A CASE-BASED REASONING SYSTEM FOR SUBJECTIVE ASSESSMENT*

William M. Bain

Yale University
Computer Science Department

Abstract

People tend to improve their abilities to reason about situations by amassing experiences in reasoning. Resorting to previous instances of similar situations for guidance is known as *case-based reasoning*. This paper presents JUDGE, a computer model of judges who sentence criminals. The task is viewed as one in which people learn empirically from the process of producing relative assessments of input situations with respect to several concerns, with little external feedback. People can perform such subjective tasks by at least trying to keep their assessments consistent. For assessment tasks, this reasoning style involves comparing a previous similar situation with an input one, and then extracting an assessment for the new input, based on both the assessment previously assigned to the older example, and differences found between them. The system also stores input items to reflect their relationships to situations already contained in memory.

1 Introduction

When people run out of rules to guide them, they reason about problems subjectively. Domains where expert reasoning of this type occurs usually come packaged with a starter kit of traditions, prototypes and precedents; such is the case, for example, with legal reasoning, real estate assessment, various methods of scientific discovery, and art. Beyond such initial guidelines, however, a person often finds himself in uncharted territory.

This paper describes research which has been directed at modelling by computer the behavior of judges who sentence criminals. Our effort has not been to examine sentencing as a representative example of legal reasoning. Instead, we have viewed it as a more generic reasoning task in which people learn empirically from having to produce relative assessments of input situations with respect to several different concerns. Judges receive little external feedback from sentencing that they can directly apply to future cases, so studying this task can help us to understand better the nature of subjectivity, and how to get computer programs to reason subjectively, relying on experience.

Unlike medical tasks, the task of sentencing is not usually considered by judges to be diagnostic. As a result, we have not taken a traditional classification-style approach to modeling judges [3]; instead, our implementation, the JUDGE program, uses a method called *case-based reasoning* [5], [2], which relies on its own experiences to dictate reasons for making certain assessments.

2 Case of the 16-year old offender

To facilitate building a sentencing model, we considered how judges face the task of fashioning sentences by talking with judges who were sitting on the bench in Connecticut at the Superior Court level. An excerpt from a discussion which I had with a judge follows. I described to him briefly an augmented version of a real case which was new to him. This crime was unusual in that it involved a child molestation where the offender was himself only sixteen-years old.

Interviewer: This is a Risk of Injury to a Minor case, against a boy who is sixteen-years old himself...a first offender...with no juvenile record. The details of the crime were that this boy was babysitting for the two kids, and he molested both of them (details given). The kids told their mother and she called the police. Neither of them needed any psychiatric treatment or care for their trauma other than some talking to by their mother–some reassurance.

---

* This paper is a greatly shortened version of [2]; see that source for more extensive discussion. This research was supported in part by the Air Force Office of Scientific Research under contract F49620-82-K-0010 and contract 85-0343.

Judge: If he were presented to me for determination (of youthful offender status), I would feel very strongly against it, because basically, I'd have to... it's very hard to judge, I mean, some people just goof up sexually as they're that age or so, and it's hard to tell with no prior record, I would tend to want to give someone the benefit of the doubt, especially since there is no severe trauma to the victims. I'd do an awful lot of agonizing, but I might give him youthful offender status, and give him three years suspended, with psychiatric treatment throughout the probation. If he were not treated as a youthful offender, then I might well place him on five years, give him a suspended sentence, five years probation with the psychiatric treatment.

Although the judge first began to say that he would not wish to treat the offender as a juvenile, he very abruptly changed his mind. Even so, he formulated two alternative sentences, depending on the ultimate status granted to the offender.

About six weeks after this discussion, I met with the same judge again to discuss a number of cases, including the one above. This time his reaction to the issue of youthful offender status was markedly different, even after such a short period of time.

Interviewer: One of the cases we discussed before dealt with a sixteen-year old boy, who was charged with two counts of Risk of Injury. The facts of this case are ...(*same facts given*).

Judge: What did I say about that? I don't remember what I said. As you talked, the fact situation sounds very familiar.

Interviewer: OK. The boy was sixteen, so one of the things you wanted to know was whether he should, whether he was being treated as a juvenile or an adult.

Judge: Yeah, he should be treated as an adult. He's not a kid. That's a situation where I would find it hard not to consider a suspended sentence and a long, perhaps maximum period of probation with psychiatric treatment, if that's possible. Notwithstanding my feeling that it's going to simply be a waste of time. But who knows, you know? You're giving someone the benefit of the doubt at that age.

Interviewer: How long a period of probation?

Judge: At least five years.

Interviewer: Would you treat him as a juvenile if he were presented to you for determination as a juvenile.

Judge: No I would not. Not at sixteen.

Interviewer: Why not?

Judge: Because I don't think he should be treated as a juvenile at sixteen. I don't think we should be saying, "Oh, he's nothing but a little kid." And besides that, nothing happens in juvenile court. Absolutely nothing. I mean you go through a charade and kids walk out of there laughing. I don't think that's a laughing matter. I mean you can commit murder in our society if you're a juvenile, and get tapped on the wrists; where you gonna send 'em? What are you going to do with them?

2.1 Reasoning from Experience-based Generalizations

The judge's staunch attitude against the offender on the second occasion differed dramatically from his previous position of feeling uncertain, yet beneficent, toward him. Part of the judge's earlier uncertainty was apparent when he proposed two possible sentences for the offender—one for the condition in which he would grant youthful offender status and one if he were to treat him as an adult.

In the second discussion, however, he made no such provisions for doubt. He made the strong statement that this offender was an adult and not a child. He did soften this position a bit by suggesting, as before, that the offender should be given the benefit of the doubt; however, his sentence proposal of "at least five years" was substantially harsher than the tentative sentences which he had mentioned the first time (either 3 years or 5 years). Moreover, his attitude in denying youthful offender status this time could only be described as hostile.

The only explanations which the judge gave for his changed attitude was that in the meantime he had presided in a juvenile case which had been particularly agonizing for him; unfortunately, he gave no details of the case. However, it is noteworthy that the case he heard during that month and a half contributed to his using a different perspective for dealing with juveniles than he had used before, to the point that he reacted to the same set of facts quite differently the second time. From this we note that *the extent to which a judge considers certain features of cases and of offenders to be significant is a function of similar experiences he has in dealing with those features.*

3 The JUDGE Program

The JUDGE system was written to develop sentences for certain crimes, including cases of murder, assault,

and manslaughter, by reasoning about similar situations which the program has previously sentenced. The program compares crimes to ascertain differences in heinousness, which it then maps onto differences in sentence severity. Heinousness is determined relatively, by comparing the causal structures used to represent and interpret crimes.

Some of the 55 cases presented to the system so far were based on real cases of manslaughter and assault, similar to the type of case analyzed in [4]; however, the majority of the cases were constructed to include a variety of actions, results and degrees of justification. JUDGE does not use all of the information available for real cases to determine its sentences. For example, it isn't concerned with whether the offender pleaded guilty, the likelihood that he'll repeat, or the gravity of his prior record. The sentence which the system gives in each case is simplified to apply to only one count of one statutory violation.

The JUDGE system has five stages of operation which lead it to derive sentences. These include the following:

1. An *interpretation phase* for inferring the motives of the actors in the input case and to determine the extent to which each person was justified in acting violently;
2. *Retrieval procedures* for finding similar cases in memory;
3. *Difference analysis procedures* for comparing retrieved cases with the input to help determine how severe the new sentence should be;
4. *Strategy application and modification functions* which both map sentences from old cases onto new cases, and help to insure that differences in sentence severity between crimes corresponds to the relative degrees of heinousness of the crimes;
5. *Generalization capabilities* which enable the system to form sentencing rules when it finds similar cases that require similar sentences.

In addition to being able to generalize rules from processing its input cases, JUDGE can also further modify its own rules. Each of these processes is described in detail in [2].

3.1 Interpretation

The interpretation phase in JUDGE assigns an interpretation to each set of input actions and results. Interpretations provide the system with inferences about the motivations of actors and expand greatly on the representation given initially for each crime; they also serve as indices to cases in memory.

For example, CRIME0, the first case we gave to JUDGE, is an instance of a murder (fictitious), which the system interpreted initially as shown below:

CRIME0 Facts:

First, Ted slashed at Al with a knife one time. (Interpreted as an UNPROVOKED-VIOLATION.) Next, Al slashed back at Ted with a knife one time. (PARITY-SELF-DEFENSE with an ACHIEVED-RESULT.) Finally, Ted stabbed Al with a knife several times. Al died. (ESCALATED-RETALIATION, ACHIEVED-RESULT.)

An UNPROVOKED-VIOLATION means that no other violative actions occurred before the act in question (where Ted slashed at Al); furthermore, this interpretation indicates that Ted's intent to act violently was not justified by any other input knowledge.

The final action in this crime, where Ted stabbed Al to death, was found to be RETALIATION with ESCALATED force and an ACHIEVED-RESULT. To the system, this means that the actor used greater force against his opponent than was previously used against him (escalation); there was at the time of the action no outstanding threat of harm that the actor might have perceived which could justify his action by self-defense (hence, it was retaliatory); and the actor achieved his apparent violative goal. These interpretive structures supply JUDGE with inferences such as that Ted intended to kill Al, thus eliminating other possible inferences (e.g., that the killing was accidental).

3.2 Retrieving Previous Instances from Memory

JUDGE uses the results of its interpretation, including both the interpretive structures themselves and certain of the inferences they provide, to find similar episodes and accompanying strategies in memory for sentencing cases. CRIME0, described above, must be sentenced with initial rules provided to the system, since no other cases are in memory yet. JUDGE's rules assign it a sentence of 40-50 years imprisonment.

In general, when other cases are stored in memory, it is difficult to decide which of the many features of a situation are the most salient and crucial ones to focus on. The system is provided with a set of criteria for determining feature salience derived from the causal structure that it builds for each case during the interpretation phase. This set includes the statute that was

violated, who started the fight, the violative actions and results, and the interpretations assigned to those actions and results by the program. JUDGE looks for crimes in memory which involved these same features. If any crime in memory is found to be similar, using these criteria, the system will begin to consider differences between the input and retrieved crimes.

3.3 Differentiating Cases

Once JUDGE has found a crime from memory similar to the input case, it begins to look in-depth for differences between the two crimes. The system begins by comparing the extent of harm caused by the last actions of each case; then it compares the intentions which led to each action.

CRIME1 Facts:

First, Randy struck Chuck with his fists several times. Next, Chuck struck Randy back with his fists several times. Then, Randy slashed at Chuck with a knife one time. Next, Chuck slashed back at Randy with a knife one time. Finally, Randy stabbed Chuck with a knife several times. Chuck died.

Comparing CRIME1 with CRIME0...
In both crimes, the victim was killed. Not only were both of these outcomes the result of direct intentions, but the actors intended and caused the same amount of harm.

Ted demonstrated an extreme use of force against Al when he acted to stab Al to death in response to having his skin cut in CRIME0. Randy demonstrated an extreme use of force against Chuck when he acted to stab Chuck to death in response to having his skin cut in CRIME1.

Unable to distinguish whether the extreme force used in either crime was worse. The intent of both offenders was to act repeatedly to stab the victim to death. In addition, neither actor's intentions were justified, and both escalated the level of violence about the same degree.

(*At this point,* JUDGE *cannot find a substantial difference between the two cases. As a result, it backs up to compare events that led to these intentions, actions, and results.*)

****** Considering actions of the offenders which led to subsequent victim actions. . . ******

Ted demonstrated an extreme use of force against Al when he acted to slash at Al with a knife in CRIME0. This action was unprovoked. Randy demonstrated an extreme use of force against Chuck when he acted to slash at Chuck with a knife in response to being hit hard in CRIME1.

The magnitude of the extreme force used in CRIME0 was greater than that in CRIME1, and so CRIME0 will be considered worse.

Comparison finished with result that the old crime, CRIME0, is somewhat worse.

It took the program several iterations to determine that CRIME0 was worse than CRIME1. What it found was a difference between the extent to which the offenders escalated the violence in their respective crimes. In general, the system continues to compare the events of two crimes until some notable difference is found or until one or both crimes has been fully scrutinized one event at a time. Notable differences include such features disparities as greater intended harm in one crime, greater caused harm, more justification to respond, extreme force used in one crime, greater relative force, and greater relative escalated force.

3.4 Generalization

JUDGE produces its own rules to generalize certain knowledge about sentencing. General rules are formed only when cases retrieved from memory match a substantial part of the set of features of the input case. The features which commonly describe both situations are extracted and used as indices to store the rule in memory, and a sentence is inherited from the older case. The output below shows this inheritance, along with the set of features common to both CRIME0 and CRIME1 which form the left-hand side of the rule. The sentence given to CRIME1 was 40-50 years—the same as for CRIME0.

FORMING GENERAL SENTENCING RULE:

FOR violation of Murder...
FOR causing result of kill...
FOR using action of stab-knife...
FOR offender starting the fight...
FOR responding to slash-knife harm...
FOR using escalated force and retaliation...
FOR intending to cause the result...

The sentence for this violation will be 40-50 years.

Rules stored in memory can be quickly used to create a sentence for any situation where the rule applies. Thus, in most circumstances the system avoids making an in-depth comparison of input and retrieved cases.

3.5 Rule Differentiation

When JUDGE finds a rule stored with a retrieved case, it tries to apply it if it finds that key features of the rule match with features of the input case. If all of the input case differs from features of the rule, the sentence associated with the rule is modified accordingly. An example of rule modification is shown below.

(*The input crime*-CRIME2:

The offender, Tim, is charged with one count of Murder. Tim was involved in a fight with David, which David started. They traded blows, and after David knocked Tim to the ground, Tim stabbed David several times and killed him.)

Using general rule indexed under CRIME1.

Checking for feature similarity:

FOR victim starting the fight – failure.
General rule in CRIME1 applies offender starting fight.
FOR responding to harm at knock-down level – failure.
General rule in CRIME1 applies to response to slash-knife level.

Handling failure to match on features—features will be added as new indices to rule.

– Reducing the sentence slightly because the eventual victim started the violence in the current situation.

– Increasing the sentence moderately because the offender responded to a lesser degree of violence in the current act than the rule accounts for.

The system stores a new rule in memory with features that reflect the modifications it makes, including those to the sentence. The sentence for the above case changed to 45-50 years.

4 Conclusions

The process described here involves subjective reasoning and learning in a task of relative empirical assessment with little external feedback. This *case-based* reasoning involves comparing a previous similar situation with an input one, and then extracting an assessment for the new case, based both on the assessment previously assigned to the older case, and on the differences found between them. Rules which generalize assessments for particular feature combinations can also be derived, and can refer illustratively back to underlying cases.

The case-based process requires several kinds of knowledge and functional abilities.

1. Previous situations must be kept at hand to compare with input cases.
2. Some notion of similarity must be defined, such that only similar previous situations will be retrieved.
3. A related notion of *significant difference* must be defined so that cases may be compared (and thus differentiated) in a meaningful way.
4. The outcome of comparisons must correlate with the assignment of relative assessments.
5. Finally, the new situation must be stored along with the older ones with respect to its relationship with them, and in such a way that it can be located and used in the future.

These steps are required in general for learning from several examples, as opposed to one-shot or single-instance learning.

Acknowledgements

I am indebted to Roger Schank, Chris Riesbeck, Bob Abelson, Eduard Hovy and Larry Birnbaum for extensive comments and discussion on this work.

References

1. Bain, W. M. *Toward a Model of Subjective Interpretation.* Technical Report 324, Yale University Department of Computer Science, July 1984.
2. Bain, W. M. *Case-based Reasoning: A Computer Model of Subjective Assessment.* PhD thesis, Yale University, 1986.
3. Clancey, W. J. Heuristic Classification. *Artificial Intelligence*, 1985, *27*, 289-350.
4. Pennington, N., and Hastie, R. *Juror Decision Making: Story Structure and Verdict Choice.* Technical Report, Americal Psychological Association, August 1981.
5. Simpson, R. L. *A Computer Model of Case-based Reasoning in Problem-solving: An Investigation in the Domain of Dispute Mediation.* PhD thesis, School of Information and Computer Science, Georgia Institute of Technology, 1985.