Hierarchical *a Fortiori* Reasoning with Dimensions

Wijnand VAN WOERKOM [a,1], Davide GROSSI [b,c,d], Henry PRAKKEN [a] and Bart VERHEIJ [b]

[a] *Department of Information and Computing Sciences, Utrecht University, The Netherlands*

[b] *Bernoulli Institute for Mathematics, Computer Science and Artificial Intelligence, University of Groningen, The Netherlands*

[c] *Amsterdam Center for Law and Economics, University of Amsterdam, The Netherlands*

[d] *Institute for Logic, Language and Computation, University of Amsterdam, The Netherlands*

Abstract. In recent years, a model of a fortiori argumentation, developed to describe legal reasoning based on precedent, has been successfully applied in the field of artificial intelligence to improve interpretability of data-driven decision systems. In order to make this model more broadly applicable for this purpose, work has been done to expand the knowledge representation on the basis of which it functions, as the original model accommodates only binary propositional information. In particular, two separate expansions of the original model emerged; one which accounts for non-binary input information, and a second which accommodates hierarchically structured reasoning. In the present work we unify these expansions to a single model, incorporating both dimensional and hierarchical information.

Keywords. a fortiori reasoning, explainable artificial intelligence, precedential constraint, dimensions, hierarchy

1. Introduction

In [1] Horty introduced a formal model of *a fortiori* reasoning, which he called the *result model* (RM), for describing the type of reasoning performed by a court when citing past decisions called precedent cases. The model describes when a new decision is, or is not, consistent with respect to the precedent. In other words, it describes the way in which a set of precedent constrains future decision-making. The RM works on the basis of a knowledge representation using *factors*—legally relevant fact patterns that are assumed to favor either a decision for the plaintiff or the defendant of the case. Two shortcomings of the RM have been pointed out on the basis of this form of knowledge representation.

The first is that the factors used in the model are *binary*, while in practice there may be relevant legal information which is encoded as, for example, a natural number. Horty presented an extended version of the RM in [2], which uses a knowledge representation

[1] Corresponding author; e-mail: w.k.vanwoerkom@uu.nl.

that allows for dimensional information to be encoded. For the purpose of comparison we will refer to this modified version of the RM as the *dimensional result model* (DRM).

The second shortcoming (as already pointed out by Horty himself in [1]) is that in practice factors often have a hierarchical structure, which the RM does not take into account. The court uses this hierarchical structure to move from low-level factors through a series of intermediate concepts, called *abstract* factors, before arriving at some final conclusion. Earlier work on formal models of precedential constraint by Roth and Verheij [3,4] did include such hierarchical information. Building on those ideas, the RM was extended in [5] to operate on a knowledge representation using hierarchical structure. We will refer to this model as the *hierarchical result model* (HRM). We note that Horty's *reason* model of precedential constraint—also introduced in [1], and which builds upon the RM by adding a notion of reason for the decision of the court—was also recently extended to a model including hierarchical information in [6].

In the present work we unify the HRM and DRM into a single model which accounts for both dimensional information and hierarchical structure, and which we will at present refer to as the *dimensional hierarchical result model* (DHRM). This model subsumes the HRM and the DRM in the same way they subsume the RM—for example, every instance of the RM can also be considered an instance of the HRM, or of the DRM.

Our motivation for this work stems from recent applications of these a fortiori reasoning models to the improvement of interpretability of AI systems. For example, starting with work by Prakken and Ratsma in [7], the DRM has been used as the basis of a post-hoc explanation method for black-box AI systems. This work has since been continued in e.g. [8,9,10]. Additionally, the DRM is used as part of an interpretable decision support system by the Dutch national police force [11]. To aid in this application, notions of *justifiability* and *relevance* were recently added to the theory in [12]. The limits of the knowledge representations used by the DRM and HRM in turn limit the scope of these applications, which is why we presently propose the DHRM extension.

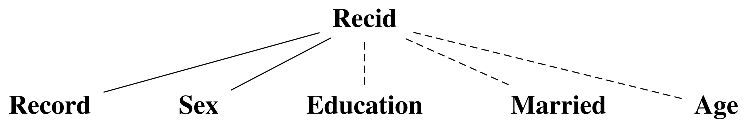
Following [10], we motivate and illustrate the applicability of the various a fortiori models discussed in this work by use of a running example from the legal domain of criminal sentencing. More specifically, we consider the tasks of judging recidivism risk and granting bail. We show that a fortiori reasoning is applicable to these tasks, and requires the use of a knowledge representation incorporating both dimensional and hierarchical information. Criminal sentencing is a highly relevant domain for our purposes—decisions surrounding criminal sentences have the potential to greatly affect peoples' lives, and AI is increasingly being used to compliment or even replace human decision-making for these tasks. For example, in recent years there has been much discussion surrounding allegations put forth in [13] that the COMPAS system, widely used in the United States for automatic recidivism risk assessment, was making racially biased decisions.

In Sections 2 through 5 we sequentially describe the RM, HRM, DRM, and DHRM. Each of these models consists of roughly the same two components: some form of knowledge representation which is used to represent *fact situations* and *cases*, and a notion of the way in which a set of such cases *constrains* decision-making about new, unseen fact situations. Each of the aforementioned sections is structured according to this fact—they begin with an intuitive explanation of the model on the basis of our running example, and then give formal definitions of their knowledge representation and the associated notion of constraint. After having presented the models, we end in Section 6 with some closing thoughts.

2. Result Model

2.1. An Example of Factors

We illustrate the various kinds of models discussed in this work through a running example on the criminal sentencing domain. In this case, we consider a judgement of whether a convict is at low or high risk of recidivism—a primary task of the COMPAS software. Much research has been done on the factors influencing recidivism, see e.g. [14] for a recent meta-study. Below is a graphical representation of a number of such factors:



(1)

The factors in the bottom row respectively indicate whether the defendant has a criminal record, is male, has a high school diploma, is married, and is over the age of 21. A solid line between a factor and the **Recid** node indicates that the presence of that factor suggests a higher risk of recidivism, while a dotted line indicates that its presence suggests a lower risk. For instance, having a criminal record indicates a higher risk, while being married indicates a lower risk. Now, suppose a 30 year-old unmarried male defendant with a pre-existing criminal record and no high school diploma was judged to be at high risk of recidivism. Given our assumption that older people tend to recidivate less, it follows *a fortiori* that a defendant who is on all accounts similar, but is 20 years old instead of 30, should also be judged to be at high risk of recidivism.

2.2. Knowledge Representation

A *factor* is a propositional variable, i.e. a variable which is either true (denoted **t**) or false (denoted **f**). We denote factors using lowercase letters p, q, r etc. The domain is modeled by a finite set of factors F. A *fact situation* is a valuation of F, i.e. a function $X : F \to \{\mathbf{t}, \mathbf{f}\}$ assigning true or false to every factor in F. We use upper case letters X, Y, Z etc. to denote fact situations, and write $X \vDash p$ for $X(p) = \mathbf{t}$ and $X \vDash \neg p$ for $X(p) = \mathbf{f}$.

Cases are decided for either of two *sides*; the plaintiff, denoted by π; or the defendant, denoted by δ. Each factor $p \in F$ has a *preference* for exactly one of the two sides, which is modeled by two sets $\text{Pro}(\pi)$, $\text{Pro}(\delta)$. These sets should constitute a disjoint union of F, meaning $\text{Pro}(\pi) \cup \text{Pro}(\delta) = F$ and $\text{Pro}(\pi) \cap \text{Pro}(\delta) = \emptyset$. If a factor is pro-$\pi$ (δ) it is con-δ (π) so we define $\text{Con}(\delta) = \text{Pro}(\pi)$ and $\text{Con}(\pi) = \text{Pro}(\delta)$. A *case* is a pair (X, s) with X a fact situation and s a side; a *case base* CB is a finite set of cases.

2.3. Constraint

The idea behind the RM is that a decision of a fact situation X for a side s constitutes a balancing of the pro-s factors in X against the con-s factors in X. The support that factors provide for an outcome is defeasible and unquantified, which makes it difficult to weigh sets of pros against sets of cons. However, once a set of pros was deemed to outweigh a set of cons, any superset of the set of pros should also outweigh any subset of the set of cons. This intuition is formalized by the following definition.

Definition 1. *The decision of a fact situation X for a side $s \in \{\pi, \delta\}$ is* forced *by a case base CB, denoted $CB, X \vDash s$, if there exists a case $(Y, s) \in CB$ such that:*

- *for all $p \in \mathrm{Pro}(s)$: if $Y \vDash p$ then $X \vDash p$, and*
- *for all $p \in \mathrm{Con}(s)$: if $X \vDash p$ then $Y \vDash p$.*

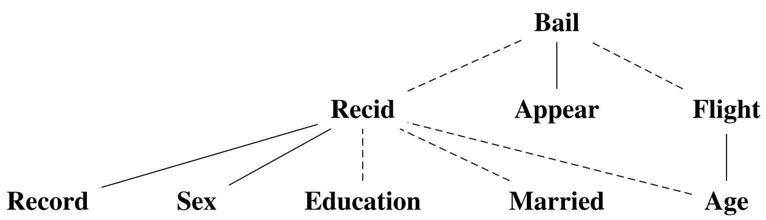
Note carefully that the RM was not designed as a method for weighing pros and cons against each other. Instead, it normatively prescribes a principle of what it means for such a method to act in accordance to the a fortiori principle and a set of previous cases.

3. Hierarchical Result Model

3.1. An Example of a Factor Hierarchy

A downstream purpose of recidivism risk assessment, and an example of the purposes for which the COMPAS program is used in practice, is to determine whether a defendant should be released on bail. Bail is a sum of money that the defendant must pay to the court as a guarantee that they will appear at their trial—if the defendant does not appear, the bail is forfeited. The decision to grant bail, like recidivism risk, is influenced by several factors; e.g. a defendant with a high risk of flight is less likely to be granted bail, while one with a history of appearing to court is more likely to be granted bail.

In other words, determining bail is a domain to which the result model can be applied, but this time one of the input factors—risk of recidivism—can itself be determined on the basis of a fortiori reasoning. This situation is called a *factor hierarchy* in the AI & law literature, a concept which was first used in the CATO program [15]. We expand our example from Section 2.1, graph (1), to a hierarchy including a bail decision:



(2)

The **Bail** node corresponds to a decision to grant bail. The links are either solid or dotted, which carries the same meaning as it did in our earlier example. For instance, the dotted line from **Recid** to **Bail** indicates that high risk of recidivism suggests bail should be denied. Two other factors are added: **Appear**, which stands for a low or high chance of appearing at the next trial, and **Flight**, which stands for a low or high risk of fleeing. This means that in this example we assume that older people are more likely to flee before trial than younger people. Whether this assumption holds in practice is debatable—it is added primarily to exemplify that factors can influence multiple higher level factors.

3.2. Knowledge Representation

We again assume the domain is modeled by a set F of factors. However, we now consider the case outcome to be just one of the factors. Therefore, to be able to distinguish between fact situations that have been decided for a side and those that have not, we also drop the requirement that a fact situation is defined on all factors. More specifically, a *fact situation* is now a valuation of a subset of F. The *domain* of a fact situation X, i.e. the factors to which it assigns true or false, is denoted $\mathrm{dom}(X)$.

A *factor hierarchy* is a set of factors F with a binary relation H on F satisfying:

1. the transitive closure of H is irreflexive, and
2. it is equal to a disjoint union of two relations Pro and Con.

A hierarchy H is called *flat* if all factors in it are H-minimal or H-maximal. An H-minimal factor is called *base-level*. A factor that is not base-level is called *abstract*, and the set of abstract factors is denoted by A. Factors are assumed to support or oppose each other in hierarchical fashion, as indicated by the relations Pro and Con. When $\mathsf{H}(p,q)$ holds between factors p and q then either $\mathsf{Pro}(p,q)$, which means p is a pro-q factor, or $\mathsf{Con}(p,q)$, which means p is a con-q factor.

A hierarchy can have one or more maximal elements—factors which are not subordinate to any other factor. In a hierarchy modelling a common-law system there should be a single maximal element corresponding to the case outcome, as in the RM. However, when modelling a civil-law system there may be multiple maximal elements corresponding to the *issues* of the domain—inputs to a legal rule which together determine the case outcome. See [16] for an analysis on the role of issues in precedential constraint.

3.3. Constraint

For the notion of constraint for HRM we introduce negations of factors as a notational device. Given a factor $p \in F$ we denote its negation by $\neg p$. We extend the set of factors F to a set $\mathcal{F}$ including these negations, so $\mathcal{F} = F \cup \{\neg p \mid p \in F\}$. Similarly, we define $\mathcal{A} = A \cup \{\neg p \mid p \in A\}$. We extend a fact situation X to operate on negations in the obvious way—if $p \in \mathrm{dom}(X)$ then $X(\neg p) = \neg X(p)$. We also define Pro, Con as in the RM:

$$\mathrm{Pro}(p) = \{q \in F \mid \mathsf{Pro}(q,p)\}, \qquad \mathrm{Con}(p) = \{q \in F \mid \mathsf{Con}(q,p)\},$$
$$\mathrm{Pro}(\neg p) = \mathrm{Con}(p), \qquad \mathrm{Con}(\neg p) = \mathrm{Pro}(p).$$

Lastly we define $\mathrm{Pro}_X(p) = \mathrm{Pro}(p) \cap \mathrm{dom}(X)$—the set of pro-$p$ factors on which a fact situation X is defined.

This brings us to the definition of constraint for the HRM.

Definition 2. *The decision of a fact situation X for a factor $p \in \mathcal{F}$ is* forced *by a case base CB, denoted $CB, X \vDash p$, if and only if either*

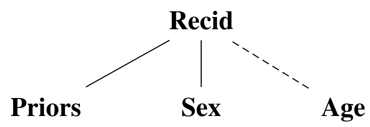
- *$X \vDash p$, or*
- *$p \in \mathcal{A}$ and there is a fact situation $Y \in CB$ with $Y \vDash p$ and*
  - *∗ for all $q \in \mathrm{Pro}_Y(p)$: if $Y \vDash q$ then $CB, X \vDash q$, and*
  - *∗ for all $q \in \mathrm{Con}_Y(p)$: if $CB, X \vDash q$ then $Y \vDash q$.*

Finally, we note that we can now consider any instance of the RM as a flat hierarchy, in which the factors of the RM are the base-level factors of the hierarchy, and the case outcome is the (single) maximal element of the hierarchy. In other words, every instance of the RM can be translated to an instance of the HRM. Moreover, this translation preserves the RM's notion of constraint, which is to say the HRM subsumes the RM.

4. Dimensional Result Model

4.1. An Example of Dimensions

Whereas a factor can be seen as a proposition, a dimension can take on a set of possible values. Usage of this terminology in the field of AI & law dates back to CATO's predecessor HYPO [17]. For an example of dimensions, we return to our running example:



(3)

Previously, the **Age** factor represented whether the defendant was over the age of 21. Viewed instead as a dimension, **Age** can take any value above 0. Similarly, we replace the **Record** factor with a dimension **Priors**, indicating the number of previous convictions.

It is not possible to say directly of a dimension whether it favors one of the two outcomes of a case. Instead, we require the dimension to come with a relation expressing the relative preference the values have for the final judgement. For instance, we know that in general older people tend to recidivate less, and so for the **Age** dimension we can say that the value 30 is less indicative than the value 21. In the graph above we have again used solid and dotted links to indicate whether higher values of the dimension are suggestive of high or low risk of recidivism. Dimensions with two values, such as the **Sex** dimension, can be considered as factors in the RM and HRM models.

4.2. Knowledge Representation

A *dimension* is a nonempty set d. We denote dimensions by lower case letters d,e,f etc. The domain is modeled by a finite set of dimensions D. A *fact situation* X is a choice function on D, i.e. a function $X : D \to \bigcup D$ such that $X(d) \in d$ for every $d \in D$.

Cases are again decided for one of the two sides π or δ, and again we assume that specific values of dimensions have a preference for either of these sides, but this is now modeled by a binary relation on the dimension. More specifically, we assume there is for each dimension $d \in D$ a *preference* relation $\preceq$ on d, which we require to be a partial order. Given values $v,w \in d$ such that $v \preceq w$, we say w prefers outcome π relative to v, and v prefers outcome δ relative to w. For this reason we will also denote $\preceq$ by $\preceq_\pi$, and $\succeq$ (the converse of $\preceq$) by $\preceq_\delta$.

4.3. Constraint

The notion of constraint for the DRM can now be stated succinctly as follows.

Definition 3. *The decision of a fact situation X for a side s is* forced *by a case base CB, denoted $CB, X \vDash s$, if there is a case $(Y,s) \in CB$ such that $Y(d) \preceq_s X(d)$ for all $d \in D$.*

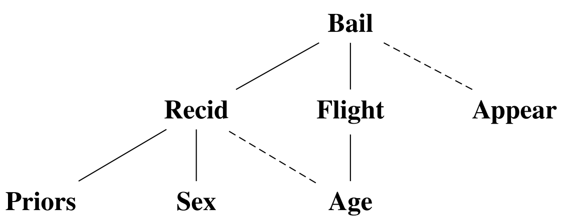
The factors of the knowledge representation for the RM can be modeled in the DRM as two-element dimensions, of which one of the elements is strictly greater than the other, depending on the preference of the factor. Again, this preserves the notion of constraint in the RM, so that the DRM subsumes the RM.

5. Dimensional Hierarchical Result Model

We now unify the knowledge representations of the previous sections, by considering a set of dimensions D together with a hierarchical structure H—a *dimension hierarchy*.

5.1. An Example of a Dimension Hierarchy

Consider the following modification of our running example, graph (2):



(4)

In the setting of a dimension hierarchy we can consider recidivism risk as a dimension, for instance as a score ranging from 1 to 10. The COMPAS system also outputs a score from 1 to 10 indicating severity of the risk, so this is a more realistic example than those in the previous sections. Additionally, **Bail** can now be considered a dimension, specifying the amount of bail in, say, USD. Note that denial of bail can still be modeled as an ‘infinite’ amount of bail. **Appear**, too, can be considered a dimension, indicating the relative frequency of past trial appearances by the defendant.

To begin building some intuition for an appropriate notion of constraint in this setting we illustrate a difference with the DRM, which is that dimensions now affect other dimensions instead of the case outcome directly. To this end, we consider the subgraph of graph (4) consisting of just the dimensions **Recid**, **Priors**, **Sex**, and **Age**, and the fact situations X, Y, and Z, listed in Table 1. The situation Z concerns a 25-year-old female with 2 prior offenses. What recidivism risk score may be consistently assigned to Z, given the previous judgements that a 30-year-old female with 1 prior offense received score 5 (situation X), and that a 20-year-old male with 4 prior offenses received score 8 (situation Y)? Comparing the situation Z to X we see that Z is dimension-wise equal or more indicative of recidivism risk than X: Z is younger, both Z and X are female, and Z has more prior offenses. Since X received a recidivism risk score of 5, it seems sensible to require that Z would get *at least* a score of 5, but possibly higher since Z is indicative of higher risk on some dimensions. This exemplifies one of the key differences between the DHRM and the previous models—decisions are not forced exactly, but constrained to lie within an interval. Comparing Z to the situation Y we get the opposite picture; Y has received a risk score of 8, but Z is dimension-wise equal or less indicative of recidivism risk than Y. Therefore, we expect Z to receive a score of *at most* 8. In sum, the case base $\{X,Y\}$ should produce the constraint that $5 \preceq Z(\textbf{Recid}) \preceq 8$.

We now turn our attention to the full hierarchy, depicted by graph (4), involving a downstream judgement of bail amount. In such a scenario, we can apply a recursive notion of constraint as in the HRM. Consider, again, the fact situations listed in Table 1. We have seen that X and Y bind the recidivism score of Z to the integer range $[5,8]$. In addition, we now have two situations V and W for which a bail amount was determined on the basis of their recidivism risk assessment, risk of flight, and relative frequency of

Table 1. Five example fact situations V, W, X, Y, and Z for the bail domain.

	Age	**Sex**	**Priors**	**Recid**	**Flight**	**Appear**	**Bail**
X	30	F	1	5	-	-	-
Y	20	M	4	8	-	-	-
V	-	-	-	2	low	0.8	$2,500
W	-	-	-	9	high	0.3	$20,000
Z	25	F	2	?	low	0.5	?

previous trial appearances. Defendant V was granted a bail amount of \$2,500, on the basis of a recidivism risk score of 2, a perceived low risk of flight, and an 80% appearance rate at previous trials. Defendant Z has a lower appearance rate at previous trials and is similarly perceived as unlikely to flee, but is not yet assigned a definitive recidivism risk score. However, since we know that Z should receive a risk score of at least 5 it will in any case be higher than V's score of 2. Therefore, we would ultimately expect Z to receive a bail amount which is equal or greater than that of V—so $\$2{,}500 \preceq Z(\textbf{Bail})$. Similarly, we can deduce from the case W that, since Z should receive a risk score of at most 8, the amount of bail for Z should not exceed \$20,000—so $Z(\textbf{Bail}) \preceq \$20{,}000$. In sum, the case base $\{V,W,X,Y\}$ should produce the constraints $\$2{,}500 \preceq Z(\textbf{Bail}) \preceq \$20{,}000$.

This use of recursion is useful, because it allows the use of the forcing relation despite some dimensions not having been assigned an exact value. Consider, for instance, the decision support system used by the Dutch national police force [11]. It is argued in [12] that determining the values of dimensions for a specific case can be costly, and the aforementioned use of recursion can alleviate this need for abstract factors.

5.2. Knowledge Representation

A *dimension hierarchy* is a set D of dimensions, together with a hierarchy H satisfying the familiar conditions listed in Section 3.2. We maintain the terminology from the HRM—a dimension is *base-level* if it is H-minimal, and *abstract* otherwise; A is the set of abstract dimensions. A *fact situation* X is a choice function on a subset of D; we denote its domain by $\mathrm{dom}(X)$. Lastly, we assume each dimension $d \in D$ is assigned a partial order $\preceq$ on d.

5.3. Constraint

As in the HRM, we define $\mathrm{Pro}(d) = \{e \in D \mid \mathsf{Pro}(e,d)\}$ and $\mathrm{Pro}_X(d) = \mathrm{Pro}(d) \cap \mathrm{dom}(X)$; the sets $\mathrm{Con}(d)$ and $\mathrm{Con}_X(d)$ are defined analogously. Using these, we now define by mutual recursion two relations $CB \vDash v \preceq X(d)$ and $CB \vDash X(d) \preceq v$.

Definition 4. *Given a case base CB and a value v in some dimension d, a fact situation X is* lower bounded *by v and CB, written* $CB \vDash v \preceq X(d)$*, if and only if either*

- $v \preceq X(d)$*, or*
- $d \in A$ *and there is* $Y \in CB$ *such that* $v \preceq Y(d)$ *and*
  - ∗ *for all* $e \in \mathrm{Pro}_Y(d)$: $CB \vDash Y(e) \preceq X(e)$*, and*
  - ∗ *for all* $e \in \mathrm{Con}_Y(d)$: $CB \vDash X(e) \preceq Y(e)$.

The upper bound *by v, written* $CB \vDash X(d) \preceq v$*, is defined analogously.*

The idea behind the recursive clause is that there is a precedent Y which, by the a fortiori principle, forces $X(d)$ to take a value which is at least $Y(d)$, and therefore $v \preceq X(d)$ follows by transitivity from $v \preceq Y(d) \preceq X(d)$.

Let us now verify that Definition 4 correctly captures the intuition of the example in Section 5.1. We consider a dimension hierarchy as depicted in graph (4), with a case base $\{W,Y\}$ and novel fact situation Z as listed in Table 1. The question is now whether $\{W,Y\} \vDash Z(\mathbf{Bail}) \preceq \$20{,}000$. To check this, we first verify that $\{W,Y\} \vDash Z(\mathbf{Recid}) \preceq 9$:

$$
\begin{aligned}
&\{W,Y\} \vDash Z(\mathbf{Recid}) \preceq 9 && (5)\\
&\text{if there is } T \in \{W,Y\} \text{ such that } T(\mathbf{Recid}) \preceq 9 \text{ and} && (6)\\
&\quad * \text{ for all } d \in \mathrm{Pro}_T(\mathbf{Recid})\text{: } \{W,Y\} \vDash Z(d) \preceq T(d), \text{ and} &&\\
&\quad * \text{ for all } d \in \mathrm{Con}_T(\mathbf{Recid})\text{: } \{W,Y\} \vDash T(d) \preceq Z(d) &&\\
&\text{if } \{W,Y\} \vDash Z(\mathbf{Priors}) \preceq 4, \text{ and} && (7)\\
&\quad \{W,Y\} \vDash Z(\mathbf{Sex}) \preceq \mathrm{M}, \text{ and} &&\\
&\quad \{W,Y\} \vDash 20 \preceq Z(\mathbf{Age}) &&\\
&\text{if } Z(\mathbf{Priors}) \preceq 4 \text{ and } 20 \preceq Z(\mathbf{Age}). && (8)
\end{aligned}
$$

Since $Z(\mathbf{Recid})$ is undecided, we begin by unfolding (5) to the recursive clause of Definition 4, which gives (6). We then substitute $T = Y$, as $Y(\mathbf{Recid}) = 8 \preceq 9$, yielding (7). Since all the dimensions subordinate to **Recid** are base-level, and any value v of **Sex** satisfies $v \preceq \mathrm{M}$, we can simplify (7) to (8). Indeed, defendant Z of Table 1 satisfies these conditions, and so we have verified that $\{W,Y\} \vDash Z(\mathbf{Recid}) \preceq 9$.

Next, we proceed in the same fashion to confirm that $\{W,Y\} \vDash Z(\mathbf{Bail}) \preceq \$20{,}000$:

$$
\begin{aligned}
&\{W,Y\} \vDash Z(\mathbf{Bail}) \preceq \$20{,}000 && (9)\\
&\text{if there is } T \in \{W,Y\} \text{ such that } T(\mathbf{Bail}) \preceq \$20{,}000 \text{ and} && (10)\\
&\quad * \text{ for all } d \in \mathrm{Pro}_T(\mathbf{Bail})\text{: } \{W,Y\} \vDash Z(d) \preceq T(d), \text{ and} &&\\
&\quad * \text{ for all } d \in \mathrm{Con}_T(\mathbf{Bail})\text{: } \{W,Y\} \vDash T(d) \preceq Z(d) &&\\
&\text{if } \{W,Y\} \vDash Z(\mathbf{Recid}) \preceq 9, \text{ and} && (11)\\
&\quad \{W,Y\} \vDash Z(\mathbf{Flight}) \preceq \text{high}, \text{ and} &&\\
&\quad \{W,Y\} \vDash 0.3 \preceq Z(\mathbf{Appear}) &&\\
&\text{if } \{W,Y\} \vDash Z(\mathbf{Recid}) \preceq 9 \text{ and } 0.3 \preceq Z(\mathbf{Appear}). && (12)
\end{aligned}
$$

The reasoning proceeds in the same way as in the previous derivation, except now we substitute W for T. Line (12) holds for defendant Z listed in Table 1, as we have already shown that $\{W,Y\} \vDash Z(\mathbf{Recid}) \preceq 9$, and $0.3 \preceq Z(\mathbf{Appear})$ holds by assumption.

6. Conclusion

In this work, we have proposed an extension of Horty's result model of precedential constraint [1] which accounts for both dimensional and hierarchical information, thereby subsuming two of its extensions given in [2] and [5]. We exemplified and motivated practical use of such models through an example of the legal domain of criminal sentencing.

Using this example, we showed that our formal model correctly captures the intuition of some examples from this domain.

In the future, we intend to use the theory developed in this work and apply it to the improvement of interpretability and responsible use of black-box AI systems. In particular, it would be interesting to see systems like COMPAS adhere to the notion of constraint for the DHRM.

Acknowledgements

We kindly thank the reviewers for their careful corrections and helpful suggestions. This research was (partially) funded by the Hybrid Intelligence Center, a 10-year programme funded by the Dutch Ministry of Education, Culture and Science through the Netherlands Organisation for Scientific Research, grant number 024.004.022.

References

[1] Horty J. Rules and reasons in the theory of precedent. Legal Theory. 2011;17(1):1–33.

[2] Horty J. Reasoning with dimensions and magnitudes. Artificial Intelligence and Law. 2019:309-45.

[3] Roth AC. Case-based reasoning in the law. A formal theory of reasoning by case comparison. [Dissertation]. Universiteit Maastricht; 2003.

[4] Roth B, Verheij B. Dialectical Arguments and Case Comparison. In: Gordon TF, editor. Legal Knowledge and Information Systems. JURIX 2004: The Seventeenth Annual Conference; 2004. p. 99-108.

[5] van Woerkom W, Grossi D, Verheij B, Prakken H. Hierarchical Precedential Constraint. In: Nineteenth International Conference on Artificial Intelligence and Law. ICAIL'23; 2023. p. 333-42.

[6] Canavotto I, Horty J. Reasoning with hierarchies of open-textured predicates. In: Nineteenth International Conference on Artificial Intelligence and Law. ICAIL'23; 2023. p. 333-42.

[7] Prakken H, Ratsma R. A top-level model of case-based argumentation for explanation: Formalisation and experiments. Argument & Computation. 2022;13(2):159-94.

[8] Peters J, Bex F, Prakken H. Justifications derived from Inconsistent case bases using authoritativeness. In: Proceedings of the 1st International Workshop on Argumentation for eXplainable AI. vol. 3209; 2022. p. 1-13.

[9] Peters J, Bex F, Prakken H. Model- and data-agnostic justifications with A Fortiori Case-Based Argumentation. In: Nineteenth International Conference on Artificial Intelligence and Law. ICAIL'23; 2023. p. 207-16.

[10] van Woerkom W, Grossi D, Prakken H, Verheij B. Landmarks in Case-Based Reasoning: From Theory to Data. In: Schlobach S, Pérez-Ortiz M, Tielman M, editors. HHAI2022: Augmenting Human Intellect. IOS Press; 2022. p. 212-24.

[11] Odekerken D, Bex F, Villata S, Harašta J, Křemen P, et al. Towards Transparent Human-in-the-Loop Classification of Fraudulent Web Shops. Legal Knowledge and Information Systems. 2020;334:239-42.

[12] Odekerken D, Bex F, Prakken H. Justification, stability and relevance for case-based reasoning with incomplete focus cases. In: Nineteenth International Conference on Artificial Intelligence and Law. ICAIL'23; 2023. p. 177-86.

[13] Larson J, Mattu S, Kirchner L, Angwin J. How We Analyzed the COMPAS Recidivism Algorithm; 2016. Available at `https://archive.ph/gPvQP`, accessed 7 Jun 2023. ProPublica.

[14] Yukhnenko D, Blackwood N, Fazel S. Risk factors for recidivism in individuals receiving community sentences: a systematic review and meta-analysis. CNS spectrums. 2020;25(2):252-63.

[15] Aleven V. Using background knowledge in case-based legal reasoning: A computational model and an intelligent learning environment. Artificial Intelligence. 2003;150(1):183-237.

[16] Bench-Capon T, Atkinson K. Precedential Constraint: The Role of Issues. In: Proceedings of the Eighteenth International Conference on Artificial Intelligence and Law. ICAIL'21; 2021. p. 12–21.

[17] Ashley KD. Reasoning with cases and hypotheticals in HYPO. International Journal of Man-Machine Studies. 1991;34(6):753-96.