

Volume 7, Issue 2, August 2010

ZOMBAIS: LEGAL EXPERT SYSTEMS AS REPRESENTATIVES “BEYOND THE GRAVE”

*Burkhard Schafer**

Abstract

This paper explores the possibility of computer assistance for the interpretation of wills and testaments. It draws from experience with legal expert systems developed for the interpretation of laws and other legal norms.

DOI: 10.2966/scrip.070210.384



© Burkhard Schafer 2010. This work is licensed under a [Creative Commons Licence](#). Please click on the link to read the terms and conditions.

* Professor of Computational Legal Theory, University of Edinburgh.

It is a truth universally acknowledged that a zombie in possession of brains must be in want of more brains.

Seth Grahame-Smith

1. Introduction

The desire to exercise control beyond the grave is deeply rooted in the human psyche. Before we die, we try to create cues that preserve our identity in the minds of the survivors.¹ The survivor is left with images, materials, and wishes of the deceased that allow, or force, them to act on upon information and behaviours that were part of the deceased when he or she was alive.² This is nowhere more obvious than in the law governing wills and testaments. Even if (most of us) realise that we cannot take our wealth with us, many of us hope nevertheless to control at least in part how our financial assets are used when we are not around any longer.³ This too is in part an identity preservation strategy, as for exercising control beyond the grave there must be a something as substratum of this control. Once a prerogative for the powerful and wealthy whose testaments, famously like Caesar's, could shape the fate of entire nations, the testament's historical roots in the west can be traced back to the law reforms of Solon. It then became a mainstream tool for the disposal of assets in Roman law. Roman law also gave us the blueprint for the trust in the legates and *fideicommissa* and with that the instrument, not just for controlling who should inherit, but also an enforceable means for controlling how assets were to be used. "Communication technologies" played for obvious reasons an important part for wills and testaments from the beginning. Since the testator is not around by definition and cannot be asked for his or her opinion, he or she needs to find ways to reliably communicate his or her intentions to the executors in a will. The advent of writing and improvements of writing and document storage facilities in archives were from the beginning a driving force in the development of wills and testaments as tools to engineer and control ones future. More recently, use of video recordings added a new dimension of "immediacy" to the way in which a testator can communicate with the executor and the heirs. In a variation of this theme, US soldiers often make video recordings for their children prior to deployment into a battle zone, with the idea that if they do not return, the children will get parental advice at predetermined points in time.

This paper will explore if the "artificial brains" software developed in AI research could become the next generation of tools to exercise control "beyond the grave" and create identify maintaining cues in the way Unruh described which is similar to the "personal backup" popularised in the novels of the Scottish writer Iain Banks. It will argue that such an approach could revitalise previously abandoned themes in legal AI research. In the first part, we develop an analysis of the methodological challenges encountered by legal AI research in developing systems that can autonomously interpret legal norms. In the second part we describe a new application, the use of expert systems in inheritance law, which can use the positive insights that were gained

¹ R Butler, "Looking Forward to What? The Life Review, Legacy and Excessive Identity versus Change" (1970) 14 *American Behavioural Scientist* 121-128.

² DR Unruh, "Death and Personal History: Strategies of Identity Preservation" (1983) 30 *Social Problems* 340-351.

³ See e.g. J Rosenfeld, "Old Age, New Beneficiaries: Kinship, Friendship and (Dis)inheritance" (1980) 64 *Sociology and Social Research* 86-95.

in the early days of research into legal AI, while avoiding the systemic methodological problems earlier, more ambitious projects had encountered.

2. Back to the Future

In the early days of legal AI and legal expert system research, the image of the computer judge provided a powerful metaphor that brought the hopes and aspirations of the research community to the point. The image did not just influence research in legal AI, but also jurisprudence⁴ and popular literature.⁵ But despite several promising results, the ultimate goal of fully automated judicial decision making remained elusive. The early enthusiasm was followed by a period of introspection in the 1980s and 1990s, which led to an increasingly critical reassessment of the computerised decision maker in law, and a “new modesty” in legal AI research. As a result of this phase of introspection, systemic problems in the project of developing computer judges were identified that seem to make any attempt to revive the notion in 2010 unfeasible on conceptual, philosophical, methodological and ethical grounds. The idea of a fully automated reasoner that interprets legal rules and suggests solutions in specific cases seems today distinctly like an anachronistic return to the 1980s. The idea is dead and we should let it rest – or so it seems but, as one would expect in a special section on zombies, the dead do not always stay in their graves. This article will then try to revive the idea and to come up with a new “business model” for computer assisted norm interpretation that is fully informed by methodological debates within the AI and law community.

3. From Automated Legal Interpretation to Decision Support System

To achieve its aims, a computer judge should have been capable of applying general, abstract norms correctly to the facts of a specific case. But attempts to formalise this process of subsuming specific cases under general norms soon ran into apparently insurmountable difficulties:

1. The inherent vagueness of legal texts. To be sufficiently flexible and capable of regulating situations unforeseen by the legislator at the point of law making, legal language is necessarily vague to a certain degree. This results in a need for interpretation and with that the capture of the meta-level rules that guide the interpretative process.
2. The value-ladenness of law and legal language. To be able to give an adequate interpretation, judges need to refer to values implicit in the legal system. Their own moral, political and philosophical commitments play a necessary, albeit problematic, part in this process.
3. The symbiotic relationship between facts and norms. How the facts of a case are described and the factual proofs that were taken regularly preempt the legal interpretation.
4. The contested nature of law. In particular in appeal cases, both sides will have good arguments on their sides. The “one right answer” which is part

⁴ R Susskind, “Detmold’s Computer Judge Revisited” (1986) 49 *Modern Law Review* 683-684.

⁵ See http://theinfosphere.org/Computer_Judge (accessed 2 Jul 2010).

and parcel of discourses in the natural sciences hardly more than a jurisprudential abstraction in law. In particular, in cases where the court itself is divided and a decision is reached by simple majority voting, it is obvious that the opposite solution would also have been a possible that is a consistent, solution. For legal informatics, this opens up two follow-on questions:

- a. If it is possible that even the top experts rationally hold mutually contradictory opinions, what exactly is the “knowledge” that the computer models, and on what basis is the decision what to include, or which one to chose, taken?
- b. What does it mean for the evaluation of the computer judge? Under which conditions are we entitled to say that the programme is working correctly and that the “right” decisions are reached?

The wider AI community developed tentative answers for some of these questions, which were also received in research into legal AI. Layman Allen in particular developed formal approaches to model the interpretation of legal norms that drew extensively from ideas from generic natural language processing research in the 1980s.⁶ However, it became apparent that the specific methodological particularities of the legal domain impeded a wholesale adoption of approaches from general AI for the development of commercial-strength expert systems.

Modern AI research, for instance, makes extensive use of neural network and automated learning approaches, which allow computers to acquire knowledge on how to disambiguate vague terms in a given context.⁷ In legal AI, this approach has also been used with some success in systems such as *Split-Up* und and related systems.⁸ The comparative success of *Split-Up* is however based on the advantageous properties of the specific domain that it models, divorce law, and hence can only within strict confines be extended to other applications. It does not attempt to model reasoning by appeal courts on the meaning of legal terms, but takes decisions of first instance courts as input that deal primarily with “unproblematic” subsumption of fact-rich situations under legal provisions. The main difficulty a lawyer faces in these cases is the number of parameters that are relevant for the decision (such as the contributions of both parties to the acquisition of major items such as the family home during their marriage). Computers are good at keeping track of these large numbers of items under consideration and thus add real value to the practicalities of decision making. This focus on first instance decisions also guarantees a large number of training examples that are a prerequisite of automated learning based approaches. The legal rules in question often explicitly refer to mathematical operations to divide the assets between the parties, and in that sense display their formal structure so to speak on the surface. Furthermore, decisions are typically not simple yes-no answers, but permit the judge to divide the communal assets in a multitude of graded ways. It makes sense therefore

⁶ LE Allen and CS Saxon “Multiple Interpretations of the Logical Structure of Legal Rules: Impediment or Boon to Legal Expert Systems?” in RA Kowalski and KA Bowen (eds), *Logic Programming: Proceedings of the Fifth International Conference and Symposium, Seattle, Washington, August 15-18, 1988* (Cambridge, MA: MIT Press, 1988) 1609-1623.

⁷ S Lawren and S Fong, “Natural Language Grammatical Inference: A Comparison of Recurrent Neural Networks and Machine Learning Methods” (2006) 1040 *Lecture Notes in Computer Science* 33-47.

⁸ J Zeleznikow, “Building Judicial Decision Support Systems in Discretionary Legal Domains” (2000) 14 *International Review of Computers, Law and Information Technology* 341-356.

to ask for the “average” decision – what percentage of the property can a party “on average” expect, if the litigant was given custody of the children after six years of marriage? This is also one of the reasons why *Split-Up*, despite its success, is marketed as a decision support tool only, not as a full fledged computer judge. It is helpful for a party to know in advance what amount of money it can broadly expect, for instance to take a rational cost-benefit analysis. But this means of course that the decision in the individual case can still differ to a considerable extent, simply because it is reached at by a specific, not an “average” judge, and on a specific, not a “normal” set of circumstances. Taken together, recognition of the presuppositions that enabled the success of *Split-Up* allows us to identify some further methodological issues:

- c. The majority of court decisions are not published and hence are not available to train an expert system.
- d. Only a selection of cases is accepted by the appeal courts for decision, and hence there are even fewer authoritative training examples where word meaning is disambiguated.
- e. Even in these decided cases, the reasoning that informs the decision is not always sufficiently clear to make straightforward training examples, they often create ambiguities of their own, and asking the decision maker is not normally possible.

We can now reformulate the problem of computerised legal decision making. Legal language is ambiguous so that it can adjust to changing circumstances and unforeseen conditions. Therefore, a good test for a computer judge is the ability to predict decisions by real judges accurately. But since the number of relevant training examples and with that the empirical input is low, the number of variables and unknowns high – amongst them the moral, political and philosophical convictions of as yet unknown future judges – developing robust computerised decision makers is problematic. Robustness in expert system design is understood as the ability to deal with new and unforeseen circumstances correctly. But as we have seen, evaluating the “correctness” of the answers is in itself problematic. What counts as a “correct” decision is often contested within the relevant legal community. That is a problem in particular when the result must be a binary decision and the average outcomes are unhelpful.

We said above that the lessons that the AI and law community learned from these methodological reflections could be described as “new modesty”. Rather than aiming at computers that can interpret legal norms autonomously and reach a decision, computers are now mainly described as decision or argumentation support tools. Most of the actual interpretations of legal norms – the core skill of the legal profession – are done by the user, often in an ex-post-facto analysis. In this way, a user trained in law will be able to check if his or her proposed argument meets some formal minimum requirements of consistency and completeness, e.g. checking if all pertinent questions are answered, and no circular use of premises occurs.⁹

⁹ See also TF Gordon “Juristische Argumentation als Modellierungsprozess” in R Traunmüller and M Wimmer (eds), *Informatik in Recht und Verwaltung: Gestern - Heute - Morgen* (Bonn: Gesellschaft für Informatik, 2009) 104–112; FJ Bex et al “Sense-Making Software for Crime Investigation: How to Combine Stories and Arguments?” (2007) 6 *Law, Probability & Risk* 145-168.

This reorientation, however, also meant that many of the older ideas and approaches in legal AI, despite their theoretical validity, became irrelevant. We argue that this might have been premature. As the success of *Split-Up* shows, it is often the choice of the right domain that decides about the validity and success of a legal AI system. Can we find applications of legal reasoning that minimise the methodological problems identified above and nevertheless preserve the focus on the interpretation of legal norms and the disambiguation of terms? This could result in a revival of some of the older approaches, and a theoretically more ambitious and rewarding field of study.

4. Know Thyself – The User as Norm Giver and Norm Interpreter

Hauser: Howdy, stranger! I'm Hauser. If things haven't gone wrong, I'm talking to myself and you don't have a wet towel around your head. Now, whatever your name is, get ready for the big surprise. You are not you, you're me.

Douglas Quaid: [to himself] No shit.

Total Recall (1990)

As we have seen, early approaches to legal expert system design tried to find solutions for decision making situations that are governed by a large number of imponderables and unknowns. The legislator necessarily communicates with citizens and the judiciary using vague terms that need interpretation. How exactly parliament would have wanted its own laws to be understood under new and unpredicted circumstances can however only be determined through an indirect, complex and ultimately contestable process of interpretation. This becomes even more complicated if we try to determine not just abstractly how parliament, were it asked, would want a term to be interpreted, but also when we try to anticipate the values, convictions and methodological preferences of future judges as decision makers. To reduce this complexity, we would ideally want a situation where we have to deal only with a single norm-giver, and a single, already known “judge” or norm-interpreter. This person could then be systematically interrogated to establish how she or he would interpret the terms of her or his own norms under a variety of hypothetical conditions, giving us a potentially unlimited number of training examples. Once the system has been trained sufficiently, we could then rely on this person as an objective benchmark for the evaluation of our system: it is correct if it predicts how the judge would have interpreted the norm in question correctly.

The above quote from *Total Recall* is a fictitious example that comes close to the application suggested here. Hauser is about to get his memory wiped out to be transformed into the infiltrator Quaid. He needs to find a way to communicate to his own future self, which will have preferences and values intentionally designed to be very different from his present set of convictions about how to act. Technology gives him the means to do so in a recorded video clip where he explains to his own future self the background for his assignment and how he is now to understand the orders given to him. In this case, norm giver and norm interpreter are the same person (for a given value of “same person”) and technology acts as a mediator to ensure that the present incarnation, Quaid, interprets the norms (“infiltrate the rebellion”) in a way that the earlier persona, Hauser, would have approved of.

At this point, an analogy from another field of computer science research might be helpful. Research in computer assisted speech recognition distinguishes between

speaker dependent and speaker independent approaches.¹⁰ In speaker independent applications, the aim is to develop software for arbitrary, unknown users who can use the system immediately without needing training. This type of system tries for instance to handle calls to call centres. You can assume in advance that every caller will use one of several terms to identify his or her problem (“overdraft”, “charges”, “mortgage”, etc) but needs to be robust enough to predict how an arbitrary user the system has never encountered before will pronounce these words. Whenever the vocabulary can be kept small, e.g. in telephone inquiries, successful systems have been developed. Even specialist and technical vocabulary can be identified, but even the best systems currently available can only identify several thousand words.

By contrast, speaker dependent voice recognition software is “tailor made” for the individual user who can train the system on his personal particularities in pronunciation or dialect prior to use.¹¹ Typically, he will be given test sentences to train the system with, repeating them often enough until the computer recognises the sentence. On the down side, this means that the system will only work properly with one specific user. On the positive side, these systems have a much larger vocabulary and higher degrees of accuracy and reliability than speaker-independent systems.¹²

The traditional approach to legal expert system design was similar to speaker independent voice recognition. It does not matter who the user is, or who the judges will be who are going to evaluate and interpret a norm, the system will correctly predict their decision. As with voice independent speaker recognition, this is only feasible if from the beginning, the number of possible answers is highly restricted and a very small vocabulary suffices. Are there now applications conceivable that are more similar to speaker dependent voice recognition?

For these, the user himself would have to be at the same time norm giver and interpreter of his own laws. The first condition is easy to fulfill if we remember that the main purpose of private law is to enable citizens to establish their own rules between themselves. The contracts that we make with each other, the wills and testaments that we write, the property dispositions that we undertake all create legal norms which bind on the one hand ourselves, but also in an indirect way the judges who might have to adjudicate if a party fulfilled their contractual obligations.

As with parliament, private parties need to formulate the rules they agree to abide by in sufficiently general and vague terms to allow for possibly changing circumstances. This of course leads to familiar problems in the interpretation of contracts. Normally of course, in case of conflict the parties can communicate their respective understandings to each other, or an adjudicator if necessary. Implementing this process on a computer seems to offer little additional value. Are there however conceivable situations where the parties cannot themselves contribute to the interpretation of the terms of contract and cannot inform the decision maker how they intended the terms to be understood, their wishes, preferences and values that informed the norms and can disambiguate any problematic clause? In such a situation,

¹⁰ XD Huang and KF Lee, “On Speaker-Independent, Speaker-Dependent, and Speaker-Adaptive Speech Recognition” Proceedings of the International Conference on Acoustics, Speech, and Signal Processing, ICASSP-91 1991, available at <http://doi.ieeecomputersociety.org/10.1109/ICASSP.1991.150478> (accessed 28 Jul 2010).

¹¹ H Beigi, *Fundamentals of Speaker Recognition* (Springer: New York, 2010).

¹² J-C Junqua and J Haton, *Robustness in Automatic Speech Recognition: Fundamentals and Applications* (Berlin: Kluwer Academic Publishers, 1995).

an expert system that is sufficiently trained on the preferences of the user could assist in disambiguating the term in question.

In medical, inheritance and trust law, we find just this kind of situation. A living will establishes for instance general rules about how I want to be treated in case an accident or illness permanently deprives me of my ability to communicate or make decisions for myself. In inheritance and trust law, I can create rules regarding who should benefit from my property after my death by leaving it for instance to trustees to act on my behalf and to interpret the rules I laid down. In both cases, I postulate rules that I hope are clear and are able to deal with all possible eventualities. Yet we know of course that this is not always possible. To take one example, I might decree that my property should go after my death to “my grandchildren”. After my demise, it transpires that my son had an illegitimate child which he kept hidden from the family that I was unaware of.

A typical question in inheritance law in this case is the meaning of “my grandchildren”. Was the illegitimate but biological grandchild covered when I wrote “my grandchildren”?

For obvious reasons, asking me will not be possible.¹³ Hence, it will be necessary to interpret my rules, in the same way in which a judge would interpret an Act of Parliament.

Of course, evidence about my value system, convictions or religious and ethical beliefs are relevant pointers for that process, but inevitably a degree of speculation would be necessary. But what if I had trained a computer system to learn about my personal values in the same way in which I can train a computer to understand my voice? Such a system could then be interrogated in my stead, to hear an authentic account of “my” opinion on that matter. The neural network at the heart of the system would have been trained by asking me standardised questions (Do you think nature is more important than nurture – yes or no?) or generating typical ethical problems (Would I save someone who is related to me or someone who is my personal friend, from a burning building?) and asking for my opinion. Current research in “experimental philosophy”, with its emphasis on large datasets of answers from large numbers of people across cultures, plays an important role in developing the necessary methodology.¹⁴ The questions can become increasingly subtle and detailed and the user can spend as much time on training the software as he or she wishes.

In the next step, the system can be validated by the user, developing its own solutions that try to mimic the learned behaviour and attitudes, with user feedback for correct and incorrect answers. Based on this feedback the system can then model more and more accurately my convictions and the ethical rules that govern my behaviour. It is my decision when I consider the answers of the system “good enough” to entrust it with acting as my legal representative. Here we can see two important differences

¹³ In Ryūnosuke Akutagawa’s *Rashomon*, the ghost of the victim is allowed to give evidence in the trial against his murderer. Jacques Orneuve is a famous, real life (?) example of a zombie asking to be permitted to give evidence in court (see http://thefullzombie.com/topics/us_law_and_haitian_zombie (accessed 2 Jul 2010)). However, cross examining the undead is generally frowned upon by modern legal systems.

¹⁴ See e.g. J Knobe and S Nichols (eds), *Experimental Philosophy* (Oxford: OUP 2008); KA Appiah, *Experiments in Ethics* (Cambridge, MA: Harvard University Press, 2008); B Musschenga “Was ist empirische Ethik?” (2009) 21 *Ethik in der Medizin* 187-199.

from traditional legal AI: My decision is the only relevant benchmark for judging the correctness of the answer and I can generate as large a number of training examples as I wish – just like voice recognition software that never stops learning. Insights into the logical aspects of norm interpretation that have been developed in legal AI would still build the formal basis of such a system, revitalising and reusing older AI and law research. The knowledge base, on the other hand, is provided by the user. Taken together they should be capable of interpreting norms in the light of new situations based on the values and preferences of the user, and it would be trivially possible to quantify the accuracy that the system has acquired in its predictive power. Theoretically, the system would already be a success if it predicted how its owner would reinterpret his or her norms in the light of new circumstances better than a third party, like a judge, who had no personal knowledge of the deceased and had to base its decision exclusively on the textual basis of the will and testament. However, for legal and evidentiary reasons, one might require that the system is better than chance in its results, meeting the “preponderance of evidence” standard. Whether the system has the necessary level of proficiency can of course be documented easily as part of the learning process by simply keeping count of the ratio between wrong and right answers that the system gives.

5. Legal and Ethical Implications

The article has so far described in broad outlines the business model and basic formal features for an expert legal system that is capable of assisting a judge in interpreting norms created by a private party that has since died or is otherwise incapable of disambiguating the rules and norm it created. It uses methods that were first proposed for developing a universal legal reasoner or norm interpreter while avoiding most of the methodological and practical pitfalls that prevented such a system from becoming robust and reliable enough to be of practical use in the past.

The result would be a system that, like the zombies of lore, falls well short of the full intelligence of the person whose behaviour it models. But it would preserve nonetheless task specific knowledge that would make it a capable representative.

The proposal raises some interesting technical, ethical and legal questions. We anticipate that the greatest technical problem will be in formulating suitable training questions and examples. These need to be capable of generating the right general rules to model the value system of the person who trains it correctly. Also of crucial importance is ensuring that the system is secure and cannot be hacked into or otherwise taken over by a third party. If such a security breach were possible, and if we would really permit computers to act as legal representatives of their (deceased) owners, such a compromised system would indeed be a zombie, or rather “ZombAI” – acting as if it is the voice of its owner, while in reality being under control of a malicious agency.

From a legal perspective, we would need to address if a computer should be permitted to act as a representative – or indeed if conceptualising the computer as a representative is the most appropriate way to think about such an application. Discussions on autonomous agents and the law have recently debated similar issues in

some depth,¹⁵ with some writers arguing that, in order to form a legally valid offer, an autonomous agent software programme would need recognition as legal person.¹⁶ Others have maintained that it is much more appropriate to think of them just as a new delivery method for the will of its owner.¹⁷ While our system would be semi-autonomous, and capable of dealing with situations that were not foreseen by (and hence not covered by the intent of) its owner, it seems more appropriate, if less spectacular, to think of them not so much as the disembodied mind of the person who trained it, but simply as a new way to record ones wishes and intentions.

Furthermore, it is assumed throughout this paper that establishing the “true” intent of the testator is a desirable outcome under all circumstances. German law at least seems to indicate that this is the case.¹⁸ But as a society, we might actually not want to give the dead too much control over the present and restrict our ability to act as we see fit. After all, the dead outnumber the living by some margin.

The role of the interpreter in this case would be not just to establish the true intent of the deceased but also to find a sustainable compromise between the needs of the present and respect for the past. The more time moves on, the greater that conflict can potentially become. Our proposed system is based on the assumption that a person’s ethical and moral preferences remain stable and that only the set of circumstances to which it is applied changes. This is of course highly unrealistic, as anyone who briefly reflects on the ideals of his or her youth will realise. Also in this respect, our system is eerily reminiscent of the zombies from literature: only living things can learn and change, the undead by contrast are doomed to remain unchanging, incapable of learning and static. Despite these concerns, as a first step to find practical applications for the sophisticated methods and models developed in AI and law research, developing systems to assist in interpreting a person’s will when they are no longer able to speak for themselves is a promising reorientation for research.

¹⁵ IR Kerr, “The Legality of Software-Mediated Transactions” in *Proceedings of IASTED International Conference: Law and Technology* (Calgary: ACTA Press, 2000) 87-96; J Shaheed and J Cunningham, “Agents making moral decisions” presented at the ECAI2008 Workshop on Artificial Intelligence in Games (AIG’08), Patras, 2008, available at http://www.doc.ic.ac.uk/~rjc/jss00_ecai08.pdf (accessed 8 Jul 10).

¹⁶ S Wettig, E Zehender, “A Legal Analysis of Human and Electronic Agents” (2004) 12 *Artificial Intelligence and Law* 111- 135.

¹⁷ G Sartor, “Cognitive Automata and the Law: Electronic Contracting and the Intentionality of Software Agents” (2009) 17 *Artificial Intelligence and Law* 253-290.

¹⁸ R Foer, *Die Regel “Falsa demonstratio non nocet” unter besonderer Berücksichtigung der Testamentsauslegung* (Frankfurt am Main: Lang, 1987).