Tools for Rapid Prototyping of Legal Case Based Reasoning

Trevor BENCH-CAPON[1], Thomas GORDON[2]
[1] *Department of Computer Science, University of Liverpool, UK*
[2] *Fraunhofer FOKUS, Berlin, Germany*

Abstract. In this paper we discuss tools for rapid prototyping of legal CBR. We describe how a recent highly quantitative analysis can be realised using a spreadsheet, while a more nuanced approach at a finer level of granularity can be prototyped with the web service Carneades, which displays its results as an argument graph.

Keywords. case based reasoning, dimensions, Carneades.

1. Introduction

Reasoning with legal cases has been a central focus of AI and Law since its beginnings. Perhaps the central example system is CATO [3], which developed from HYPO [4]. Since them there have been theoretical developments, mostly concerned with finding a logic of precedent, [19], [17], [21], and practical modelling the cases, and producing implementations (e.g. [5], [2]). There have also been some empirical studies of the effective of different approaches, such as [11]. Investigation of approaches, and models of particular bodies of case law, could be greatly aided by means to rapidly prototype an implementation, to validate and compare approaches, and to refine and compare models and analyses. In this paper we consider two tools which could be used: a standard spreadsheet (Excel) and the web service Carneades [14]. We will take as our starting point the domain analysis in [1] which extended the analysis of [2]) to allow for consideration of degrees of presence through the use of dimensions. The domain modelled was that of the wild animals cases introduced in [8], and which, together with the case of *Popov v Hayashi* introduced in [23], have subsequently been widely discussed and represented.

2. An Analysis of the Domain as Dimensions

In [2] the cases of interest were analysed in terms of factors. The analysis yielded a CATO-style factor hierarchy, which, when augmented by truth conditions can be used as an Abstract Dialectical Framework (ADF) [10]. In [2], the ADF was rewritten as a *2-regular ADF*, in which every non-leaf node has exactly two children (we will refer hereafter to a node in such an ADF as the parent, and its two children as *elder* and *younger*). One advantage of the 2-regular ADF is that the truth conditions take one of three forms: *parent if elder and younger*, *parent if elder or younger*, and *parent if elder and not younger*.

In that analysis the leaf nodes were factors, and so were considered either present or absent. The truth conditions can be used to determine whether the parent, representing a more abstract factor, is present or absent. Towards the top of the hierarchy we encounter issues: such parents are found for either plaintiff or defendant and related through a logical model (as in IBP [11]). The analysis was intended to be realised as a Prolog program. It was, however, recognised that in many circumstances the distinction between presence and absence is not clear cut, and that representing degrees of presence was desirable.

An analysis in terms of dimensions was offered in [1]. The idea here is that factors represent points on (or, more generally, sub-ranges of) dimensions. As we move from extreme pro-plaintiff to extreme pro-defendant the plaintiff is less and less favoured by the dimension. At some point on the dimension, the plaintiff will cease to be favoured entirely, and at some point (which may be the same point) the defendant will start to be favoured, and favoured increasingly thereafter. Thus factors, or the facts which lead to the ascription of a factor, can be mapped to a point on a dimension. If we treat the dimension as a numeric scale (e.g extreme pro-plainiff = 10, extreme pro-defendant = 0) we can thus replace the factor with a number, representing the strength of reason it affords.

Nine dimensions were identified in [1]. The dimensions are: LandOwnership (LO), ApplicableConvention (AC), ClosenessOfPursuit (COP),QuarryValue (QV), QuarryLandConnection (QLC) NatureOfAct (NOA), PlaintiffMotive (PM), DefendantMotive (DM) and DefendantRole (DR). Five cases ((*Pierson v. Post*, *Keeble v. Hickergill*, *Young v. Hitchens*, *Ghen v. Rich*) and *Popov v Hayashi*) were represented as shown in Table 1. Each of the points is associated with a number representing its position of the dimension (see [1]).

Table 1. Dimensions from [1]

	Pierson	*Keeble*	*Young*	*Ghen*	*Popov*
LandOwnership	Unowned	P-Freehold	Unowned	Unowned	Other Owned
Convention	Social Preference	Nothing	Nothing	Full Possession	Informal Right
ClosenessOfPursuit	In Hot Pursuit	Certain To Capture	Certain To Capture	Captured	In Hot Pursuit
QuarryValue	Social	Market	Market	Market	Market
QuarryConnection	Regular	Frequent	Resident	Resident	Frequent
NatureOfAct	Impolite	Nuisance	Impolite	ActOk	Violently Illegal
PMotive	Sport	Commerce	Commerce	Commerce	Gain
DMotive	Impulse	Malice	Commerce	Commerce	Gain
DefendantRole	Solely Responsible	Solely Responsible	Solely Responsible	Innocent	Innocent

Each case can now be represented as a vector of nine numbers: ⟨LO, AC, COP, QV, QLC, NOA, PM, DM, DR⟩. We can determine the values of abstract factors by propagating these numbers up the tree using the truth conditions, but interpreting *and* and *or* as in fuzzy logic [24] (i.e min and max respectively). The third type of truth condition, parent if elder and not younger, becomes if *elder > younger then elder else 0*. As we approach issues, we may need to introduce thresholds. For example in [6], it was argued that the automobile exception to the fourth amendment of the US constitution involved consideration of two issues: whether the exigency of the situation exceeded a certain threshold, and whether the reasonable expectations of privacy fell short of a given threshold. Three thresholds are required for our cases. First, to determine how close the pursuit needs to be to give possession to the plaintiff: several authorities were cited in *Pierson v Post* with different views as to where the line should be drawn. Second we need a threshold to determine whether the land belongs to the plaintiff in a way which gives possession of the animals on it to the plaintiff (e.g. if the land is rented, do the animals belong to the tenant or the landlords?). Third a threshold for where the law should intervene (e.g. commerce only, any kind of valuable quarry, or even in social matters). The thresholds are represented by setting appropriate flags.

A custom GUI for the application was proposed in [1], but often it will be desirable to use available tools to realise the analysis, either instead of (or preparatory to) writing dedicated code. In the next sections will be discuss the implementation of the anaysis with a spreadsheet. Note that although we have chosen to model an analysis in terms of dimensions an alysis in terms of factors as in [] could be implemented in exactly the same way. The only differences would be that the vector would be considerably longer (27 factors rather than 9 dimensions) and that it would comprise only 10s and 0s, since factors are either present or absent and no intermediate degrees are allowed. .

3. Spreadsheet Implementation

The spreadsheet[1] has four areas: a *data* area to hold the case vectors, a *flags* area, an *abstract factors* area, used to propagate the initial values, and an *issues* area to resolve the issues. In the data area each case has a column and there is an additional column ("Current") for the case being considered. The data for an individual case can be pasted into the Current column to make that the case under consideration. The data area is shown in Table 2. The current case in Table 2 (and Tables 3,4 and 5 is *Pierson v Post*).

Similarly the flags area show the available settings and has a "current" column to hold the selected flags. The flags may simply be fixed values, or may be calculated: for example, the Livingston flag depends of the value of the quarry and: if the quarry value is greater than or equal to 7. the flag is 5, else it is 8. Thus the Livingston cell has an associated formula $= IF(G5 >= 7, 5, 8)$.

The heart of the model is in the abstract factors area, shown in Table 4. Here each cell represents one or two arguments, justifying the content of the cell. Thus

[1] The spreadsheet can be downloaded from www.csc.liv.ac.uk/t̃bc/FTP/popovDimensions.xlsx

Table 2. Data Area with *Pierson* as Current case

Dimensions	pierson	keeble	young	ghen	popov	current
LO	5	8	5	5	4	5
AC	3	0	0	10	5	3
COP	5	7	7	8	7	5
QV	7	10	10	10	10	7
QLC	8	8	10	10	10	8
NOA	3	5	3	0	10	3
PM	6	10	10	10	8	6
DM	2	0	10	8	8	2
DR	0	0	0	6	10	0

Table 3. Flags Area with *Pierson* as Current case

Flags							
Authority	Justinian	Putendorf	Barbeyrac	Livingston	Current		
	10	8	7	5	10	RTP	Current
Ownership	Strict	Standard				3	10
	10	7			7		
Litigation	Business	Gain	Social				10
	10	8	6		10		

Table 4. Abstract Factors Area with *Pierson* as Current case

Abstract Factors						
FC	ER	FP	DB	PPI	RTP	Current
0	6	5	3	5	3	10
ByC	ByL	CR	CQ			
3	0	3	5			10

the presence of FP (favoured pursuer) can be established through the presence of either LO (a person has the right to pursue animals on his own land) or COP (a person may establish a right to pursue by coming sufficiently close). Thus the degree of presence of FP will be LO fuzzy-or COP, i.e *=MAX(G2,G4)*. ER (exclusive right) depends on the motives of the parties involved: if the plaintiff has the better motive, then ER will take that value, otherwise it will be zero. Thus the cell is filled by two arguments:

- ER = PM because plaintiff motive is better than defendant motive (i.e. PM >DM.
- ER = 0 because plaintiff motive is no better than defendant motive (i.e. PM =<DM.

This can be realised using an IF...THEN...ELSE formula (i.e *=IF(G8 > B14,G8,0)*).

Finally we reach the issue area, shown in Table 5. We have three issues: ownership, whether to find for the plaintiff and whether to find for the defendant. Asn ca be seen from Table 5 in *Pierson v Post* we have FFP = 5 and FFD = 8.

Table 5. Issues Area with *Pierson* as Current case

Issues				
	Ownership	FFP	FFD	FFP
	7	7	6	5

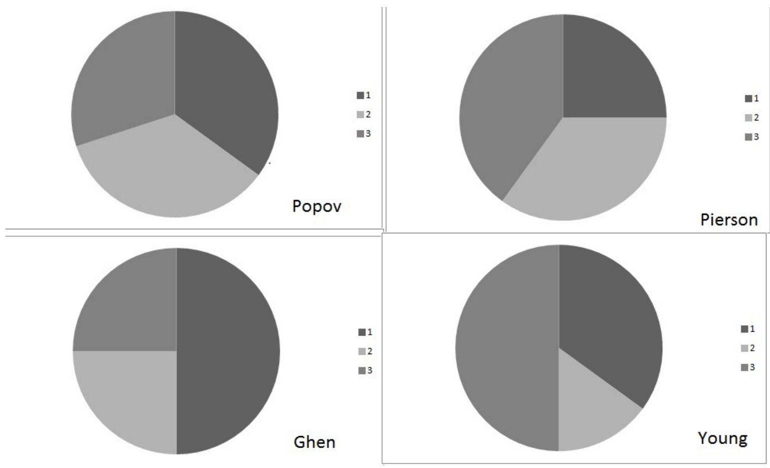

Figure 1. Pie Charts showing FFP, Uncertainty and FFD for three cases

For issues, the numbers represent to the degree to which the plaintiff is favoured. It is then up to the user to decide whether the plaintiff is sufficiently favoured: for our cases, 8 seems to be the right cut-off (FFP = 5 for *Pierson*, 8 for *Keeble*, 7 for *Young*, 10 for *Ghen* and 7 for *Popov*). FFD is a separate issue, since in *Popov* the judge wanted to find neither for Popov not for Hayashi but to effect a compromise. FFD is 8 for *Pierson*, 5 for *Keeble*, 10 for *Young*, 5 for *Ghen* and 6 for *Popov*. Again 8 would be an acceptable cut-off.

It is now possible to use various chart facilities to visualise aspects. For example we could use a bar chart of the case vector in current to visualise the "shape" of a case. Or, as shown in Figure 1, we could use pie charts to show the extents to which the parties are favoured by the facts, and the remaining uncertainty.

The spreadsheet provides a systematic way to rapidly prototype an implementation of a dimensional analysis as found in [1]. As such it would make an excellent environment for the exploration of the effects of different flags and refinement of the rule set. Of course, it does require proficiency in the use of spreadsheets, both to develop and to use the model. Thus it is probably a tool for the software professional rather than end users.

4. The Carneades Argumentation System

The Carneades argumentation system provides an integrated set of software tools for argument (re)construction, evaluation, mapping and interchange. Carneades has been in continuous development since 2006 and several versions have been developed over the years.

The first version was based on the computational model of structured argument in [16]. Its main innovations were its support for multiple burdens of proof, based on prior work in [13,15] and its method of modelling different kinds of critical questions of argumentation schemes. However, this computational model of structured argument is limited to acyclic argument graphs and did not provide sufficient support for *conductive arguments* [9] which accrue, aggregate and balance evidence and reasons pro and con alternative positions (options) of an issue . Moreover, statements in argument graphs of this first version of Carneades modelled Boolean issues about whether or not a claimed proposition or its complement (negation) are acceptable (justified), or neither, given the chosen proof standard. The model did not therefore provide good support for arguing about issues which have more then two options or practical reasoning about which action to perform, when more than one action is being considered.

The latest version of Carneades, 4.x, overcomes these limitations. The limitation to cycle-free argument graphs is overcome using fixpoint semantics, inspired by Abstract Argumentation Frameworks [12] and Abstract Dialectical Frameworks [10], but without using either of these as models of abstract argumenation directly, as "middleware". Better support for issues with more than one option is provided by integrating ideas from Issue-Based Information Systems [22,15] and Multi-criteria Decision Analysis [18].

In order to model conductive arguments the new model makes the weight of arguments a function of, and thus dependent on, the labels of statements in the argument graph. Since the label (in, out, undecided) of a statement is also dependent, recursively, on the weights of the arguments pro and con the statement, it is not sufficient to weigh or order the arguments *before* they are evaluated, as in done in ASPIC+ [20].

Semi-formal and completely formal specifications of the new model, called the Carneades Argument Evaluation Structures (CAES) are available from the second author butare not yet published. The Isabelle/HOL proof assistant is being used for the the complete formalization, which allows all proofs to be automatically checked.[2]. The formalizations are available by request from the second author.

Carneades 4.1 is available online at `http://carneades.fokus.fraunhofer.de` and provides a way to import structured argument graphs in a number of formats. The "native" format of Carneades 4.x, supporting all the features of the new CAES model of structured argument, is based on YAML, a human readable data serialization format.[3] Argument graphs are evaluated and then visualized in argument maps. Various graphic formats for outputting the argument maps are available, including dot, PNG, SVG, and GraphML. Like previous versions of Carneades, the software is open source, available for downloading on Github.[4]

Carneades 4.x does not yet include a language for defining functions for weighing arguments, such as weighted sums of the values of criteria or dimensions for particular kinds of issues. The formal model allows the weights to depend on the labelling of statements and the properties of any arguments in the argument graph. In particular, the weight of an argument can depend on the weights of

[2] `https://isabelle.in.tum.de/`

[3] https://en.wikipedia.org/wiki/YAML

[4] `https://github.com/carneades/carneades-4`

other arguments which establish its premises. Some weighing functions, built in the system, are available to choose from, but it is not yet possible to configure or customize the system to define new weighing functions without modifying and re-compiling the source code. And unlike previous versions of Carneades, it also does not yet include a language or inference engine for argumentation schemes. These are planned for versions 4.2 and 4.3, respectively. See the roadmap on Github for details. Carneades is currently being expressly developed to support case-based reasoning, in the European EAGLE project, where the goal is to develop a system enabling clerks in public administrations throughout Europe to share arguments about the application of open-textured concepts, such as "undue hardship", in concrete cases.[5]

4.1. Example

In the wild animals cases there is an issue as to whether the plaintiff has a better motive than the defendant (otherwise it can be seen as fair competition) and if so, whether the motive is of sufficient importance for the law to intervene. In the approach of [1] PM and DM are dimensions and so the motivations are assigned a number indicating a point on the respective dimensions. Who has the better motive then is just a matter of simple numeric comparison. Similarly the importance is established by comparing the winning number with a threshold indicating the extent to which the law is prepared to intervene. But this is rather a crude representation.

Motives in fact have a number of different aspects. For example, a person may be motivated by gain, by doing good for society and may have good or bad intentions. In our cases Keeble, Young and Ghen were all motivated by money, but their actions were also socially useful since they were bringing goods to the marketplace. Their intentions were also ok. Hitchens and Rich also acted with good intentions, but Hickergill acted from malice. In the Popov case, money was the motive, and the intentions were ok, but there is no social value in the souvenir baseball market.The motives in Pierson are disputed: was Post acting from pleasure or a desire to rid the countryside of vermin? Was Pierson acting from a dislike of land owning gentry, as suggested in [7]? The motivations are shown in Table 6.

In Carneades we are able to explore these nuances. By identifying the different aspects of the motivations, and using a cumulative argument scheme we can use which of the various aspects apply to the participants to determine who has the better motive. Thus in *Keeble* and *Pierson2* the plaintiff has the better motive. We can then use the proof standard to decide whether the motive is sufficient. Applying a weaker standard, as advocated by Livingston in *Pierson*, we find the difference sufficient, whereas the stronger standard (advocated by Tompkins in *Pierson* to discourage excessive litigation), will find the difference insufficient.

The input to Carneades takes the form a text (.yml) file specifying:

- **Statements**. This section declares the statements to be used in the model and a text annotation that will appear in the output. For example

[5] http://www.eagle-learning.eu/

Table 6. Motives in the Cases

Plaintiff	Money Value	Social Value	Intentions Ok	Defendant	Money Value	Social Value	Intentions Ok
PvP (1)	No	No	Yes		No	No	Yes
PvP (2)	No	Yes	Yes		No	No	Yes
PvP (3)	No	No	Yes		No	No	No
KvH	Yes	Yes	Yes		No	No	No
YvH	Yes	Yes	Yes		Yes	Yes	Yes
GvR	Yes	Yes	Yes		Yes	Yes	Yes
PvH	Yes	No	Yes		Yes	No	Yes

```
PBM: Plaintiff has sufficiently better motive
DBM: Defendanthas sufficiently better motive
```

- **Issues** These are two or more propositions which cannot be true together. In our example the issue is whether either PBM or DBM (or neither) should be taken as true. Here we also specify the standard to apply, to indicate the extent of difference demanded..
- **Arguments**.The various rules are represented as arguments. For example selecting a case will determine the values for the various aspects of motive. For example, the plaintiff's intentions in *Pierson1* are ok:

  ```
  a1:
        conclusion: PI
        premises: [Pierson1]
  ```

  The argument for PBM, however, is specified as *cumulative*, so that it has a strength dependent on how many of its premises are satisfied, permitting comparison with the strength of the argument for DBM, so that the standard for the issue can be applied.

  ```
  a0:
        scheme: cumulative
        conclusion: PBM
        premises: [PM, PS, PI]
  ```

- **Assumptions**. This allows an assertion of the case to be considered.

Once the file modelling the domain has been produced, it can be submitted to the web service (http://carneades.fokus.fraunhofer.de/eval-form) and evaluated to produce the corresponding graph. Argument graphs for Pierson2 with the two different standards are shown in Figure 2. Note that the plaintiff has a sufficiently better motive in one case, but not the other. Figure 3 shows the graphs for *Pierson1* with the weak standard and *Keeble*. The same graph for *Keeble* (except for the issue label) is produced whichever standard is used.

By varying the case using the assumption line, and the standard in the issue, variations in cases and standards can be readily and visually explored.

5. Concluding Remarks

Whether exploring an approach to modelling reasoning with legal cases, or constructing a domain analysis it is enormously useful to be able to execute a corresponding implementation in order to see whether the behaviour is as desired and expected, and to explore refinements if this is not so. This requires the availability of tools which enable rapid prototypes to be produced without undue programming effort.

In this paper we have shown how systems based on hierarchies of concepts (such as the abstract factor hierarchy of [3] or the ADFs in [2] and [1]) in which truth values are propagated up a tree from leaf nodes to root can be straightforwardly expressed in a spreadsheet. We have also shown how we can use a web service such as Carneades to explore more nuanced approaches which are not susceptible to representing cases as vectors of numbers.

Acknowledgments

We would like to thank Latifa Al-Abdulkarim and Katie Atkinson for their role in producing the dimensional analysis of the cases used in this paper.

References

[1] L. Al-Abdulkarim, K. Atkinson, and T. Bench-Capon. Adding dimensions and facts to adf representation of legal cases. In *Submitted*, 2015.

[2] L. Al-Abdulkarim, K. Atkinson, and T. Bench-Capon. Factors, issues and values: Revisiting reasoning with cases. In *Proceedings of the Fifteenth International Conference on Artificial Intelligence and Law*, pages 3–12, 2015.

[3] V. Aleven. *Teaching case-based argumentation through a model and examples*. PhD thesis, University of Pittsburgh, 1997.

[4] K. Ashley. *Modelling Legal Argument: Reasoning with Cases and Hypotheticals*. Bradford Books/MIT Press, Cambridge, MA, 1990.

[5] K. Atkinson. Introduction to special issue on modelling Popov v. Hayashi. *Artificial Intelligence and Law*, 20(1):1–14, 2012.

[6] Trevor Bench-Capon and Henry Prakken. Using argument schemes for hypothetical reasoning in law. *Artificial Intelligence and Law*, 18(2):153–174, 2010.

[7] Bethany R Berger. It's not about the fox: The untold history of pierson v. post. *Duke Law Journal*, pages 1089–1143, 2006.

[8] D. Berman and C. Hafner. Representing teleological structure in case-based legal reasoning: The missing link. In *Proceedings of the Fourth International Conference on Artificial intelligence and Law*, pages 50–59, 1993.

[9] J. Anthony Blair and Ralph H. Johnson, editors. *Conductive Argument: An Overlooked Type of Defeasible Reasoning*. College Publications, 2011.

[10] G. Brewka and S. Woltran. Abstract dialectical frameworks. In *Principles of Knowledge Representation and Reasoning: Proceedings of the Twelfth International Conference*, 2010.

[11] S. Brüninghaus and K. Ashley. Predicting outcomes of case-based legal arguments. In *9th International Conference on Artificial Intelligence and Law*, pages 233–242, 2003.

[12] Phan Minh Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence*, 77(2):321–357, 1995.

[13] Arthur M. Farley and Kathleen Freeman. Burden of Proof in Legal Argumentation. In *Fifth International Conference on Artificial Intelligence and Law (ICAIL-95)*, pages 156–164, New York, NY, USA, 1995. ACM Press.

[14] Thomas F. Gordon. Introducing the carneades web application. In *Proceedings of the 13th International Conference on Artificial Intelligence and Law*, pages 243–244, 2013.

[15] Thomas F Gordon and Nikos Karacapilidis. The Zeno argumentation framework. In *Proceedings of the Sixth International Conference on Artificial Intelligence and Law*, pages 10–18, Melbourne, Australia, 1997. ACM Press.

[16] Thomas F. Gordon, Henry Prakken, and Douglas Walton. The Carneades Model of Argument and Burden of Proof. *Artificial Intelligence*, 171(10–11):875–896, 2007.

[17] J. Horty and T. Bench-Capon. A factor-based definition of precedential constraint. *Artif. Intell. Law*, 20(2):181–214, 2012.

[18] Marc Lauritsen. On balance. In *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Law*, ICAIL '13, pages 83–91, New York, NY, USA, 2013. ACM.

[19] H. Prakken and G. Sartor. Modelling reasoning with precedents in a formal dialogue game. *Artif. Intell. Law*, 6(2-4):231–287, 1998.

[20] Henry Prakken. An abstract framework for argumentation with structured arguments. *Argument & Computation*, 1:93–124, 2010.

[21] A. Rigoni. An improved factor based approach to precedential constraint. *Artif. Intell. Law*, 23(2):133–160, 2015.

[22] HorstW.J. Rittel and MelvinM. Webber. Dilemmas in a general theory of planning. *Policy Sciences*, 4(2):155–169, 1973.

[23] A. Wyner, T.Bench-Capon, and K. Atkinson. Arguments, values and baseballs: Representation of popov v. hayashi. In *Proceedings of Jurix 2007*, pages 151–160, 2007.

[24] L. A Zadeh. Fuzzy sets. *Information and control*, 8(3):338–353, 1965.

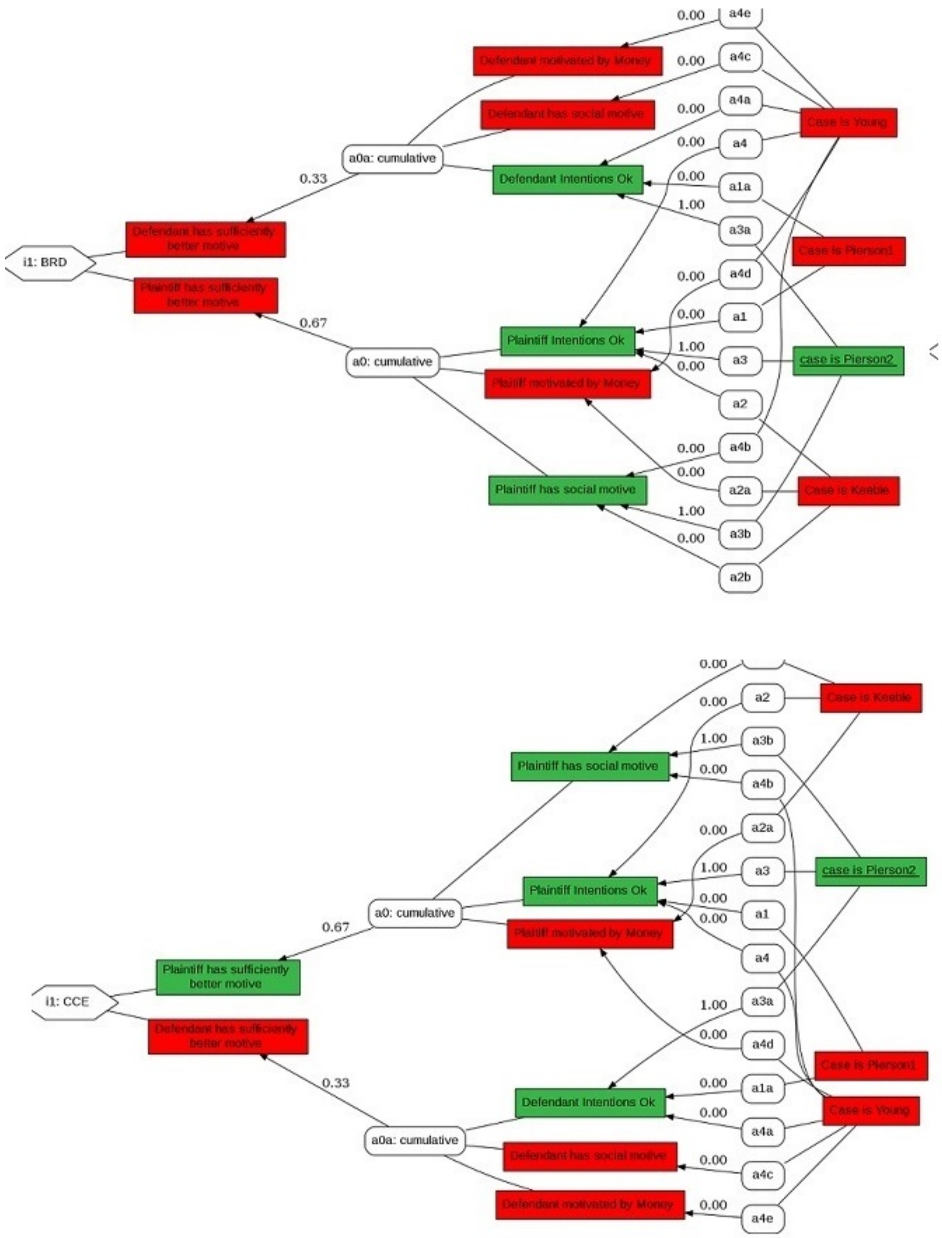

Figure 2. Carneades graphs for *Pierson2* with strong (above) and weaker (below) standards

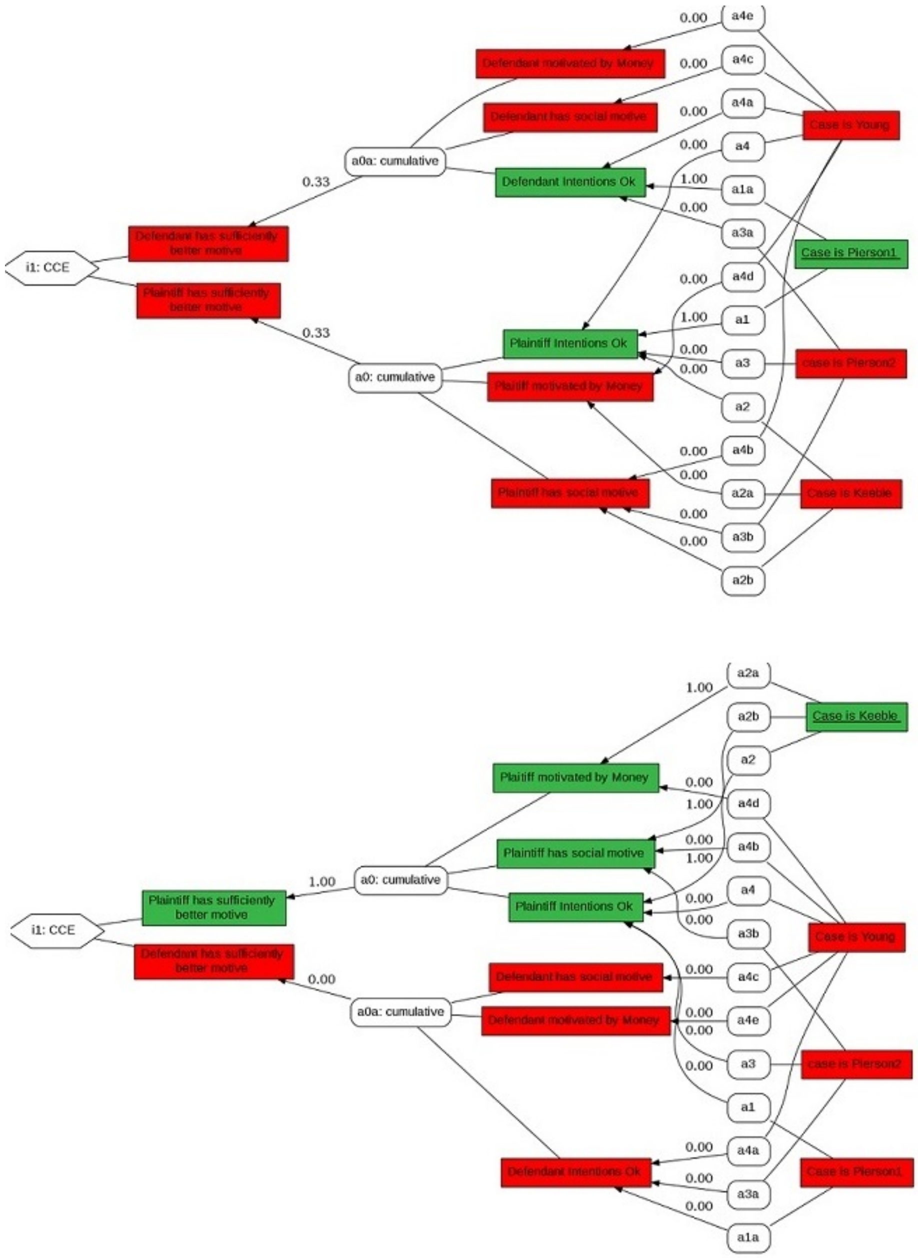

Figure 3. Carneades graph for *Pierson1* with weaker standard (above) and *Keeble* (below)