases tell us that in determining the percentage of property awarded to the husband, contributions must be taken into account as well as the level of wealth of the marriage. Legislation (Section 79[4]) tells us that needs must be taken into account. But only the use of KDD from commonplace cases provides a guide as to how judges balance needs, contributions and the level of wealth of the marriage when determining an equitable distribution of the common pool.

Currently, the Split-Up tree of arguments is being modified in conjunction with domain experts from Victoria Legal Aid to accommodate recent changes in legislation and practice—in particular

(a) The recent tendency by Family Court judges to view domestic violence as a negative financial contribution to a marriage.
(b) The re-introduction of spousal maintenance as a benefit to one of the partners. Under the clean-break philosophy, judges of the Family Court of Australia were reluctant to award spousal maintenance, since it would mean one partner would continue to be financially dependant on his/her ex-partner. However the increasing number of short, asset-poor, income-rich marriages has led to a re-consideration of the issue of spousal maintenance.
(c) The need to consider superannuation and pensions separately from other marital property.

5. Conclusion

In this paper, we have focused on how we can model decision-making in discretionary domains. To do so, we have examined the notion of different levels of stare decisis, and how cases can be used in modelling domains in which traditional stare decisis is not strongly followed.

We have observed that whilst most traditional case-based reasoners use landmark cases to model norms in legal domains, a different approach is needed for discretionary domains. This has led us to the distinction between landmark and commonplace cases. We then investigated what computing techniques can be modelled in discretionary legal domains, and indeed, whether it is unwise to build reasoning systems for certain legal domains.

The over-arching claim we make is that, provided the data that reflect the decision-making processes are collected (data selection, data pre-processing and data transformation), KDD techniques are particularly adept at discovering patterns of judicial decision-making in discretionary fields of law. We introduced the Split-Up system as an example application of what may be achieved with the use of KDD.

The development of the Split-Up system required a detailed underlying knowledge of how Australian marital property is distributed following divorce. We claim that a detailed knowledge of the context of the domain is necessary when using KDD to learn about decision-making in discretionary domains. However, given this knowledge and an abundance of commonplace cases, KDD can be effectively used to provide decision support in discretionary legal domains.

Acknowledgements

The Split-Up system was developed by Andrew Stranieri as part of his doctoral and post-doctoral work under the supervision of John Zeleznikow. Mark Gawler, Bryn Lewis and Andrew Skabar, also graduate students of John Zeleznikow, provided minor programming support. Dan Hunter and Richard Ingleby (both then at the University of Melbourne) and Renata Alexander and Domenico Calabro (Victoria Legal Aid) provided legal advice.
The two anonymous referees provided excellent advice which greatly improved the quality of the paper.

REFERENCES


STRANIERI, A. 1998 Automating legal reasoning in discretionary domains, Ph.D. Dissertation, La Trobe University, Melbourne, Australia.


