Hardness of case-based decisions: A formal theory

Zheng, H.; Grossi, D.; Verheij, B.

DOI
10.1145/3462757.3466071

Publication date
2021

Document Version
Final published version

Published in
Eighteenth International Conference on Artificial Intelligence and Law : proceedings of the conference

Hardness of Case-Based Decisions: a Formal Theory

Heng Zheng
Artificial Intelligence, Bernoulli Institute, University of Groningen
The Netherlands
h.zheng@rug.nl

Davide Grossi
Artificial Intelligence, Bernoulli Institute, University of Groningen
ILLC/ACLE, University of Amsterdam
The Netherlands
d.grossi@rug.nl

Bart Verheij
Artificial Intelligence, Bernoulli Institute, University of Groningen
The Netherlands
bart.verheij@rug.nl

ABSTRACT

Stare decisis is a fundamental principle of case-based reasoning. Yet its application varies in complexity and depends, in particular, on whether relevant past decisions agree, or exist at all. The contribution of this paper is a formal treatment of types of the hardness of case-based decisions. The typology of hardness is defined in terms of the arguments for and against the issue to be decided, and their kind of validity (conclusive, presumptive, coherent, incoherent). We apply the typology of hardness to Berman and Hafner's research on the dynamics of case-based reasoning and show formally how the hardness of decisions varies with time.

CCS CONCEPTS

• Computing methodologies → **Knowledge representation and reasoning**; **• Applied computing** → *Law*.

KEYWORDS

Case-based reasoning, computational argumentation, hard cases

ACM Reference Format:
Heng Zheng, Davide Grossi, and Bart Verheij. 2021. Hardness of Case-Based Decisions: a Formal Theory . In *Eighteenth International Conference for Artificial Intelligence and Law (ICAIL'21), June 21–25, 2021, São Paulo, Brazil.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3462757.3466071

1 INTRODUCTION

Legal decision-making can be hard, very hard. Its complexities, which have been well-recognized in AI and Law since its early days, are numerous. An early contribution to the discussion of the hardness of legal decision-making in AI and Law is 'An Artificial Intelligence Approach to Legal Reasoning', a book by Gardner [9]. In that book, Gardner addresses the distinction between hard and easy cases using ideas from jurisprudence. Following Rissland's review [18] of the landmark work by Gardner, legal decisions are guided rather than governed by existing law; legal terms are open textured; legal questions can have more than one answer, but a reasonable and timely answer must be given; and the answers to legal questions can change over time. These, and other, such complexities of legal decision-making have been object of much work in AI and Law (cf. also [23]).

But some cases are easier to decide than others. For instance, when all past cases agree on a given legally relevant fact situation, decision-making using the principle of *stare decisis* can be straightforward. When relevant past cases disagree, things get harder. Sometimes such a conflict of precedents can be resolved, for instance when one precedent is considered a landmark case overturning existing doctrine, or when a precedent comes from a higher level court. But not all conflicts can be resolved, making decision-making harder. Also it can happen that a legally relevant fact situation has no matching precedent, so the *stare decisis* principle gives no answer.

Paper contribution. As the above examples show, the hardness of case-based decisions comes in different types. It is the topic of this paper to provide a formal theory of the hardness of cases in case-based decision-making. Significant work has been devoted to the nature and dynamics of case-based reasoning (e.g., [1, 2, 5–8, 11, 13–17, 21]), and to the topic of hard cases (e.g., [4, 9, 12]). Yet, to the best of our knowledge, no formal theory has been proposed so far of what makes a current case harder, or less hard, than other cases. We provide such a theory here, by focusing on the following question: *is there a typology of how hard it is to make a decision about an issue in case-based reasoning?* To answer this question, we propose a formal approach based on the case model formalism [20–22, 25, 26]. We describe a decision-making issue as an argument and its counterargument, and formalize its hardness with the validity of these arguments. We also illustrate the approach with a case study in the dynamics of case-based reasoning (following an example by Berman and Hafner [6–8, 11] as formalized in [21]), and show how hardness varies over time.

Paper outline. Section 2 introduces earlier work in the case model formalism. Section 3 develops a formal theory of hardness. Section 4 shows an application of our approach to a series of concrete legal cases highlighting the development of hardness over time. Section 5 positions our theory within existing literature on case-based decision-making in law. Section 6 concludes. Detailed proof sketches of relevant formal properties are provided throughout the paper.

2 PRELIMINARIES: CASE MODELS

Our approach uses case models [20], a formalism based on a propositional logic language L generated from a finite set of constants. We write $\neg$ for negation, $\wedge$ for conjunction, $\vee$ for disjunction, $\leftrightarrow$ for equivalence, $\top$ for a tautology, and $\bot$ for a contradiction. The associated classical logical consequence relation is denoted $\models$. Cases can

ICAIL'21, June 21–25, 2021, São Paulo, Brazil

ACM ISBN 978-1-4503-8526-8/21/06. https://doi.org/10.1145/3462757.3466071

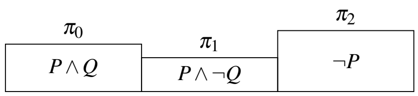

Figure 1: Example of a case model. Larger boxes denote cases that are preferred over the cases denoted by smaller boxes.

be compared through the preference relation between cases in case models. A *case model* is a set of logically consistent, incompatible cases forming a total preorder (i.e., a transitive, total binary relation) representing a preference relation among the cases.

Definition 1 (Case models [20]) *A* case model *is a pair* $\mathscr{C} = (C, \geq)$ *with finite* $C \subseteq L$*, such that, for all* π, π' *and* $\pi'' \in C$*:*

(1) $\not\models \neg\pi$*;*
(2) If $\not\models \pi \leftrightarrow \pi'$*, then* $\models \neg(\pi \wedge \pi')$*;*
(3) If $\models \pi \leftrightarrow \pi'$*, then* $\pi = \pi'$*;*
(4) $\pi \geq \pi'$ *or* $\pi' \geq \pi$*;*
(5) If $\pi \geq \pi'$ *and* $\pi' \geq \pi''$*, then* $\pi \geq \pi''$*.*

As customary, the asymmetric part of $\geq$ is denoted $>$. The symmetric part of $\geq$ is denoted $\sim$. Intuitively, $\geq$ means 'at least as preferred as'. For instance, $P \wedge Q \geq P \wedge \neg Q$ means that the case expressed by $P \wedge Q$ is higher than or equally high in the preference ordering than the case expressed by $P \wedge \neg Q$.

Example 1 ([21]) Cases $\pi_0 = P \wedge Q$, $\pi_1 = P \wedge \neg Q$, and $\pi_2 = \neg P$, and preference relation $\pi_2 > \pi_0 > \pi_1$ form a case model (Figure 1).

We move now to the definition of arguments and their validities.

Definition 2 (Arguments [20]) *An* argument *from* χ *to* ρ *is a pair* (χ, ρ) *with* χ *and* $\rho \in L$*. For* $\lambda \in L$*, if* $\chi \models \lambda$*,* λ *is a* premise *of the argument; if* $\rho \models \lambda$*,* λ *is a* conclusion*; if* $\chi \wedge \rho \models \lambda$*,* λ *is a* position *made by the argument. We say that* χ *expresses the* full premise *of the argument,* ρ *the* full conclusion*, and* $\chi \wedge \rho$ *its* full position *made by the argument.*

Example 2 For instance, (P, Q) is an argument with P as its full premise, and Q as its full conclusion. Sentence $P \wedge Q$ is the full position by the argument (P, Q).

Arguments have three kinds of validities. If the full position made by an argument is logically implied by one of the precedents in a case model, then the argument is *coherently valid* in the case model. If an argument's conclusion is logically implied by a precedent which is weakly preferred over all precedents that logically imply the argument's full premise, then the argument is *presumptively valid* in the case model. If all precedents that logically imply the full premise of a coherently valid argument, also logically imply its full conclusion, then the argument is *conclusively valid* in the model.

Definition 3 (Validity of arguments [20]) *Let* $\mathscr{C} = (C, \geq)$ *be a case model and* (χ, ρ) *an argument:*

(1) (χ, ρ) *is* coherent *w.r.t.* $\mathscr{C}$ *if and only if* $\exists \pi \in C : \pi \models \chi \wedge \rho$*. We then write* $\mathscr{C} \models \chi \top \rho$*. If* $\mathscr{C} \not\models \chi \top \rho$*, then* (χ, ρ) *is* incoherent *w.r.t.* $\mathscr{C}$*. We then write* $\mathscr{C} \models \chi \bot \rho$*.*
(2) (χ, ρ) *is* presumptive *w.r.t.* $\mathscr{C}$ *if and only if* $\exists \pi \in C$ *s.t.:*
   (a) $\pi \models \chi \wedge \rho$*; and*
   (b) for all $\pi' \in C$*: if* $\pi' \models \chi$*, then* $\pi \geq \pi'$*.*
   We then write $\mathscr{C} \models \chi \rightsquigarrow \rho$*.*
(3) (χ, ρ) *is* conclusive *w.r.t.* $\mathscr{C}$ *if and only if* $\exists \pi \in C : \pi \models \chi \wedge \rho$ *and for all* $\pi \in C$*: if* $\pi \models \chi$*, then* $\pi \models \chi \wedge \rho$*. We then write* $\mathscr{C} \models \chi \Rightarrow \rho$*.*

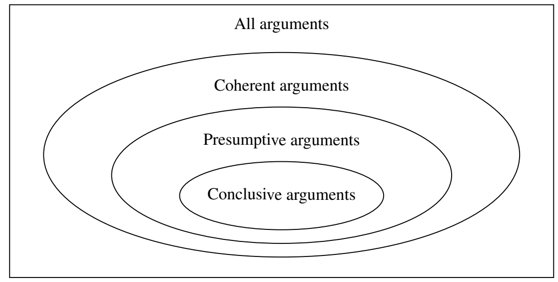

Figure 2: Validity of arguments in a Venn diagram.

Example 3 If we consider arguments (P, Q), $(\neg P, Q)$, and $(P \vee R, P)$ with the case model $\mathscr{C}$ shown in Figure 1, we have:

(1) $\mathscr{C} \models P \top Q$, since $\pi_0 \models P \wedge Q$;
(2) $\mathscr{C} \models \neg P \bot Q$, since there is no case in $\mathscr{C}$ can imply $\neg P \wedge Q$;
(3) $\mathscr{C} \models P \rightsquigarrow Q$, since $\pi_0 \models P \wedge Q$, and π_0 is the most preferred case in $\mathscr{C}$ that implies P;
(4) $\mathscr{C} \models P \vee R \Rightarrow P$, since for all cases in $\mathscr{C}$, if they imply $P \vee R$, they also imply its conclusion P.

The following proposition shows relations between validities.

Proposition 1 *Let* $\mathscr{C}$ *be a case model,* (χ, ρ) *be an argument:*

(1) If $\mathscr{C} \models \chi \Rightarrow \rho$*, then* $\mathscr{C} \models \chi \rightsquigarrow \rho$*, but not in general vice versa;*
(2) If $\mathscr{C} \models \chi \rightsquigarrow \rho$*, then* $\mathscr{C} \models \chi \top \rho$*, but not in general vice versa.*
(3) For all $\chi, \rho \in L$*, either* $\mathscr{C} \models \chi \top \rho$ *or* $\mathscr{C} \models \chi \bot \rho$*.*

Proof. We prove the first claim. The other cases are similar. Observe that since (χ, ρ) is conclusively valid, by Definition 3, for all $\pi \in C$, if $\pi \models \chi$, then $\pi \models \chi \wedge \rho$. Furthermore, a case $\pi \models \chi \wedge \rho$ exists again by the definition of conclusive validity. There are two cases. Either such a π is maximally preferred among the cases logically implying χ, and we have $\pi \models \chi \wedge \rho$ as desired. Or it is not. There exists then a maximally preferred $\pi' \in C$ such that $\pi' \models \chi$ and $\pi' \geq \pi$. But by the above observation we also have that $\pi' \models \rho$ as desired. We conclude that (χ, ρ) is presumptively valid. □

Figure 2 is an illustration of Proposition 1. Conclusive arguments are presumptively valid, and presumptive arguments are coherently valid. All arguments are either incoherent with respect to the case model (outside the set of coherent arguments), or coherent with respect to the model (inside the set of coherent arguments).

3 A FORMAL THEORY OF HARDNESS

We turn to the main theoretical contribution of the paper: a formal approach to compare issues in case models by their hardness.

We start by introducing the key definitions underpinning our theory, in two steps. First we introduce a natural way of ordering arguments by their strength, in essence based on Proposition 1. We show then how this notion can be used also to develop ways in which issues can be compared. We think of issues essentially as *pairs* of arguments (e.g., the plaintiff's and the defendant's arguments): the argument that points to the truth of the issue, an the one that points to its falsity. Here arguments are premise-conclusion pairs, without considering a possible internal stepwise structure. Based on two natural ways to compare issues, we will define and study an ordering of issues representing their relative hardness.

3.1 Comparing arguments

First we introduce labels for arguments based on their validity.

Definition 4 (Validity labels for arguments) *Let (χ,ρ) be an argument and $\mathscr{C}$ a case model. Then the validity label of (χ,ρ) in $\mathscr{C}$ is denoted $A_{\mathscr{C}}(\chi,\rho)$ and is defined as follows:*

(1) $A_{\mathscr{C}}(\chi,\rho)$ = conc if $\mathscr{C} \models \chi \Rightarrow \rho$;
(2) $A_{\mathscr{C}}(\chi,\rho)$ = pres if $\mathscr{C} \models \chi \rightsquigarrow \rho$ and $\mathscr{C} \not\models \chi \Rightarrow \rho$;
(3) $A_{\mathscr{C}}(\chi,\rho)$ = coh if $\mathscr{C} \models \chi \top \rho$ and $\mathscr{C} \not\models \chi \rightsquigarrow \rho$;
(4) $A_{\mathscr{C}}(\chi,\rho)$ = incoh if $\mathscr{C} \not\models \chi \top \rho$.

For instance, pres is the label for arguments that are presumptively valid, but not conclusive. Using Figure 2 as an illustration, these arguments are in the presumptive arguments set, but not in the conclusive arguments set. The label incoh is used for arguments that are not coherent. In Figure 2, these arguments are in the set of all arguments, but not in the set of coherent arguments.[1]

Validity labels come with a natural ordering:

Definition 5 (Validity ordering) *The validity ordering is the total order $\geq$ on set {conc, pres, coh, incoh} characterized by the following property:*[2] *conc $\geq$ pres $\geq$ coh $\geq$ incoh.*

Intuitively, based on their validity, arguments with label conc are stronger than arguments with label pres. Similarly, arguments with label pres are stronger than arguments with label coh, and arguments with label coh are stronger than arguments with label incoh.

The following proposition relates argument validities (Definition 3) to validity labels and validity ordering (Definitions 4 and 5).

Proposition 2 *Let (χ,ρ) be an argument and $\mathscr{C}$ a case model. Then the following hold:*

(1) (χ,ρ) is conclusive if and only if $A_{\mathscr{C}}(\chi,\rho)$ = conc;
(2) (χ,ρ) is presumptive if and only if $A_{\mathscr{C}}(\chi,\rho) \geq$ pres;
(3) (χ,ρ) is coherent if and only if $A_{\mathscr{C}}(\chi,\rho) \geq$ coh;
(4) (χ,ρ) is incoherent if and only if $A_{\mathscr{C}}(\chi,\rho)$ = incoh.

PROOF. Immediate using the definitions and Proposition 1. □

Using the ordering of arguments their validity label, we can quantify how far apart they are in terms of strength, as follows:

Definition 6 (Validity distance) *Let v_0 and $v_1 \in$ {conc, pres, coh, incoh}. Then we define the* validity distance *$vd(v_0,v_1)$ as the length of the shortest path from v_0 to v_1 in the validity ordering.*

Intuitively, the validity distance between two arguments is the number of steps in the validity ordering $\geq$ that are needed to go from the label of the weaker argument to the label of the stronger one (of course avoiding loops). Clearly, as $\geq$ has length 3, the maximal validity distance is 3 and the minimal is 0.

[1] These labels and their ordering in the following Definition 5 are related to the quantitative representation of [20, Section 3.3]. Arguments with label conc correspond to those arguments with strength equal to 1 [20]. Arguments with label pres corresponds to those arguments with strength above a given threshold but less than 1. Arguments with label coh corresponds to arguments with strength less than the given threshold but still above 0. And arguments with label incoh corresponds to arguments with strength equal to 0.
[2] Recall that a total order is a binary relation which is transitive, total and antisymmetric.

Example 4 In the case model of Figure 1, argument (P,Q) has type pres, and argument $(P,\neg Q)$ has type coh. Hence the validity distance between the validity labels of these two arguments is 1.

3.2 Issues

We introduce now the notion of issue, that is, the specific proposition a case model is supposed to decide about. Our theory concerns precisely the hardness of deciding about the truth or falsity of an issue given a situation in a case model.

Issues and issue types. So an issue is a sentence whose truth or falsity we would like to establish, given a situation. Formally:

Definition 7 (Issues) *A situation is a sentence $\sigma \in L$. A sentence $\iota \in L$ is an* issue *given situation σ (denoted $\sigma \pm \iota$) if and only if $\sigma \not\models \iota$ and $\sigma \not\models \neg\iota$.*

In other words, an issue for a given situation is a sentence whose truth or falsity is not logically settled by the situation. It is worth observing that $\sigma \pm \iota$ is an issue if and only if $\sigma \pm \neg\iota$ is an issue.

Example 5 For instance, $P \pm Q$ represents an issue Q with respect to a situation P.

Importantly, for every issue $\sigma \pm \iota$ there are two naturally associated arguments: (σ,ι) and $(\sigma,\neg\iota)$. We will study hardness as a relation on the types of those pairs of arguments that correspond to issues.

Given the pair of arguments induced by an issue, we call a *type* the multi-set (i.e., a set admitting multiple copies of an element) consisting of the validity labels of the two arguments. Formally:

Definition 8 (Types of issues) *Let $\mathscr{C}$ be a case model and $\sigma \pm \iota$ an issue. Then the* type *of the issue, denoted $T_{\mathscr{C}}(\sigma \pm \iota)$, is the multiset:*

$$\{A_{\mathscr{C}}(\sigma,\iota), A_{\mathscr{C}}(\sigma,\neg\iota)\}.$$

Example 6 The type of issue $P \pm Q$ with respect to the case model in Figure 1 is {pres,coh}, as $A_{\mathscr{C}}(P,Q)$ = pres and $A_{\mathscr{C}}(P,\neg Q)$ = coh.

Not all types are logically possible. The following proposition shows there are in total 7 possible issue types.

Proposition 3 *Let $\mathscr{C}$ be a case model and $\sigma \pm \iota$ an issue. Then $T_{\mathscr{C}}(\sigma \pm \iota)$ equals one of the following:*

(1) {conc, incoh};
(2) {pres, pres};
(3) {pres, coh};
(4) {pres, incoh};
(5) {coh, coh};
(6) {coh, incoh};
(7) {incoh, incoh}.

PROOF. By Definition 4, an argument can have 4 possible labels conc, pres, coh, incoh. There are therefore $\frac{(4^2-4)}{2}+4 = 10$ possible multisets of size 2. Let a case model $\mathscr{C} = (C,\geq)$ and an issue $\sigma \pm \iota$ be given. We split the proof in two parts. First, we reason by cases and for a multiset $\{v_0, v_1\}$ (for which we can assume $v_0 \geq v_1$) we show what values among conc, pres, coh, incoh v_1 can take once we fix v_0. Then we display examples showing the types listed in the statement are possible.

$\boxed{v_0 = \text{conc}}$ Then $v_1 \in$ {conc, pres, coh, incoh}. If $A_{\mathscr{C}}(\sigma,\iota)$ = conc, by Definition 3, for each $\pi \in C$ with $\pi \models \sigma$, we have $\pi \models \sigma \wedge \iota$. Then since cases are consistent, $\pi \not\models \sigma \wedge \neg\iota$, so $A_{\mathscr{C}}(\sigma,\neg\iota)$ = incoh and the types {conc, conc}, {conc, pres}, and {conc, coh} are not possible. Similarly, when $A_{\mathscr{C}}(\sigma, \neg\iota) = v_0$. Therefore, $T_{\mathscr{C}}(\sigma \pm \iota)$ is equal to {conc, incoh}.

$\boxed{v_0 = \text{pres}}$ Then $v_1 \in$ {pres, coh, incoh} and $A_{\mathscr{C}}(\sigma, \neg\iota) \in$ {pres, coh, incoh}. Similarly, when $A_{\mathscr{C}}(\sigma, \neg\iota) = v_0$. Therefore, $T_{\mathscr{C}}(\sigma \pm \iota)$ is equal to {pres, pres}, {pres, coh}, or {pres, incoh}.

$\boxed{v_0 = \text{coh}}$ Then $v_1 \in$ {coh, incoh}, and $T_{\mathscr{C}}(\sigma \pm \iota)$ is equal to {coh,coh} or {coh,incoh}.

$\boxed{v_0 = \text{incoh}}$ Then v_1 = incoh and $T_{\mathscr{C}}(\sigma \pm \iota)$ is {incoh,incoh}.

Now we give an example for each possible type. For {conc, incoh}, let $\mathscr{C}$ be a case model with only one case $\pi_0 = P \wedge Q$, $P \pm Q$ an issue. $T_{\mathscr{C}}(P \pm Q)$ = {conc, incoh}.

Let $\mathscr{C}$ be a case model with cases $\pi_0 = P \wedge Q$ and $\pi_1 = P \wedge \neg Q$, $P \pm Q$ an issue. For {pres, pres}, if $\mathscr{C}$ has a preference relation $\pi_0 \sim \pi_1$, then $T_{\mathscr{C}}(P \pm Q)$ = {pres, pres}. For {pres, coh}, if $\mathscr{C}$ has a preference relation $\pi_0 > \pi_1$, then $T_{\mathscr{C}}(P \pm Q)$ = {pres, coh}.

For {pres, incoh}, let $\mathscr{C}$ be a case model with cases $\pi_0 = P \wedge Q$ and $\pi_1 = P \wedge R$, $P \pm Q$ an issue. If $\mathscr{C}$ has a preference relation $\pi_0 > \pi_1$, then $T_{\mathscr{C}}(P \pm Q)$ = {pres, incoh}.

For {coh, coh}, let $\mathscr{C}$ be a case model with cases $\pi_0 = P$, $\pi_1 = P \wedge Q$, and $\pi_2 = P \wedge \neg Q$, $P \pm Q$ an issue. If $\mathscr{C}$ has a preference relation $\pi_0 > \pi_1 \sim \pi_2$, then $T_{\mathscr{C}}(P \pm Q)$ = {coh, coh}.

For {coh, incoh}, let $\mathscr{C}$ be a case model with cases $\pi_0 = P$ and $\pi_1 = P \wedge Q$, $P \pm Q$ an issue. If $\mathscr{C}$ has a preference relation $\pi_0 > \pi_1$, then $T_{\mathscr{C}}(P \pm Q)$ = {coh, incoh}.

For {incoh, incoh}, let $\mathscr{C}$ be a case model only with case $\pi_0 = P$, $P \pm Q$ an issue. Then $T_{\mathscr{C}}(P \pm Q)$ = {incoh, incoh}. □

Proposition 3 depends on the preference relation between cases in $\mathscr{C}$ to be general. If restrictions are imposed on that preference relation, fewer types may be possible. In particular, if such a preference relation is trivial (in the sense that all cases are at least as preferred as all other cases), like in the case models representing HYPO examples [25], then only 4 types are possible: {conc,incoh}, {pres,incoh}, {pres,pres} and {incoh,incoh} since then a coherent argument is always presumptive.

Comparing issues. There are two natural ways in which to compare types. They can be compared by the relative strength of their validity labels, or by the distance of their labels. Formally:

Definition 9 (Type orderings) *Let $\{v_0, v_1\}$ and $\{v'_0, v'_1\}$ be types. We define two binary relations $\succeq_v$ and $\succeq_d \in$ {conc, pres, coh, incoh}^{2} as follows:*

(1) $\{v_0, v_1\} \succeq_v \{v'_0, v'_1\}$ if and only if: $v_0 \geq v'_0$ and $v_1 \geq v'_1$, or $v_0 \geq v'_1$ and $v_1 \geq v'_0$;
(2) $\{v_0, v_1\} \succeq_d \{v'_0, v'_1\}$ if and only if $vd(v_0, v_1) \geq vd(v'_0, v'_1)$.

The asymmetric parts of $\succeq_v$, $\succeq_d$ are respectively denoted $\succ_v$ and $\succ_d$. Their symmetric parts are respectively denoted $\sim_v$ and $\sim_d$.

Intuitively, relation $\succeq_v$ orders types by the strength of their validity labels defined in Definition 5. So higher types in $\succeq_v$ pertain issues involving stronger arguments, while lower types in $\succeq_v$ pertain issues involving weaker arguments. Instead, relation $\succeq_d$ orders types by how far apart the labels within the type are from each other. So higher types in $\succeq_d$ pertain issues involving arguments containing a strong and a weak argument, while lower types in $\succeq_d$ pertain issues involving arguments of similar strength.

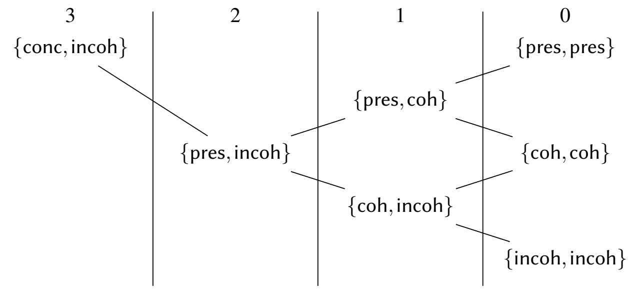

Figure 3: Hasse diagram of $\succeq_v$. The numbered columns depict the equivalence classes of $\succeq_d$ (one per distance from 0 to 3).

Example 7 Continuing on Example 6, we have:

{pres,coh} $\succeq_v$ {incoh,incoh};
{pres,coh} $\succeq_d$ {incoh,incoh}.

We will see these relations at work in the definition of hardness ordering provided in the next subsection. For now it is important to observe that these relations are well-behaved:

Proposition 4 *We have that:*

(1) $\succeq_v$ is a partial order, which is not total;
(2) $\succeq_d$ is a total preorder, which is not antisymmetric.

PROOF. Claim 1. By Definition 9, $\succeq_v$ inherits the properties of $\geq$ (Definition 5) and is therefore reflexive, antisymmetric, and transitive; hence a partial order. However, it is not total since some types are incomparable, for instance {conc,incoh} and {pres,coh} can not be compared since conc$\geq$pres while incoh$\leq$coh .

Claim 2. By Definitions 9 and 6, an integer is associated to each type. Then $\succeq_d$ inherits the properties of the integer $\geq$ relation: reflexivity, transitivity and totality. However, it is not antisymmetric, for instance, {pres,coh} $\succeq_d$ {coh,incoh} and {coh,incoh} $\succeq_d$ {pres,coh} (both have validity distance 1), but {pres,coh} $\neq$ {coh,incoh}. □

Relations $\succeq_v$ and $\succeq_d$ are depicted in Figure 3.

3.3 Hardness

We view hardness as a relation between issues: harder vs. easier issues. The intuition that guides our definition is based on the relations $\succeq_d$ and $\succeq_v$ introduced in the previous section. An issue is 'easy' if the validity labels of the two arguments involved in the issue are far apart by their validity distance: the stronger argument prevails. By means of illustration, an issue where the two arguments are one of type conc and one of type incoh will be easy to decide. It is 'hard' if, vice versa, the validity labels of the two arguments involved in the issue are close by validity distance. The prototypical case consists of arguments with the same validity label. It is in such cases that one can then use relation $\succeq_v$ to distinguish among issues whose arguments have same validity distance. Intuitively, if two issues both involve arguments with the same validity distance—like for instance {pres, coh} and {coh, incoh}—it is the issue involving stronger arguments that is arguably 'easier'.

These intuitions back the definition of hardness as an ordering over issue types, which is based on the lexicographic combination of $\succeq_d$ and $\succeq_v$. In our definition of the hardness ordering, an issue type that is higher in the ordering is 'easier'.

Definition 10 (Hardness ordering) *Let $\mathscr{C}$, $\mathscr{C}'$ be two case models and $\sigma \pm \iota$ and $\sigma' \pm \iota'$ two issues. The hardness ordering $\succeq_h$ is a binary relation over types defined as follows.*

$$T_{\mathscr{C}}(\sigma \pm \iota) \succeq_h T_{\mathscr{C}'}(\sigma' \pm \iota')$$

if and only if:

(1) $T_{\mathscr{C}}(\sigma \pm \iota) \succ_d T_{\mathscr{C}'}(\sigma' \pm \iota')$, or
(2) $T_{\mathscr{C}}(\sigma \pm \iota) \sim_d T_{\mathscr{C}'}(\sigma' \pm \iota')$ and $T_{\mathscr{C}}(\sigma \pm \iota) \succeq_v T_{\mathscr{C}'}(\sigma' \pm \iota')$.

The asymmetric (symmetric) part of $\succeq_h$ is denoted $\succ_h$ ($\sim_h$).

Example 8 In the case model shown in Figure 1, $T_{\mathscr{C}}(P \pm Q) = \{\mathsf{pres}, \mathsf{coh}\}$, $T_{\mathscr{C}}(P \pm R) = \{\mathsf{incoh}, \mathsf{incoh}\}$. By the hardness ordering, $T_{\mathscr{C}}(P \pm Q) \succ_h T_{\mathscr{C}}(P \pm R)$.

We now show that our definition of hardness is well-behaved. In particular, it is transitive and can compare any pair of issue types.

Theorem 1 *The relation $\succeq_h$ is a total order.*

PROOF. We need to prove that $\succeq_h$ is transitive, antisymmetric and total. Let x, y, z be issue types.

$\boxed{\text{Transitivity}}$ Assume $x \succeq_h y$ and $y \succeq_h z$. There are 4 circumstances:

(1) If $x \succ_d y$ and $y \succ_d z$, then by Proposition 4, $x \succ_d z$, hence by Definition 10, $x \succeq_h z$;
(2) If $x \succ_d y$ and $y \sim_d z$, then by Definition 9 $x \succ_d z$, hence by Definition 10, $x \succeq_h z$;
(3) If $x \sim_d y$ and $y \succ_d z$, then by Definition 9, $x \succ_d z$, hence by Definition 10 $x \succeq_h z$;
(4) If $x \sim_d y$ and $y \sim_d z$, then by Definition 10, $x \succeq_v y$ and $y \succeq_v z$. Then by Proposition 4, $x \succeq_v z$, hence $x \succeq_h z$;

$\boxed{\text{Totality}}$ For types x and y, if the validity distance between the arguments in x is not equal to the validity distance in y, then by Definition 10 and Proposition 4, x and y are comparable via $\succeq_h$ because of the relation $\succeq_d$. If the validity distance between the arguments in x is equal to the validity distance in y, then by Definition 10, x and y are comparable via $\succeq_h$ using $\succeq_v$.

$\boxed{\text{Antisymmetry}}$ If $x \succeq_h y$ and $y \succeq_h x$, then by Definition 10, $x \succeq_d y$, $y \succeq_d x$, hence $x \sim_d y$ and therefore $x \succeq_v y$, $y \succeq_v x$. Then using Proposition 4, we find $x = y$. Hence $\succeq_h$ is antisymmetric. □

The hardness ordering is depicted in Figure 4. The formal notion of hardness captured by Definition 10 provides us with a systematic way to categorize easy and hard decisions in case-based reasoning, once represented within the case model formalism. The remaining of the paper is dedicated to putting this formal notion of hardness to the test by further discussing the intuitions underpinning it, and putting it at work in concrete examples of case-based decisions.

3.4 An illustration of the theory

We now give examples that are meant to illustrate all possible types of an issue (summarized in Figure 5). In the next section, we will give a realistic example for discussing the hardness of issues.

We assume there are two possible outcomes for the current situation to be considered. We represent the current situation as `SITUATION`, the two possible outcomes as `OUTCOME_1` and `OUTCOME_2`.

easier ↑

$\{\mathsf{conc}, \mathsf{incoh}\}$
$\{\mathsf{pres}, \mathsf{incoh}\}$
$\{\mathsf{pres}, \mathsf{coh}\}$
$\{\mathsf{coh}, \mathsf{incoh}\}$
$\{\mathsf{pres}, \mathsf{pres}\}$
$\{\mathsf{coh}, \mathsf{coh}\}$
$\{\mathsf{incoh}, \mathsf{incoh}\}$

harder ↓

Figure 4: Hasse diagram of $\succeq_h$.

The issues to consider are as follows:

(1) Whether the current situation should have `OUTCOME_1` or not, represented as `SITUATION` $\pm$ `OUTCOME_1`;
(2) Whether the current situation should have `OUTCOME_2` or not, represented as `SITUATION` $\pm$ `OUTCOME_2`.

The situation is decided based on testimony, where the witness can have expert knowledge, or not. Therefore, depending on the knowledge of the witness, the value of the witness varies. If the witness is an expert (represented as `EXPERT`), the outcome has stronger support than if the person is not an expert (represented as $\neg$`EXPERT`).

Now we focus on issue `SITUATION` $\pm$ `OUTCOME_1`. Figure 5 shows examples of case models for each type of the issue listed in Proposition 3, from which we can see that the cases with an expert witness (`EXPERT`) are more preferred than the cases with $\neg$`EXPERT`, as suggested by the size of boxes.

In the {conc,incoh} model, there is only one case, namely, an expert predicts that the situation should have `OUTCOME_1`. In this model, argument (`SITUATION, OUTCOME_1`) in `SITUATION` $\pm$ `OUTCOME_1` is conclusive, namely all decisions in this model are for having `OUTCOME_1`, which makes `OUTCOME_1` seem like a natural consequence. And there is no support for other decisions.

Comparing the {conc,incoh} model with the {pres,incoh} model, where there are decisions based on both an expert and a non-expert. For issue `SITUATION` $\pm$ `OUTCOME_1` in the {pres,incoh} model, argument (`SITUATION, OUTCOME_1`) has label pres, which makes the decision of `OUTCOME_1` stronger. However, comparing with the same issue in the {conc,incoh} model, it becomes less strong, since there is also a case that implies another outcome (`OUTCOME_2`) in the {pres,incoh} model, even though the argument for `OUTCOME_2` is only with coh. As the cases do not all point to `OUTCOME_1`, issue `SITUATION` $\pm$ `OUTCOME_1` is harder in the {pres,incoh} model.

In the {pres,coh} model, the issue `SITUATION` $\pm$ `OUTCOME_1` is again harder than in the {pres,incoh} model. Argument (`SITUATION, OUTCOME_1`) is still stronger (pres), however, the other argument about `SITUATION` $\pm$ `OUTCOME_1`, i.e., (`SITUATION`, $\neg$`OUTCOME_1`), is not incoherent anymore. There is a case in the model with $\neg$`OUTCOME_1`, which indicates that it is possible that the situation should not have `OUTCOME_1`, even though it is from a non-expert source. The coherent but opposite decisions make the issue become harder.

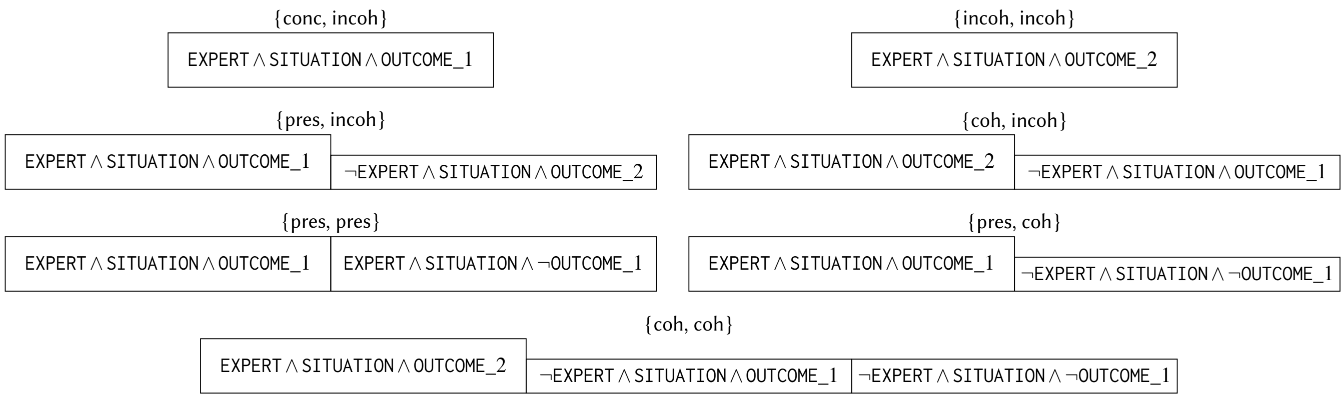

Figure 5: The hardness of SITUATION $\pm$ OUTCOME_1 in 7 case models characterized by the type of issues

In the {coh,incoh} model, SITUATION $\pm$ OUTCOME_1 is harder than in the {pres,coh} model, since for argument (SITUATION, OUTCOME_1), it is less preferable as the testimony for OUTCOME_1 is made by a non-expert. For (SITUATION, ¬OUTCOME_1), there is no support. Comparing with the {pres,coh} model, there is no expert testimony about the issue, which makes the consideration of this issue harder.

In the {pres,pres} model, issue SITUATION $\pm$ OUTCOME_1 is harder than in the {coh,incoh} model, even though the terstimony for OUTCOME_1 is from an expert, namely a more preferable source. This is because its counterargument (SITUATION, ¬OUTCOME_1) is also based on an expert, which makes both having OUTCOME_1 and not having OUTCOME_1 have strong support, hence harder to solve than in the {coh,incoh} model, where there is no one who testifies that the situation should not have OUTCOME_1.

In the {coh,coh} model, issue SITUATION $\pm$ OUTCOME_1 is harder than in the {pres,pres} model. Even though in both of the models, the arguments for having OUTCOME_1 and for not having it are as strong as each other, type {coh,coh} still indicates that the testimonies are from less preferable sources (non-expert), and because of this, the consideration of the issue becomes harder.

Type {incoh,incoh} is the hardest one. As shown in Figure 5, there is completely no decision about issue SITUATION $\pm$ OUTCOME_1, hence the consideration of the issue is the hardest as there is nothing that can be referred to.

4 HARDNESS OVER TIME

In this section, we apply our approach to model case-based decision-making in a real legal domain from the United States, and discuss the development of precedential value in a series of relevant cases by following the research developed by Berman and Hafner [8, 11] and Verheij [21].

The cases we show here are tort cases from New York, which are about car accidents, and which rule should be applied when different jurisdictions are relevant. For instance, when people drive from New York and have an accident in Ontario, which rule should be followed, Ontario's or New York's?

Smith v. Clute 277 N.Y. 407, 14 N.E.2d 455 (1938): The claim was in tort law (driver negligence). The territorial rule applies.

Kerfoot v. Kelley 294 N.Y. 288, 62 N.E.2d 74 (1945): The claim was in tort law (driver negligence). The territorial rule applies.

Auten v. Auten 308 N.Y. 155, 124 N.E.2d 99 (1954): The claim was in contract law (enforce a child support agreement). The center-of-gravity rule applies.

Kaufman v. American Youth Hostels 5 N.Y.2d 1016 (1959): The claim was in tort law (travel guide negligence). The territorial rule applies.

Haag v. Barnes 9 N.Y.2d 554, 175 N.E.2d 441, 216 N.Y.S. 2d 65 (1961): The claim was in contract law (reopen a child support agreement). The center-of-gravity rule applies.

Kilberg v. Northeast Airlines 9 N.Y.2d 34, 172 N.E.2d 526, 211 N.Y.S.2d 133 (1961): The claim was in tort law (common carrier negligence). The territorial rule is partly applied, and there is an exception for the damages part of the case.

Babcock v. Jackson 12 N.Y.2d 473, 191 N.E.2d 279, 473 N.Y.S.2d 279 (1963): The claim was in tort law (driver negligence).

Considering the case model constructed in [21], also shown in Figure 6, where it consists of 7 cases. They are represented by factors for the plaintiff's name (SMITH, KERFOOT, etc.), the year of the decision (1938, 1945, etc.), the kind of case (TORT for a tort case, CONTRACT for a contract case), and the jurisdiction choice rule (TERRITORY for entirely applying the territorial rule, EXCEPTION for partly applying the territorial rule while making an exception for the damages part of the case, and GRAVITY for applying the center-of-gravity rule).

The preference relation among these cases is denoted by the size of the boxes directly, namely, the *Babcock* case is more preferred

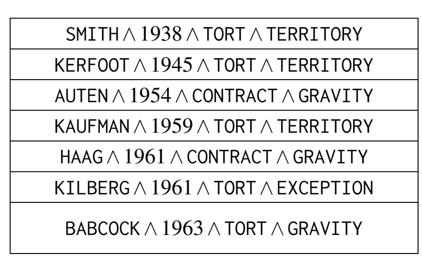

Figure 6: The development of precedential values in cases [21]

than other cases, which are preferentially equivalent. Since the *Babcock* case is a landmark case that overriding previous cases, by which the center-of-gravity approach is established for tort law [11, 21]. We also apply the background theory of all cases in the case model set in [21], namely, the plaintiff names exclude each other pairwise ($\neg(\mathtt{SMITH} \wedge \mathtt{KERFOOT})$, etc.), and similarly for the decision years ($\neg(1938 \wedge 1945)$, etc.), the kinds of cases ($\neg(\mathtt{TORT} \wedge \mathtt{CONTRACT})$) and the choice rules ($\neg(\mathtt{TERRITORY} \wedge \mathtt{EXCEPTION})$, etc.).

We analyze the development of the jurisdiction choice rule by restricting the case model to the cases up and until a particular year. For instance, we write $\mathscr{C}(1954)$ for the set consisting of the three cases *Smith*, *Kerfoot* and *Auten* dating from 1954 or before [21].

The issues that we want to analyze in this series of cases are about the development of the jurisdiction choice rule in tort law cases. They are shown as follows:

(1) $\mathtt{TORT} \pm \mathtt{TERRITORY}$, which is associated with arguments: $(\mathtt{TORT}, \mathtt{TERRITORY})$ and $(\mathtt{TORT}, \neg\mathtt{TERRITORY})$;
(2) $\mathtt{TORT} \pm \mathtt{GRAVITY}$, which is associated with arguments: $(\mathtt{TORT}, \mathtt{GRAVITY})$ and $(\mathtt{TORT}, \neg\, \mathtt{GRAVITY})$;
(3) $\mathtt{TORT} \pm \mathtt{EXCEPTION}$, which is associated with arguments: $(\mathtt{TORT}, \mathtt{EXCEPTION})$ and $(\mathtt{TORT}, \neg\, \mathtt{EXCEPTION})$;
(4) $\top \pm \mathtt{GRAVITY}$, which is associated with arguments: $(\top, \mathtt{GRAVITY})$ and $(\top, \neg\, \mathtt{GRAVITY})$.

Issue $\mathtt{TORT} \pm \mathtt{TERRITORY}$ is about whether a tort law case should entirely apply the territorial rule or not. $\mathtt{TORT} \pm \mathtt{GRAVITY}$ is about whether a tort law case should apply the center-of-gravity rule or not. $\mathtt{TORT} \pm \mathtt{EXCEPTION}$ is about whether a tort law case should partly follow the territorial rule and make an exception for the damages part of the case. And $\top \pm \mathtt{GRAVITY}$ is about the applied status of the center-of-gravity rule in a general sense.

The validity of the arguments listed above has been discussed in [21]. As we show in Section 3, the hardness of an issue is determined by the validity of the arguments that it associates with. For instance, the hardness of issue $\mathtt{TORT} \pm \mathtt{TERRITORY}$ in 1938 with respect to case model $\mathscr{C}(1938)$ is:

$$T_{\mathscr{C}(1938)}(\mathtt{TORT} \pm \mathtt{TERRITORY}) = \{\mathsf{conc}, \mathsf{incoh}\}$$

which is determined by the validity of the following arguments:

$$A_{\mathscr{C}(1938)}(\mathtt{TORT}, \mathtt{TERRITORY}) = \mathsf{conc}$$
$$A_{\mathscr{C}(1938)}(\mathtt{TORT}, \neg\mathtt{TERRITORY}) = \mathsf{incoh}$$

$\mathtt{TORT} \pm \mathtt{TERRITORY}$ becomes harder in 1961, since the hardness of the issue in $\mathscr{C}(1961)$ is $\{\mathsf{pres}, \mathsf{pres}\}$. Notice that according to Definition 4, we consider the labels of arguments $(\mathtt{TORT}, \mathtt{TERRITORY})$ and $(\mathtt{TORT}, \neg\mathtt{TERRITORY})$ in $\mathscr{C}(1961)$ as pres rather than coh.

Based on the validity of the relevant arguments, we summarize the hardness of issues with respect to case models by years in Table 1. The trends of the hardness of issues is shown in Figure 7, from which we have the following observations about the hardness of issues:

(1) When the center-of-gravity rule is introduced in general (1954) and into the tort law domain (1963), the issues about the $\mathtt{GRAVITY}$ rule become harder.
(2) The issues related to the territorial rule become harder correspondingly when the court shows doubt on the rule by making an exception for the damages part of a tort case.
(3) When the center-of-gravity rule is introduced by a landmark case (with higher preference), which makes the rule more preferred, not only the issue about this rule becomes harder, but also other issues are affected (become easier).
(4) In general, after finally making the $\mathtt{GRAVITY}$ rule as the primary one in 1963, the 4 tort law-relevant issues remain the same hardness, as in 1945 when the primary one is the territorial rule. However, more options makes the consideration of applying rules becomes harder.

The first observation can be illustrated by the introduction of the center-of-gravity rule. In 1954, the center-of-gravity rule starts to be considered by the New York courts in a general sense, even it has no effect on the hardness of issues $\mathtt{TORT} \pm \mathtt{TERRITORY}$ and $\mathtt{TORT} \pm \mathtt{GRAVITY}$, it does make issue $\top \pm \mathtt{GRAVITY}$ become harder than in 1945:

$$T_{\mathscr{C}(1954)}(\top \pm \mathtt{GRAVITY}) \succ_h T_{\mathscr{C}(1945)}(\top \pm \mathtt{GRAVITY}).$$

This is because, before 1954, it is clear that the $\mathtt{GRAVITY}$ rule is not considered in the court as argument $(\top, \mathtt{GRAVITY})$ is with incoh and $(\top, \neg\mathtt{GRAVITY})$ is with conc. However, the *Auten* case introduces this rule to the series of case models and makes both of the arguments become presumptive. The introduction not only makes $(\top, \mathtt{GRAVITY})$ stronger and $(\top, \neg\mathtt{GRAVITY})$ weaker but also shortens the validity distance between the validity labels of the two opposite arguments in the issue. Because of the shorter distance, considering whether the center-of-gravity rule should be generally considered or not becomes harder than before. Similarly, after introducing the center-of-gravity rule into the tort law domain (by the *Babcock* case in 1963), we can see the same trend as in 1954.

The second observation is about the exception in a tort law case that applied the territorial rule, which is introduced by the *Kilberg* case in 1961. After this case is added into the model, it has no effect on the hardness of issues that related to the $\mathtt{GRAVITY}$ rule:

$$T_{\mathscr{C}(1961)}(\mathtt{TORT} \pm \mathtt{GRAVITY}) \sim_h T_{\mathscr{C}(1959)}(\mathtt{TORT} \pm \mathtt{GRAVITY});$$
$$T_{\mathscr{C}(1961)}(\top \pm \mathtt{GRAVITY}) \sim_h T_{\mathscr{C}(1959)}(\top \pm \mathtt{GRAVITY}).$$

Both $\mathtt{TORT} \pm \mathtt{TERRITORY}$ and $\mathtt{TORT} \pm \mathtt{EXCEPTION}$ become harder:

$$T_{\mathscr{C}(1961)}(\mathtt{TORT} \pm \mathtt{TERRITORY}) \succ_h T_{\mathscr{C}(1959)}(\mathtt{TORT} \pm \mathtt{TERRITORY});$$
$$T_{\mathscr{C}(1961)}(\mathtt{TORT} \pm \mathtt{EXCEPTION}) \succ_h T_{\mathscr{C}(1959)}(\mathtt{TORT} \pm \mathtt{EXCEPTION}).$$

This is because the exception makes the consideration of the territorial rule in the tort law domain become more complex, as now the courts need to think of whether there will be an exception or not.

The third observation is for the introduction of the landmark case (*Babcock*), introducing the center-of-gravity rule in the tort law domain. This makes issue $\mathtt{TORT} \pm \mathtt{GRAVITY}$ harder:

$$T_{\mathscr{C}(1963)}(\mathtt{TORT} \pm \mathtt{GRAVITY}) \succ_h T_{\mathscr{C}(1961)}(\mathtt{TORT} \pm \mathtt{GRAVITY}).$$

and other relevant issues easier:

$$T_{\mathscr{C}(1963)}(\mathtt{TORT} \pm \mathtt{TERRITORY}) \succ_h T_{\mathscr{C}(1961)}(\mathtt{TORT} \pm \mathtt{TERRITORY});$$
$$T_{\mathscr{C}(1963)}(\mathtt{TORT} \pm \mathtt{EXCEPTION}) \succ_h T_{\mathscr{C}(1961)}(\mathtt{TORT} \pm \mathtt{EXCEPTION});$$
$$T_{\mathscr{C}(1963)}(\top \pm \mathtt{GRAVITY}) \succ_h T_{\mathscr{C}(1961)}(\top \pm \mathtt{GRAVITY}).$$

These trends can be explained from an intuitive perspective. Since from 1963, the $\mathtt{GRAVITY}$ rule becomes primary, for other options, the more preferable way is not applying them, hence make the issues that they associated with easier. But for the $\mathtt{TORT} \pm \mathtt{GRAVITY}$, it becomes harder as we explained in the first observation above.

In the last observation, we find that after the $\mathtt{GRAVITY}$ becomes primary in 1963, all the tort law-relevant issues remain the same hardness as they were before 1954. The only difference is that the

Table 1: Hardness of issues in different years

TORT $\pm$ TERRITORY			TORT $\pm$ GRAVITY		
1938 - 1959	{conc,incoh}	(TORT, TERRITORY): conc (TORT, $\neg$TERRITORY): incoh	1938 - 1959	{conc,incoh}	(TORT, GRAVITY): incoh (TORT, $\neg$GRAVITY): conc
1961	{pres,pres}	(TORT, TERRITORY): pres (TORT, $\neg$TERRITORY): pres	1961	{conc,incoh}	(TORT, GRAVITY): incoh (TORT, $\neg$GRAVITY): conc
1963	{pres,coh}	(TORT, TERRITORY): coh (TORT, $\neg$TERRITORY): pres	1963	{pres,coh}	(TORT, GRAVITY): pres (TORT, $\neg$GRAVITY): coh
TORT $\pm$ EXCEPTION			**$\top \pm$ GRAVITY**		
1938 - 1959	{conc,incoh}	(TORT, EXCEPTION): incoh (TORT, $\neg$EXCEPTION): conc	1938 - 1945	{conc,incoh}	($\top$, GRAVITY): incoh ($\top$, $\neg$GRAVITY): conc
1961	{pres,pres}	(TORT, EXCEPTION): pres (TORT, $\neg$EXCEPTION): pres	1954 - 1961	{pres,pres}	($\top$, GRAVITY): pres ($\top$, $\neg$GRAVITY): pres
1963	{pres,coh}	(TORT, EXCEPTION): coh (TORT, $\neg$EXCEPTION): pres	1963	{pres,coh}	($\top$, GRAVITY): pres ($\top$, $\neg$GRAVITY): coh

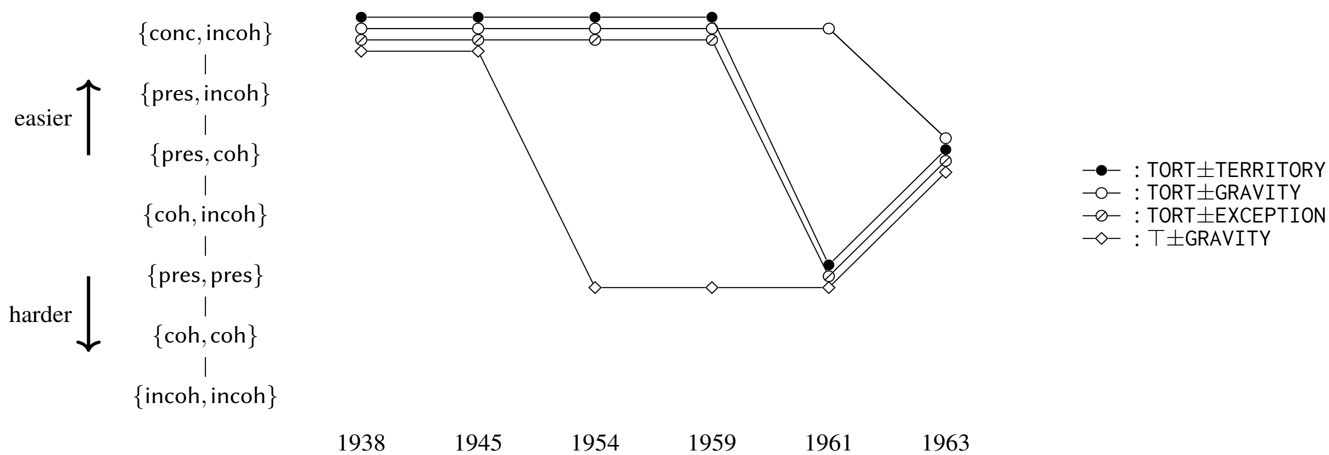

Figure 7: Development of hardness over time for different issues in the series of tort cases

more preferred rule shifts from the TERRITORY rule to the GRAVITY rule. Moreover, we can see that making the choice of which rule should be applied in 1963 is harder than before 1954. This can be connected to our first observation. Before 1954, there is only one jurisdiction choice rule to be considered, as the argument for applying the TERRITORY rule is conclusive (in $\mathscr{C}(1938)$ and $\mathscr{C}(1945)$), and the argument for applying the GRAVITY rule is incoherent (in $\mathscr{C}(1938)$ and $\mathscr{C}(1945)$). Therefore, if a current case is given in the year before 1954, by following the precedents, applying the territorial rule in the current case will be a more preferable choice. In 1963, even though the GRAVITY rule has already become more preferred than other choices, the consideration of which rule should be applied still becomes more complex. This is because there are new cases that introduce new choices about the rule application during this period. More choices make the consideration of the application harder.

5 DISCUSSION

In this section, we position our formal theory of the hardness of case-based decisions with respect to related research.

The development of hardness over time. The dynamics of case-based reasoning has for instance been addressed in [11, 13, 14] in terms of rules, values, and reasons, and there changes in these elements are associated with the handling of hard cases. As we show in Section 4, our approach is also relevant for the dynamics in case-based reasoning. We extend the analysis of a series of New York tort cases in [11, 21] with types of issues, from where we find that even though new cases can strengthen an argument's validity, the hardness of issues may also increase.

Our approach gives insight into the five temporal patterns listed in [11, Section 4.2] (also discussed in [21]), providing a different angle in terms of the hardness typology.

(1) *A general shift in the relative priority of competing purposes.* The *Auten* and *Haag* cases introduce the center-of-gravity rule into the contract cases, and let GRAVITY become a presumptive conclusion in general, whereas it was incoherent. However, argument (TORT, GRAVITY) is not yet coherent. From the issue perspective, the *Auten* case makes the general consideration of the GRAVITY rule ($\top \pm$ GRAVITY) harder (from

{conc,incoh} to {pres,pres}). After 1954, the consideration of the GRAVITY rule becomes more complicate. However, the hardness of handling GRAVITY rule in tort law cases has not changed yet, as TORT $\pm$ GRAVITY is still as hard as before.

(2) *A shift in the relative priority of competing purposes by finding exceptions.* The *Kilberg* case makes that TERRITORY, representing the entire application of the territorial rule, is no longer a conclusive consequence of tort cases, but only a presumptive consequence. From the issue perspective, after the *Kilberg* case is added to the model, the type of issue TORT $\pm$ TERRITORY shifts from {conc,incoh} to {pres,pres}, in the sense that EXCEPTION (partly applying the territorial rule) makes the territorial rule harder to handle.

(3) *The ratio decidendi of an older case is overruled, although it is significantly different.* The example of this pattern discussed in [11, 21] is that the *Babcock* case overrules the *Kaufman* case. The formal case model we use doesn't distinguish tort cases from passenger cases, thus the pattern is not visible here. But as shown in Figure 7, we can still figure out that the landmark *Babcock* case makes the consideration of some issues in a sense harder than right after the *Kaufman* case.

(4) *A case is implicitly overruled.* The rule that applied in tort cases has changed since 1961 because of the *Kilberg* case and the *Babcock* case. The territorial rule is no longer a presumptive conclusion, and the center-of-gravity rule becomes more preferred. If the *Kerfoot* case is decided after *Kilberg* or *Babcock*, it may have come with a different outcome. As shown in Figure 7, both issues are with type {pres,coh} in 1963, but notice that the conclusions of presumptive arguments in the issues are GRAVITY and $\neg$TERRITORY. Furthermore, both issues in 1963 are harder than in the period 1938 $\sim$ 1959, if the *Kerfoot* case is decided after 1963, though it will more likely apply the GRAVITY rule, the decision-making process still becomes more complicated than before.

(5) *A case is explicitly overruled.* As discussed by [8, 11], this pattern occurs rarely, and is not shown in our case model.

Comparing with other research developed recently, in particular Horty's reason model for precedential constraint [14], Henderson and Bench-Capon's model [13] following Levi's idea to consider case-based reasoning as a "moving classification system" [15], the approach we applied here is based on a different mechanism in terms of an elementary propositional logical language. Moreover, their research focuses more on the development of cases and the involved rules, and the hardness of issues as they occur in the decision-making process—the main contribution of this paper—have not been formally analyzed. It seems interesting to follow up on Levi's three stage life cycle of rules (creation, refinement, replacement) as discussed in [13] using the hardness typology in this paper.

The focus of other research on case-based reasoning is often rather different from our paper, and the hardness approach presented here may supplement that research. For instance, [13, 14] emphasize the role of reasons and rules in legal cases, and how they favor the different parties in the court, which may be connected with our approach, in which issues and their hardness are associated with arguments with equal premises, but opposite conclusions.

The approach we present here continues the discussion in [25], where we find that 'using an incoherent argument can make sense and break new ground. A decision based on such an argument can be considered as going beyond the current legal status modeled in the precedent model.' We can further interpret this idea with the results we get from the case study in Section 4. For instance, after introducing the center-of-gravity rule into the tort law domain in 1963 by the *Babcock* case, the validity of the argument for applying the rule in a tort law case, namely (TORT, GRAVITY), shifts from incoh to pres. Even though the validity of the argument becomes stronger, the associated issue TORT$\pm$GRAVITY becomes harder, and the validity distance between the validity labels of the two arguments in the issue is shortened, namely, making the other argument weaker.

It could be interesting to enrich the series of cases discussed in Section 4 to include what happened after the *Babcock* case. Other series of cases that are well-known in AI and Law are also interesting to look at using the hardness theory we developed, for instance, the cases about product liability and privity [3, 13, 15]. Also, since the case model we show in the case study does not have all the possible types in a real legal domain, it can be interesting to investigate whether such a complete case study can be made, in order to better understand the hardness of issues in an actual decision-making process. Natural developments are also to connect our hardness typology in terms of kinds of validity to proof standards [10] and to consider the development of hardness over time in terms of argumentation schemes for case-based reasoning [24]. It would also be interesting to explore the hardness of issues under different preference orderings than significance, for instance, in terms of court levels.

From hardness of issues to easy and hard cases. As discussed by Gardner [9], hard cases is a main topic in law. Rissland summarizes that hard cases in law can arise in three ways [18]:

(1) there exist competing legal rules;
(2) there exist unresolved predicates; and
(3) there exist competing cases.

Our formalism has the potential for modeling the hard cases discussed by Gardner. Competing cases and legal rules can be associated with issues with types {pres,coh}, {pres,pres}, and {coh,coh}. As in these types, the conclusions of arguments involved are opposite to each other, hence form a competing relation. Unresolved predicates can be associated with issues that do not contain conclusive arguments, namely, the same premise can lead to different conclusions. If we treat these predicates as the premises, their indicated meaning as the conclusions, we can analyze the meaning of unresolved predicates as leaving room for debate, hence leading to cases that are harder in the typology. Hence it seems interesting for future research to connect the hardness of issues to insights on easy and hard cases. For instance, there is a connection to Dworkin's famous idea (see e.g. [19, p. 488f.]) that for the perfect, Herculean judge, there is one right solution for all cases, including the hard ones. In our hardness typology, there is a variety of options. Sometimes there is exactly one solution, namely in the types {conc,incoh}, {pres,incoh} and {coh,incoh}. In {incoh,incoh} there is no stare decisis solution. In {pres,pres} and {coh,coh}, there are two equally preferred solutions, and in {pres,coh}, there are two of which one is strictly preferred over the other. Also, consider the characterization of hard cases in [12] that they require an a-rational decision making process. In our typology, this characterization applies to {incoh,incoh}, where there is no solution, and also to {pres,pres} and {coh,coh}, where none of the two choices is preferred. These situations require the construction of a new, persuasive theory of the case and its solution, as suggested by [16].

Also, our hardness theory uses a fixed preference ordering, and it seems relevant to consider a dynamic perspective as a topic of future research in order to address changes legal and societal changes.

Hardness of issues in case-based reasoning with factors. The discussion about current situations with precedents in case-based reasoning with factors, such as HYPO/CATO, is also relevant to our approach about hardness of issues. As discussed in [25], HYPO examples can be modeled in case models in which all cases are equally preferred. In Section 3.2, we show that this kind of case model constrains the possible types of issues. Therefore, issues that occurred in HYPO-style reasoning will have a special hardness typology, a subset of the general typology. Further investigation seems to be in place. For instance, connecting the hardness of issues to argument moves, such as analogizing and distinguishing a current situation with precedents, could be an interesting line of further research.

6 CONCLUSION

In this paper, we model the hardness of case-based decisions in terms of arguments, and their kind of validity. In the approach, we describe a decision-making problem as an issue in a situation in terms of an argument and a counterargument. The hardness of an issue is represented by the validity of the two associated arguments (conclusive, presumptive, coherent, incoherent). We also define an ordering that shows which issues are harder, and which easier. Building on work by Berman and Hafner, we apply our approach to discuss the hardness of the issues that arose in a series of legal cases. It turns out that we can formally show the varying hardness of issues in the temporal development of case-based reasoning.

The hardness approach is relevant in the understanding of case-based decision-making using stare decisis as it formally describes the complexity of decision making in different circumstances. In the discussion, we further suggested that it seems interesting to connect our hardness approach to other research. Although the approach here has been applied to issues in hard cases in law, it could be interesting to consider whether our hardness typology, based on propositional logic, is relevant in other domains where example-based reasoning is relevant (such as in medical diagnosis). As the analysis of argument validity we use, is consistent with probabilistic methods [20], the connection between the hardness of issues and the validity of arguments may also lead to insights on the hardness of decision-making in hybrid AI systems involving both knowledge and data. In this way, the approach can be developed to support the relevance of AI & Law research for AI generally (cf. [23]).

ACKNOWLEDGMENTS

The authors would like to thank the reviewers and Wijnand van Woerkom for their valuable feedback on earlier versions of this paper. 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 Organization for Scientific Research, https://hybrid-intelligence-centre.nl.

REFERENCES

[1] V. Aleven. 1997. *Teaching Case-Based Argumentation Through a Model and Examples.* Ph.D. Dissertation. University of Pittsburgh.
[2] K. D. Ashley. 1990. *Modeling Legal Arguments: Reasoning with Cases and Hypotheticals.* MIT Press, Cambridge.
[3] K. D. Ashley. 2018. Precedent and Legal Analogy. In *Handbook of Legal Reasoning and Argumentation*, G. Bongiovanni, G. Postema, A. Rotolo, G. Sartor, C. Valentini, and D. Walton (Eds.). Springer Netherlands, Dordrecht, 673–710.
[4] K. Atkinson and T. Bench-Capon. 2019. Reasoning with Legal Cases: Analogy or Rule Application? In *Proceedings of the Seventeenth International Conference on Artificial Intelligence and Law*. ACM Press, New York, 12–21.
[5] T. Bench-Capon and G. Sartor. 2003. A Model of Legal Reasoning with Cases Incorporating Theories and Values. *Artificial Intelligence* 150, 1 (2003), 97–143.
[6] D. Berman and C. Hafner. 1991. Incorporating Procedural Context into a Model of Case-based Legal Reasoning. In *Proceedings of the Third International Conference on Artificial Intelligence and Law*. ACM Press, New York, 12–20.
[7] D. Berman and C. Hafner. 1993. Representing Teleological Structure in Case-Based Legal Reasoning: The Missing Link. In *Proceedings of the 4th International Conference on Artificial Intelligence and Law*. ACM Press, New York, 50–59.
[8] D. Berman and C. Hafner. 1995. Understanding Precedents in a Temporal Context of Evolving Legal Doctrine. In *Proceedings of the Fifth International Conference on Artificial Intelligence and Law*. ACM, New York, 42–51.
[9] A. Gardner. 1987. *An Artificial Intelligence Approach to Legal Reasoning.* MIT Press, Cambridge.
[10] T. F. Gordon and D. N. Walton. 2009. Proof Burdens and Standards. In *Argumentation in Artificial Intelligence*, I. Rahwan and G. R. Simari (Eds.). Springer, Berlin, 239–258.
[11] C. Hafner and D. Berman. 2002. The Role of Context in Case-Based Legal Reasoning: Teleological, Temporal, and Procedural. *Artificial Intelligence and Law* 10, 1 (2002), 19–64.
[12] J. Hage, R. Leenes, and A. R. Lodder. 1993. Hard Cases: A Procedural Approach. *Artificial Intelligence and Law* 2, 2 (1993), 113–167.
[13] J. Henderson and T. Bench-Capon. 2019. Describing the Development of Case Law. In *Proceedings of the Seventeenth International Conference on Artificial Intelligence and Law*. ACM Press, New York, 32–41.
[14] J. Horty and T. Bench-Capon. 2012. A Factor-Based Definition of Precedential Constraint. *Artificial Intelligence and Law* 20, 2 (2012), 181–214.
[15] E. Levi. 1948. An Introduction to Legal Reasoning. *University of Chicago Law Review* 15, 3 (1948), 501–574.
[16] L. T. McCarty. 1995. An Implementation of Eisner v. Macomber. In *Proceedings of the Fifth International Conference on Artificial Intelligence and Law*. ACM Press, New York, 276–286.
[17] H. Prakken. 2019. Comparing Alternative Factor- and Precedent-based Accounts of Precedential Constraint. In *Legal Knowledge and Information Systems. JURIX 2019: The Thirty-Second Annual Conference*, M. Araszkiewicz and V. Rodriguez-Doncel (Eds.). IOS Press, Amsterdam, 73–82.
[18] E. Rissland. 1988. Artificial Intelligence and Legal Reasoning: A Discussion of the Field and Gardner's Book. *AI Magazine* 9, 3 (Sep. 1988), 45–55.
[19] G. Sartor. 2005. *Legal reasoning: a cognitive approach to the law. Vol 5 of Treatise on legal philosophy and general jurisprudence*. Springer, Berlin.
[20] B. Verheij. 2016. Correct Grounded Reasoning with Presumptive Arguments. In *15th European Conference on Logics in Artificial Intelligence, JELIA 2016. Larnaca, Cyprus, November 9–11, 2016. Proceedings (LNAI 10021)*, L. Michael and A. Kakas (Eds.). Springer, Berlin, 481–496.
[21] B. Verheij. 2016. Formalizing Value-Guided Argumentation for Ethical Systems Design. *Artificial Intelligence and Law* 24, 4 (2016), 387–407.
[22] B. Verheij. 2017. Formalizing Arguments, Rules and Cases. In *Proceedings of the Sixteenth International Conference on Artificial Intelligence and Law*. ACM, New York, 199–208.
[23] B. Verheij. 2020. Artificial Intelligence as Law. Presidential Address to the Seventeenth International Conference on Artificial Intelligence and Law. *Artificial Intelligence and Law* 28, 2 (2020), 181–206.
[24] A. Wyner and T. J. M. Bench-Capon. 2019. Argument Schemes for Legal Case-based Reasoning. In *Legal Knowledge and Information Systems: JURIX 2007: The Twentieth Annual Conference*, A. R. Lodder and L. Mommers (Eds.). IOS Press, Amsterdam, 139–149.
[25] H. Zheng, D. Grossi, and B. Verheij. 2020. Case-Based Reasoning with Precedent Models: Preliminary Report. In *Computational Models of Argument. Proceedings of COMMA 2020*, H. Prakken, S. Bistarelli, F. Santini, and C. Taticchi (Eds.). Vol. 326. IOS Press, Amsterdam, 443–450.
[26] H. Zheng, D. Grossi, and B. Verheij. 2020. Precedent Comparison in the Precedent Model Formalism: A Technical Note. In *Legal Knowledge and Information Systems. JURIX 2020: The Thirty-third Annual Conference*, S. Villata, J. Harašta, and P. Křemen (Eds.). Vol. 334. IOS Press, Amsterdam, 259–262.