



THE AUSTRALIAN NATIONAL UNIVERSITY

TR-CS-03-01

**Building and Testing the
SHYSTER-MYCIN Hybrid Legal
Expert System**

**Thomas A. O'Callaghan, James Popple and
Eric McCreath**

May 2003

Joint Computer Science Technical Report Series

Department of Computer Science
Faculty of Engineering and Information Technology

Computer Sciences Laboratory
Research School of Information Sciences and Engineering

This technical report series is published jointly by the Department of Computer Science, Faculty of Engineering and Information Technology, and the Computer Sciences Laboratory, Research School of Information Sciences and Engineering, The Australian National University.

Please direct correspondence regarding this series to:

Technical Reports
Department of Computer Science
Faculty of Engineering and Information Technology
The Australian National University
Canberra ACT 0200
Australia

or send email to:

`Technical.Reports@cs.anu.edu.au`

A list of technical reports, including some abstracts and copies of some full reports may be found at:

<http://cs.anu.edu.au/techreports/>

Recent reports in this series:

- TR-CS-02-06 Stephen M Blackburn and Kathryn S McKinley. *Fast garbage collection without a long wait*. November 2002.
- TR-CS-02-05 Peter Christen and Tim Churches. *Febri - freely extensible biomedical record linkage*. October 2002.
- TR-CS-02-04 John N. Zigman and Ramesh Sankaranarayana. *dJVM - a distributed JVM on a cluster*. September 2002.
- TR-CS-02-03 Adam Czezowski and Peter Christen. *How fast is -fast? Performance analysis of KDD applications using hardware performance counters on UltraSPARC-III*. September 2002.
- TR-CS-02-02 Bill Clarke, Adam Czezowski, and Peter Strazdins. *Implementation aspects of a SPARC V9 complete machine simulator*. February 2002.
- TR-CS-02-01 Peter Christen and Adam Czezowski. *Performance analysis of KDD applications using hardware event counters*. February 2002.

Building and Testing the SHYSTER-MYCIN Hybrid Legal Expert System

Thomas A. O'Callaghan James Popple Eric McCreath
<tom@tom-ocallaghan.com> <james@popple.net> <eric.mccreath@anu.edu.au>

Department of Computer Science
Faculty of Engineering and Information Technology
The Australian National University

Technical Report TR-CS-03-01

May 2003

Abstract

SHYSTER-MYCIN is a hybrid legal expert system created by combining rule-based and case-based reasoning. The MYCIN part uses a system of rules to reason with provisions of an Act of a parliament; the SHYSTER part uses analogy to reason with cases that explain "open-textured" concepts encountered in legislation.

The construction of the expert system is focused upon: creating and evaluating a model of legal reasoning, and improving the reporting made by the MYCIN part.

The model of legal reasoning is supported by jurisprudential discussion. The model holds that rules (in the strict sense of the word) cannot be extracted from cases. Cases should therefore be argued by analogy. The only rules that exist in law are those in legislation.

The method of evaluating the model of legal reasoning is comparative. Reports by the system are compared with reports by a test group of legally trained people. Both the system and the test group were provided with the same material on which to base their reports. This ensured that the evaluation was of the model of reasoning, rather than the depth of knowledge.

The reporting made by MYCIN was improved for use in SHYSTER-MYCIN, so that the system states how it comes to its conclusions. This reporting was then restricted to only the "interesting" conclusions.

Further information on the research described in this technical report is available in [21] and [22], and at <<http://cs.anu.edu.au/software/shyster/tom/>>.

1 Introduction

SHYSTER-MYCIN was created by the first-named author, and is the legal expert system discussed in this paper. SHYSTER-MYCIN combines rule-based reasoning with case-based reasoning. The system is designed as a legal expert system to be consulted by legally trained people. This hybrid system enables the case-based reasoner (SHYSTER) to determine open-textured concepts when required by the rule-based reasoner (MYCIN).

The system operates on a reduced version of the Australian *Copyright Act 1968*, including cases that define the term “authorization”¹ (see section 2). The Act is reasoned by a system of rules, whereas cases are reasoned by analogy. This approach is supported by jurisprudential discussions on legal reasoning (section 3).

The system was created in three progressive versions (section 4). The focus of the creation of the system was the reporting of reasons for conclusions. The second and third versions were tested against three criteria: validity, conciseness and correctness (section 5). The system performed well against those criteria, indicating that the approach taken is appropriate: that is, it is appropriate to use rules to reason with statutes and analogy to reason with cases.

2 Domain

In Australia, copyright law is governed by the Copyright Act. All rights, remedies and defences originate from the Act.² In Australian copyright law, cases provide examples of applications of the Act and interpretations of terms used in the Act.

SHYSTER-MYCIN operates upon a restricted area of the Act. Specifically the sections that are represented in SHYSTER-MYCIN are a set regarding ownership of copyrights, the acts that may be exclusively performed and the infringement of those rights. The cases that SHYSTER-MYCIN knows about are those on the meaning of the term “authorization”.

The term “authorization” is undefined in the Copyright Act. Consequently, a number of cases have been before the courts seeking answers as to what conduct amounts to authorization. The main contexts in which the issue has arisen are: “home taping of recorded materials, photocopying in educational institutions and performing works in public” [19, page 198].

In 2002, the Act was amended to include guidelines on defining “authorization”. However, this does not amount to a definition; the term remains undefined in the Act. Subsections 36(1A) and 101(1A) provide guidance when determining if an act can be said to be an authorization of an infringement of an exclusive right. These new provisions essentially codify the principles drawn from the leading Australian case.³

¹The meaning of the term “authorization” in the Copyright Act is one of the four areas of law on which SHYSTER was tested (see [25]).

²See section 8 of the Act: “[s]ubject to section 8A, copyright does not subsist otherwise than by virtue of this Act.” (Section 8A deals with the prerogative rights of the Crown in the nature of copyright.)

³*University of New South Wales v Moorhouse* (1975) 133 CLR 1, as explained in the Explanatory Memorandum to the Copyright Amendment (Digital Agenda) Bill 2000.

The facts to be considered in determining whether an infringing act was authorized should include:

- (a) the extent (if any) of the person’s power to prevent the doing of the act concerned;
- (b) the nature of any relationship existing between the person and the person who did the act concerned;
- (c) whether the person took any other reasonable steps to prevent or avoid the doing of the act, including whether the person complied with any relevant industry codes of practice.

The case-base is restricted to those cases up to 1983. The reason for this is that the focus of the construction of SHYSTER-MYCIN was on the creation of the hybrid system and the improving of the MYCIN component. The case law specification that was used with the SHYSTER part consisted only of decisions up to 1983 for reasons explained in [25, page 181].

3 Method of Reasoning

Susskind has criticised all developers of legal expert systems as failing to consider jurisprudence⁴ in their construction [33, page 194]. At best some attempted justification of their method after the construction of the system. Susskind suggests that jurisprudence should be the starting point for a legal expert system rather than merely a point of discussion after construction.

Whilst this may be a valid point, the works that Susskind suggests (Hart, Dworkin, Finnis or Raz) as starting points are not, at least immediately, useful. Jurisprudence almost solely deals with general questions such as “what is law?” and “what is good law?”. The jurists have rarely studied the question of “how do we argue with law?” or “how does a lawyer reason?” The answers or discussion on these questions must surely provide a more sturdy ground to begin the construction of a legal expert system than a theory of “what law is”. Chandler states that “[m]ore must be known about the mental operations a lawyer performs when engaging in case law research before the computer can be programmed to aid him to the full extent of its capacity” [10]. Bing is in agreement: “[a] computer program for legal reasoning cannot be created without first characterising the task to be performed and the means by which the reasoning agent performs it” [5].

Hart [14] views the legal system as heavily rule-based. He claims that rules can be extracted from all cases, and that these are “as determinate as any statutory rule” [page 135]. However, just prior to this statement, he concedes that there is “no authoritative or uniquely correct formulation of any rule to be extracted from cases” [page 134]. Hart attempts to reconcile these statements by claiming that whilst “no authoritative or uniquely correct” rule exists, there is “very general agreement” [page 134]; yet it still seems a leap of faith to characterise that “rule” as being as determinative as rules extracted from Acts of a parliament. If the “general agreement” is such that the rule is so well defined and understood by all, then why is there litigation? In Hart’s world, in every dispute both parties must already know which one will lose.

Eisenberg [13] explains in detail the process of extracting rules from cases. In his discussion on reasoning by analogy, he rejects Levi’s [17] characterisation (as described below) of legal reasoning as only involving reasoning by example/analogy:

It may be that example [analogy] has a part to play in the intuitive leap of discovery. Courts, however, cannot leave matters at that. Courts must justify their results by objective reasons that meet certain criteria, and must reject intuitive

⁴“The subject matter of jurisprudence, whether the discipline be classified as an art or science, is the nature of law and its working” [18, page 1]. Chihengo notes the etymology of the word “jurisprudence” from its Latin roots—*juris* “of the law” and *prudens* “skilled”—and indicates this vague definition of the term has meant that over the years the term has taken on a number of different meanings [11, page 2]. He continues, stating that “jurisprudence may be said to involve the study of a wide range of social phenomena, with the specific aim of understanding the nature, place and role of law within society”. The two chief divisions of jurisprudential enquiry were defined by Austin as analytical and normative, addressing general questions of “what is law?” and “what is good law?” respectively [4].

conclusions that they cannot justify in this way. In a normative context, justificatory reasoning can proceed only from standards, and “reasoning by example,” as such, is virtually impossible. Reason cannot be used to justify a normative conclusion on the basis of an example without first drawing a maxim or rule from the example (or, what is the same thing, without first concluding that the example “stands for” a maxim or rule). [page 86]

Eisenberg believes that a rule must be created somewhere in the process of reasoning by analogy. However, in his description above, analogy can occur as the first stage in the process. The “concluding that the example ‘stands for’ a maxim or rule” is simply the process of stating how the two cases are alike. The rule is created with a specific association and application in mind. It is therefore not a general, but instead a very specific “rule”.

Aarnio puts forward a view of legal reasoning in which induction is allowable, but it only provides prediction, not certainty. Aarnio contrasts law with nature, declaring that in nature there are “regularities, instances of invariability” which permit generalisation, whereas law is “volative, a result of human will” [1, page 79] and consequently generalisations can never hold.

Aarnio explains [at page 79] the inability to generalise in law by an example of the localisation of any “rule”: “If a person, then, has checked cases *a*, *b*, *c* and *d* and has stated that legal principle N_i is expressed in all of them, this does not yet entitle him to claim that the principle is *general* in nature.”

Whilst Aarnio rejects the idea of declaring a rule “general in nature”, he believes that prediction is possible. When there are several cases all of which express the same rule, then the “possibility to draw up a plausible prediction *increases*” [page 256]. He does, however, deny the possibility of making a prediction or rule from a single case.

Although more cases give a better chance of defining a rule, Aarnio warns of Dray’s paradox: “a law providing a historical explanation may be sufficiently comprehensive only when it contains such a large number of restrictive conditions that in the end it only concerns the individual case that should be explained” [page 73]. As the certainty of a prediction increases with the addition of more cases from which the prediction is inducted, the generality of the prediction decreases.

Aarnio cautiously concedes that inducting a rule may be possible from cases. However, this concession is made with such a limitation for it to almost be a denial of the appropriateness of inducting a rule.

The concept of a rule fits very neatly into computer science, as “if ... then ...” statements have been a part of computing since its inception. However, the ease of handling such representations should not be the motivation for modelling the real-world system in that way.

In the second edition of his work Allen separates legal reasoning into two categories which are neatly separated by the line between case-law and statute-law. Allen states [2, page 248]: “Whereas precedent is inductive, enactment clearly imposes the necessity of deduction upon the Courts. It is general and comprehensive in form, precedent particular and limited. A decision, whatever implications may be read into it by subsequent comparison and interpretation, exists primarily for the settling of a particular dispute: a statute purports to lay down a universal rule.”

By the time of his seventh edition [3], Allen states that, whilst the method of arguing with cases is usually termed induction, what really happens is argument by analogy. Allen regards analogy as the best and most common form of argument—“a close analogy is more convincing than a far-fetched illustration” [3, page 286]. “Every *ratio* is an *interpretation* of authorities in the light of the facts of the instant case ... The ratio is thus in a constant

state of flux ... it is not susceptible of any precise and comprehensive definition" [page 60]. When interpreting the *ratio* of a case in light of the instant case, analogy is necessarily involved. Thus there is a blurring in the four-step process typically taught to law students of "Issue-Rule-Application-Conclusion" (see for example [9, page 58–60]), that is, the rule and its application should be considered as one question—how the rule is to be applied to the facts of the instant case dictate how the rule will be formed.⁵

Levi [17] explicitly states that the process of "Issue-Rule-Application-Conclusion" is not just blurred, but in fact is reversed (at least in the middle). Levi believes that the use of analogy is *the* method of arguing with cases in law—"the finding of similarity or difference is the key step in the legal process" [page 2]. By arguing with cases using the method of analogy, "the rules arise out of a process which, while comparing fact situations, creates rules and then applies them" [page 4].

Levi admits that such a description of the process of legal reasoning will not sit well with lawyers and judges, as it "runs contrary to the pretense of the system" [page 9]. However, he sees it as much more dangerous to continue in the belief of a system of rules being established from cases: "[t]he rule will be useless. It will operate on a level where it has no meaning ... The statement of the rule is roughly analogous to the appeal to the meaning of a statute or of a constitution, but it has less of a function to perform. It is window dressing. Yet it can be very misleading" [page 9].

Leith [16] begins his discussion of the "AI Man's View of Law" with the following observation: "it is almost as though when God made computer scientists, he made them all think of law in the same way—as a system of rules."

Leith views the law as being more than simply a system of rules. He states [at page 511] that: "it seems to me to be all very well to draw up a collection of rules from legislation; but, as lawyers all know intimately, a piece of legislation is but one thing in the legal world." Leith does not explicitly state that rules are an inappropriate way of reasoning with cases, but it is an obvious conclusion to make based on this statement.

Leith therefore presents the view that more than rule-based reasoning is required to reason in "the legal world". Leith states that rule-based reasoning is appropriate for statutes, but that it is not appropriate for the rest of the law (for example, cases).

Schauer states that whilst we speak of rules in the common law, they are "so malleable so as not to even be rules" [29, page 177]. Schauer appears to be of the view that rules in the proper sense cannot be elicited from previous cases: "[precedent] cannot serve to provide the rule-like constraint" [page 184]. Schauer states that there is an ambiguity in the word "rule". This ambiguity causes some jurists to believe that the "rule of law" means that the law consists of rules. Schauer explains [at page 167] the use of the word "rule": "[i]n the sense that we have *rulers* who *rule* their subjects, 'rule' bears its closest affinity with 'reign' or 'control', and has only the remotest relationship with a form of decision-making characterized either by generality or by the entrenchment of generalizations."

At page 177, Schauer explains that whilst "rules" may claim to be applied, their application is not by way of interpretation. Rather "rules" are used as guidelines:

[a]lthough lawyers and judges can describe any number of common-law rules, and although both opinions and textbooks can state them in 'black letter' fashion, the rules have no single authoritative formulation, and accordingly the process of applying them does not involve an interpretation of the *text* of the rule ... it

⁵This method of reasoning appears to be heading the way of the rule skeptics. The rule skeptics see citing of legal rules in judgments as a *ex post facto* justification of the decision in a case rather than the sources upon which to reach the decision.

appears that common-law 'rules' are indeed descriptive rather than prescriptive, functioning merely as temporary guides.

Schauer [at page 178] agrees with the rule skeptics, in that "[t]he common law appears ... to be decision according to justification rather than decision according to rule."

Schauer identifies that there is a problem with claiming to find "rules" in cases. The problem that Schauer identifies [at page 183] is that at the outset of constructing a rule, the predicate (the facts of the case) must be stated. These facts cannot be easily stated:

What distinguishes reasoning from precedent from reasoning from rule, however, is the necessity in precedential reasoning of *constructing* the generalization/factual predicate that already exists in the case of a rule. As we have seen, the factual predicate of a rule, a generalization necessarily encompassing a multiplicity of events, is part of the rule's canonical form. But where there is only a previous decision and no rule-formulation, the source of the factual predicate is obscure, and consequently the manner in which the previous decision constrains becomes problematic.

For Schauer, the concept of precedent is to ensure the same result on the same facts. He notes [at 183]: "[n]o two events are exactly alike, but the idea of precedential constraint presupposes that a prior decision will control a subsequent set of facts that are *like* the first." Here the word "like" describes a definite association. That is, one case is "like" another. Cases are therefore compared by analogy.

To create a rule from a case would be to alter the use of "like" to describe indefinite, possible association. Schauer does not agree with the creation of rules from cases, and consequently would not condone this extension of the meaning of the word "like".

The comments that Schauer has made as to the problems with constructing "rules" are restricted to the common law. That is, Schauer does not agree with the proposition that cases can be argued with by use of rule-based reasoning. The method of reasoning that Schauer advocates is that of analogy.

Some of those who support the idea of extracting rules from cases admit their shortcomings. These "rules" are: local not general, temporal rather than permanent, and subjectively considered correct rather than universally.

Given these limitations, it hardly seems appropriate to refer to whatever is extracted from a case as a "rule". It is a guide that is extracted or a principle. The use of the word rule (in legal circles) is not designed to convey the same strictness as a rule say in mathematics—something that cannot be broken.

The use of analogy to compare cases seems to fit well with the doctrine of precedent.⁶ Such a method of reasoning has support from some of the above-mentioned jurists.

⁶To select an appropriate line of authority, the courts adhere to the "doctrine of precedent". At a practical level, the doctrine provides a method of deciding how binding or applicable a previously decided case is, based upon which court decided that case. At a policy level, the doctrine provides certainty, equality, efficiency, and the appearance of justice (in the sense that the law is seen to operate consistently). Cook et al. have summarised the general rules of precedent [12]:

- each court is bound by decisions of courts higher in its hierarchy;
- a decision of a court in a different hierarchy may be of considerable weight but will not be binding;
- only the *ratio decidendi* (the judge's decision on the material facts) of a case is binding;
- any relevant decisions, although not binding, may be considered and followed; and
- precedents are not necessarily abrogated by lapse of time.

The idea of analogy is to show how one thing is like another. The doctrine of precedent purports to treat like cases alike.

4 The Design of SHYSTER-MYCIN

SHYSTER-MYCIN combines MYCIN and SHYSTER, two previous expert systems. MYCIN⁷ is a medical expert system, which was adapted for use in SHYSTER-MYCIN. SHYSTER⁸ is a legal expert system, and was used without alteration in SHYSTER-MYCIN.

MYCIN's "certainty factor" is not used in SHYSTER-MYCIN.⁹ The reason for this is the difficulty in scientifically establishing how certain a fact is in a legal domain. In medicine, the "certainty factor" can be established by calculating the error in measurement, or statistically measured certainties of test results. In the law, the vast majority of conclusions cannot be established by scientific methods and therefore a "certainty" cannot be attached to them. For example the speed of a motor vehicle can be observed and the level of certainty of that observation stated. However, the level of certainty that (given that speed) the driver was driving at an excessive speed in all the circumstances¹⁰ cannot be determined by way of a formula.¹¹

In SHYSTER-MYCIN, the MYCIN part is used to reason with the provisions of an Act of Parliament only. The MYCIN part is not used to reason with so-called rules from decided cases. Similarly, SHYSTER does not reason with the provisions of an Act—its domain is decided cases.

This clear delineation between the use of rule-based and case-based reasoning has not always been made in hybrid legal expert systems. Some previous hybrid systems would represent so-called "clear" cases using rules, storing the remaining cases in a case-based system (see for example CABARET [26, 27, 28, 30, 31, 32] and GREBE [6, 7]).

The SHYSTER part of SHYSTER-MYCIN has been left untouched, and is only called upon for questions relating to its knowledge on the definition of "authorization". In this

⁷Created through "collaboration between the medical and AI communities at Stanford" [15], beginning in 1972. The version of MYCIN that was used for SHYSTER-MYCIN was created by Peter Norvig [20] (and is available at <http://www.norvig.com/paip.html>). Consequently, comments about the system are comments about the Norvig version of MYCIN, not necessarily the original system itself. For an account of the original system, see [8].

⁸Popple created SHYSTER as part of his PhD research at the Australian National University [24]. SHYSTER "represents the state of the art in statistical legal reasoning" [23]. See also <http://cs.anu.edu.au/software/shyster/>.

⁹Or, more precisely, all the certainty factors are set at 1.

¹⁰For example, perhaps the driver was travelling faster than the posted speed limit, but was doing so because she was transporting to hospital a person who had just suffered a heart attack. Objectively she is "speeding" as she is travelling at a speed greater than the posted limit. Supposing that there is provision to be excused from a fine for speeding if the speeding was "necessary or not excessive in the circumstances", the case described in this footnote would fit the "necessary or not excessive in the circumstances" test, as she was speeding in the hope of saving a person's life. The objective fact can be described, noting the margin of error in making the measurement of the car's speed. This can be expressed using a "certainty factor". The subjective fact is a unique, subjective weighting of a multitude of secondary facts. As the weighting of the secondary facts is also subjective, a "certainty factor" cannot be attributed to the primary subjective fact.

¹¹A statistician may claim that, with a sufficient number of cases, a certainty factor could be established. The first problem is that the majority of cases do not make it onto a public record, having been settled outside the court system. However, if we accept a large enough number of cases could be observed, taking the example in the previous footnote, the number of "valid excuse" cases could be compared with the total number of speeding cases, resulting in a "certainty factor". This would give the probability that the accused *may* have a valid excuse. Importantly it says nothing about the actual validity of the excuse. Consequently a "certainty factor", so determined, has a very limited use—it can only provide the user with an estimation of how valuable further facts may be.

way the SHYSTER part is like an expert that the MYCIN part calls upon when it cannot answer a question.

4.1 The intended use of SHYSTER-MYCIN

SHYSTER-MYCIN is created as a tool for use by lawyers or para-legals. It is designed as a system that would fall into the “internal knowledge systems” quadrant of Susskind’s “Legal Grid” [33, page 9]. That is, the system is designed to speed up internal processes of handling a matter. With an improved method of fact elicitation and classification, the system could be moved into the “online legal services” quadrant. This improvement would require the system to “get the facts right” when questioning a lay person.

4.2 The versions of SHYSTER-MYCIN

SHYSTER-MYCIN was produced in three different versions. The first version was very basic, and was used to make a preliminary assessment of the appropriateness of coupling a rule-based and a case-based system to reason with sections and cases, respectively. The second version of the system had a greatly increased rule base, and alterations were made to the reporting of results so that reasons for a conclusion, rather just the conclusion, would be reported. The third version of the system produced more concise reports, by limiting the conclusions that would be reported.

4.2.1 Version 1

The first version of SHYSTER-MYCIN (“SM-v1”) was created to provide a preliminary assessment of the approach that would be undertaken with the later versions of SHYSTER-MYCIN. That approach was to provide the MYCIN part with a rule base drawn from provisions of the Copyright Act, with SHYSTER available to be called upon when an “open textured” term¹² was encountered.

When an “open textured” term was encountered, the user was notified that it was a term that might be best answered by consulting the SHYSTER part. The user had a choice at this stage to answer the question based on their own knowledge, or to consult SHYSTER. If SHYSTER was consulted, the user answered SHYSTER’s questions, and, at the end of the consultation, was given the likely result. The user then gave this answer to the MYCIN part. This allowed the user to “over-rule” the SHYSTER part if they so wished.

SM-v1 has a rule base of 16 rules; it draws conclusions based on the values of nine parameters. The rules are used to represent subsections 13(2), 36(1), and 101(1) of the Copyright Act. These are the three provisions of the Act in which the term “authorization” is used. As explained in section 2 above, this term, whilst used in the Act, remains undefined.

The parameters for SM-v1 are the facts that are asked of the user or that are determined by applying rules to the facts obtained from the user. The parameters that SM-v1 uses are:

1. the name of the material;
2. the type of the material;
3. whether the accused was the owner of the material;
4. whether the accused had a licence to use the material;

¹²Specifically, for this system, the term “authorization”.

5. whether the accused had authorized someone else to use the material;
6. whether the use of the material occurred in Australia;
7. whether the accused had infringed the owner's rights under subsection 13(2);
8. whether the accused had infringed the owner's rights under subsection 36(1); and
9. whether the accused had infringed the owner's rights under subsection 101(1).

The last three parameters are the "goal" parameters. These are the facts that SM-v1 attempts to establish by applying the rules it knows to the facts asked of the user in relation to parameters 1–6.

Interactions with SM-v1 indicated that the approach proposed would be valid, and worth continuing. The MYCIN part of SM-v1 was able to work logically through the provisions in the Act that were provided to it as rules. However, SM-v1 only "knew" of three subsections of the Act. This meant that the answers to questions asked by SM-v1 relied upon the user being familiar with the remainder of the Act. For example, the user had to make assessments as to "ownership" of copyrights—something covered by other provisions of the Act.¹³ Also, that the copyright material was used by the accused, and that that use was an exercise of one of the exclusive rights¹⁴ for that material, were facts assumed by the system.

In advancing from SM-v1, more provisions of the Act had to be added to the system, and better reporting had to be implemented.

4.2.2 Version 2

SM-v2 uses the same approach as SM-v1 in that the MYCIN part reasons with provisions of the Copyright Act and the SHYSTER part reasons with decided cases. Version 2 differed from version 1 in three areas: the size of the rule base, the debugging of the MYCIN part, and the reporting of conclusions.

The rule base that the MYCIN part was working from was greatly expanded in version 2 as compared with version 1. The provisions of the Act that the MYCIN part knew were increased, making the system more realistic. This meant that most of the terms or concepts encountered in subsections 13(2), 36(1) and 101(1) were determined in surrounding sections. For example, to determine whether the accused was the owner of the copyright, rules representing subsection 35(2) were added to explain that a person who authored a work owned the copyright. In SM-v2 the provisions that the MYCIN part knows about are:

- sub-s 13(2) the right to authorize acts;
- s 31 the acts the owner has an exclusive right to (for works);
- s 35 determining the owner of a copyright (for works);
- s 36 how copyright is infringed (for works);
- ss 85–88 the acts the owner has an exclusive right to (for subject matter other than works);
- ss 97–100 determining the owner of a copyright (for subject matter other than works); and
- s 101 how copyright is infringed (for subject matter other than works).

The rules that represent these provisions of the Act total 273. The number of parameters used by these rules increased from 9 to 56.

¹³Sections 35 and 97–100 of the Act.

¹⁴Sections 31 and 85–88 of the Act.

In the process of expanding the rule base and altering the reporting (discussed below), it became apparent that a method of viewing how the MYCIN part stepped through its rules would be useful. To achieve this, the MYCIN reasoner was altered so that, when each question was asked, information about the rule currently under consideration was recorded in a file. This record was called a “stream of consciousness”. It detailed why the MYCIN part was asking each question, and provided information as to how the system arrived at its conclusions.

This record was useful in debugging the rule base. The file provided a step-by-step record, which assisted in checking that the rules were entered in the way that they were meant to be entered, in order to accurately represent the provisions of the Act. This record of the stream of consciousness was also the first attempt to improve the reporting. However, it was immediately obvious that the record was far too long. All rules that came under considerations were recorded. This meant many non-firing¹⁵ rules were included in the file. Thus the interesting parts were hidden by the sheer volume of the record.

Prior to altering the MYCIN part, the reporting was very limited. When a conclusion was reached, only the conclusion would be reported at the end of the consultation. Importantly, no reasons supporting the conclusion were reported.

The reporting of conclusions was improved in SM-v2 to make the MYCIN part fit with the generally accepted definition of an expert system. That is, an expert system should report on reasons for reaching conclusions, rather than simply return the conclusions.

The improvement to the reporting was made such that, when the reasoner was concluding a rule, the “report-why” function would be called upon. This function writes to a file the facts that were known, the rule that was applied to these facts, and the conclusion that was consequently made. The report is made using \LaTeX tags, so that the report from the MYCIN part can be combined with the output from the SHYSTER part (which also produces \LaTeX output), to produce a more cohesive report.

SM-v2 reported upon every conclusion drawn; this approach was slightly—but significantly—altered in SM-v3.

4.2.3 Version 3

SM-v3 operated on the same rule base that SM-v2 did, however, the reporting of conclusions was altered. The reporting done by the MYCIN part was restricted to only reporting on conclusions that were made by relying on more than one fact.

5 Testing and Results

The reporting made by SHYSTER-MYCIN was assessed against three criteria: validity, conciseness and “correctness”. The testing of SHYSTER-MYCIN was made by way of comparison with reports from three legal professionals.

5.1 The testing methodology

To determine the appropriateness of the approach taken in SHYSTER-MYCIN, the reporting that it makes was assessed. To do this the system’s report was compared with reports made by legal professionals. The test group consisted of three legal professionals: a law graduate, a practising solicitor with five years’ experience and one with 30 years’ experience.

¹⁵When a rule “fires” it is activated. That is, the premise are all found to be true, and the consequence is to reach the prescribed conclusion, or perform the required action.

None of the test group was expert in copyright law. The reason for selecting such a group of people was to make comparisons with SHYSTER-MYCIN on as level a playing field as possible. The test group was provided with the same material that SHYSTER-MYCIN was provided with.

This method of testing differs from previous evaluations of expert systems. Previously, expert systems were made to compete with human experts which were allowed to draw on years of experience and knowledge not represented in the expert system. Such an evaluation will result in comments about the inadequacy of the volume of knowledge in expert system. However, the expansion of an expert system's knowledge is simply a matter of time and resource.

The method of testing SHYSTER-MYCIN provides evaluation of the model of legal reasoning. Using this method of testing, if a model of legal reasoning is evaluated favourably, then appropriate investment can be made in expanding its knowledge.

5.2 The testing process

The test group was given a series of short questions to answer. The group was instructed to answer these questions using a short version of the Copyright Act and case summaries. The version of the Act was the same set of sections that SHYSTER-MYCIN knew (except for a few preliminary sections: 1–10). The cases summarised were those in SHYSTER's case law specification for the meaning of "authorization". The case summaries provided the test group with: the name of the case, the facts of the case and some commentary on the case. Using only these materials, the test group answered the questions.

SHYSTER-MYCIN was used to answer the same set of questions that the test group answered.

5.3 The validity

A report by SHYSTER-MYCIN was valid if it referenced the same sections as a majority of the test group. SM-v2 took a lazy or verbose approach, reporting on every section on which it made a conclusion — this was invariably every section that it knew. SM-v3, on the other hand, only reported a handful of conclusions, yet also managed to make valid reports.

By limiting the reporting of SM-v3 to conclusions made on more than one fact, SM-v3 would reference the sections that the test group would reference and usually only 2–3 extra sections. SM-v2 would make approximately 11 excess references, as compared with the test group.

5.4 The conciseness

The conciseness of each report was measured by counting the number of conclusions stated. The more concise a report was, the better it was considered to be. A concise report is considered to be the ideal because the system should be able to provide reasons for its conclusions, but should not regurgitate a copy of its rule-base as "explanation". Explanation should involve a condensing or summarising of information. The user should only be told the "interesting" parts—however that may be defined. However, a report should not become more concise at the expense of its validity.

On average, SM-v3 reported only 24% of the conclusions that SM-v2 reported. Just on this comparison, version 3 seems to have an advantage over version 2. According to SM-v3, most of the conclusions (approximately 3/4 of them) are uninteresting. The criterion that

SM-v3 uses to eliminate the uninteresting conclusions is to not report the conclusions made by applying a rule to a solitary fact (see section 4.2.3). This criterion is based upon the idea that conclusions based on a single fact are simply one-to-one mappings between one fact and another and do not give the user any real information. At best the information that the user provided is regurgitated to them with a slightly different wording. Conversely, a conclusion is interesting if it is arrived at by combining several facts.

When comparing SM-v3 with the test group, it can be seen (in Figure 1) that SM-v3 does perform much better than SM-v2, yet there is still room for improvement. On average the test group reported 12% of the conclusions that SM-v2 did, or about half as many as SM-v3 did. In answering Question 2, SM-v3 and the test group were fairly equal in the number of conclusions reported. SM-v3 reported on 6 conclusions, with the test group averaging just under 5. The greatest difference was observed in answers to Question 5: SM-v3 reported on 5 conclusions, each member of the test group reported on only one. It is suggested that by the time of Question 5, the test group decided to rely heavily upon references to their earlier answers and, as a consequence, only had a single conclusion each to report.

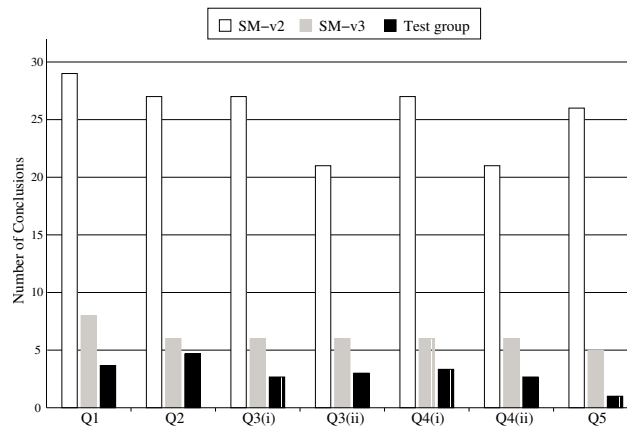


Figure 1: The number of conclusions reported for each question

5.5 The “correctness”

The “correctness” of the report is assessed by seeing if the “right” result is returned as an answer to the question raised by the factual scenario. The “right” result is not a scientifically observable fact—in the legal domain, several answers may be logically correct. That is, the answer is a logical conclusion that results from the application of Acts of Parliament and the Common Law to the facts of the present case. Therefore, to assess the “correctness” of the results reported by SHYSTER-MYCIN, there must be some method of determining the “right” result for each question.

To this end, the results given by the test group were taken to be representative of the “right” result. As there were three members of the test group, the “right” result would be the result returned by a majority of the group. In this way the test group was akin to a panel of three judges hearing a case. It is therefore assumed that the results of the test group are indicative of a good answer and, when taking the majority view, a “right” answer.

A “correct” system will provide answers that are the same as the majority of the legal experts in the test group.

SM-v2 and SM-v3 operate on the same rule base and use the same methods for accessing and concluding rules. Their difference is only in the reports they make. Therefore, because

the rules that they use and their use of them is the same, they will always return the same result for the same set of answers to their questions. For this reason, when assessing the correctness of the system, the system generally (versions 2 and 3) will be referred to.

SHYSTER-MYCIN agreed with the experts for each answer to the questions. When opinion was divided amongst the experts, the system agreed with the majority view. The test group agreed on the result for each question except for two questions. These splits in the decisions of the test group arose because of disagreements over the classification of facts. When the facts as classified by the minority were used to answer SHYSTER-MYCIN's questions, the answer given by the minority was returned.

These results indicate that the system provides "correct" answers. Further, these results also indicate that a crucial part of reaching a decision is the initial step of classifying the facts of the case.

6 Conclusions

The results of testing SHYSTER-MYCIN show that the approach taken in constructing the system is appropriate. That is, it is appropriate to use rule-based reasoning when dealing with statutes, and it is appropriate to use case-based reasoning when dealing with cases.

The results suggest that if the entire Copyright Act and associated cases were represented within SHYSTER-MYCIN (that is, with the whole Act represented in the MYCIN part, and all of the open-textured concepts resolvable by reference to an enhanced case-base in the SHYSTER part), the system would be expert in the entire body of copyright law.

6.1 Future research

Given sufficient time and resources, the entire Copyright Act and relevant cases could be represented in SHYSTER-MYCIN.

Before translating the entire Act into rules for the MYCIN part, some sort of rule management system should be created. This is because, in creating SHYSTER-MYCIN, it was observed that a great number of rules are required to represent only a small number of provisions in the Act. Without a rule management system, the rule-base is "fragile": that is, changing it would most likely result in error.

The system could be further improved in the conciseness of the reporting by the MYCIN part. By defining a "positive conclusion", the reporting could be further restricted to only report conclusions made on more than one fact which was concluded positively. A "positive conclusion" would have the effect of assigning a direction to a fact. A positive answer would be one that brings the system a step closer to a goal.¹⁶

Although left untouched in the construction of SHYSTER-MYCIN, the reasoning performed by SHYSTER could be altered. Methods of analogy other than nearest neighbour could be employed by SHYSTER to select cases.¹⁷

The facts collected by the MYCIN and SHYSTER parts should be shared between the two expert systems. At present they each store their own sets of facts, and do not pass information between each other. If the facts were stored commonly, then the potential for a user to be asked the same question twice is eliminated. This has two benefits: the user does not become annoyed or frustrated by double-questioning, and conflicting answers are not provided.¹⁸

¹⁶The answer "false" could be "positive" under this system, depending on how the question is phrased.

¹⁷As suggested by Popple [25, page 251].

¹⁸Otherwise the SHYSTER part and the MYCIN part could be working on the basis of inconsistent facts.

SHYSTER-MYCIN could be tested by comparing the results given by the system when used by lay-people with the answers given by experts. This would test whether SHYSTER-MYCIN can accurately gather facts by questioning a person. Fact elicitation/classification is one of the areas that Susskind identifies as requiring greater research [33]. A system capable of getting the facts right by questioning a lay person would fall within the “online legal service” quadrant of Susskind’s “Legal Grid” [33, page 9]. Such a system could be of great use both commercially and for society.

7 Acknowledgments

Thanks to Mr T. E. O’Callaghan, General Counsel, Citibank, N. A. for editorial assistance.

References

- [1] AARNIO, Aulis 1977, *On Legal Reasoning*, Turun Yliopisto.
- [2] ALLEN, Carleton Kemp 1930, *Law in the Making* (second edition), Oxford University Press.
- [3] ALLEN, Carleton Kemp 1964, *Law in the Making* (seventh edition), Oxford University Press.
- [4] AUSTIN, J 1832, *The Province of Jurisprudence Determined*, Weidenfeld & Nicolson.
- [5] BING, Jon 1990, “Legal decisions and computerized systems”, in *From Data Protection to Knowledge Machines: The Study of Law and Informatics*, ed. P Seipel, Kluwer Law and Taxation Publishers.
- [6] BRANTING, L Karl 1989, “Representing and reusing explanations of legal precedents”, in *Proceedings of the Second International Conference on Artificial Intelligence and Law (ICAIL-89)*, University of British Columbia, Vancouver, pp. 103–10.
- [7] BRANTING, L Karl 1991, “Reasoning with portions of precedents”, in *Proceedings of the Third International Conference on Artificial Intelligence and Law (ICAIL-91)*, St Catherine’s College, Oxford, pp. 145–54.
- [8] BUCHANAN, Bruce G and SHORTLIFFE, Edward H (eds) 1985, *Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project*, Addison-Wesley Publishing Company.
- [9] CALLEROS, Charles R. 1994, *Legal Method and Writing* (second edition), Little, Brown & Company.
- [10] CHANDLER, James P 1974, “Computers and case law”, *Rutgers Journal of Computers (Technology) and the Law*, vol. 3, pp. 202–18.
- [11] CHINHENGO, Austin 2000, *Essential Jurisprudence* (second edition), Cavendish Publishing Limited.
- [12] COOK, Catriona, CREYKE, Robin, GEDDES, Robert and HOLLOWAY, Ian 2001, *Laying Down the Law* (fifth edition), Butterworths.

- [13] EISENBERG, Melvin Aron 1988, *The Nature of the Common Law*, Harvard University Press.
- [14] HART, H. L. A. 1994, *The Concept of Law* (second edition), Oxford University Press.
- [15] JACKSON, Peter 1986, *Introduction to Expert Systems*, Addison-Wesley Publishing Company.
- [16] LEITH, Philip 1986, "Legal expert systems: Misunderstanding the legal process", *Computers and Law*, no. 49, pp. 26–31.
- [17] LEVI, Edward H. 1961, *An Introduction to Legal Reasoning* (seventh edition), The University of Chicago Press.
- [18] MCCOUBREY, Hilaire and WHITE, Nigel D. 1999, *Textbook on Jurisprudence* (third edition), Blackstone Press Limited.
- [19] MCKEOUGH, Jill, BOWERY, Kathy and GRIFFITH, Philip 2002, *Intellectual Property (Commentary and Materials)* (third edition), Lawbook Co.
- [20] NORVIG, Peter 1992, *Paradigms of Artificial Intelligence Programming: Case Studies in Common Lisp*, Morgan Kaufmann.
- [21] O'CALLAGHAN, Thomas Alexander 2003, *A Hybrid Legal Expert System*, Honours thesis, Department of Computer Science, Faculty of Engineering and Information Technology, The Australian National University, Canberra, February. <<http://cs.anu.edu.au/software/shyster/tom/thesis.pdf>>.
- [22] O'CALLAGHAN, Thomas A., POPPLE, James and MCCREATH, Eric 2003, "SHYSTER-MYCIN: A hybrid legal expert system", in *Proceedings of the Ninth International Conference on Artificial Intelligence and Law (ICAAIL-03)*, Edinburgh, Scotland, 24–28 June, ACM, pp. 103–4. ISBN 1 58113 747 8. <<http://cs.anu.edu.au/software/shyster/tom/icail-03.pdf>>.
- [23] PANNU, Ananddeep 1995, "Using genetic algorithms to inductively reason with cases in the legal domain", in *Proceedings of the Fifth International Conference on Artificial Intelligence and Law (ICAAIL-95)*, College Park, Maryland, 21–24 May, pp. 175–184.
- [24] POPPLE, James 1993, *SHYSTER: A Pragmatic Legal Expert System*, PhD thesis, The Australian National University, Canberra, April. ISBN 0 7315 1827 6. <<http://cs.anu.edu.au/~James.Popple/publications/theses/phd.pdf>>.
- [25] POPPLE, James 1996, *A Pragmatic Legal Expert System*, Applied Legal Philosophy Series, May, Dartmouth, Aldershot. ISBN 1 85521 739 2. <<http://cs.anu.edu.au/~James.Popple/publications/books/shyster.pdf>>.
- [26] RISSLAND, Edwina L 1990, "Artificial intelligence and law: Stepping stones to a model of legal reasoning", *The Yale Law Journal*, vol. 99, no. 8, June, pp. 1957–81.
- [27] RISSLAND, Edwina L and SKALAK, David B 1989, "Combining case-based and rule-based reasoning: A heuristic approach", in *Proceedings of the Eleventh International Joint Conference on Artificial Intelligence (IJCAI-89)*, Detroit, Michigan, pp. 524–530.

- [28] RISSLAND, Edwina L and SKALAK, David B 1989, "Interpreting statutory predicates", in *Proceedings of the Second International Conference on Artificial Intelligence and Law (ICAIL-89)*, University of British Columbia, Vancouver, pp. 46–53.
- [29] SCHAUER, Frederick F 1991, *Playing by the rules: A philosophical examination of rule-based decision making in law and in life*, Oxford University Press.
- [30] SKALAK, David B 1989, "Taking advantage of models for legal classification", in *Proceedings of the Second International Conference on Artificial Intelligence and Law (ICAIL-89)*, University of British Columbia, Vancouver, pp. 234–41.
- [31] SKALAK, David B and RISSLAND, Edwina L 1991, "Argument moves in a rule-guided domain", in *Proceedings of the Third International Conference on Artificial Intelligence and Law (ICAIL-91)*, St Catherine's College, Oxford, pp. 1–11.
- [32] SKALAK, David B and RISSLAND, Edwina L 1992, "Arguments and cases: An inevitable intertwining", *Artificial Intelligence and Law*, vol. 1, no. 1, pp. 3–44.
- [33] SUSSKIND, Richard E 2001, *Transforming the law: Essays on technology, justice, and the legal marketplace*, Oxford University Press.