JUSTIFICATION IN CASE-BASED REASONING

Wijnand van Woerkom,[a] Davide Grossi,[b,c,d] Henry Prakken,[a,e] Bart Verheij[b]

September 12, 2022

[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
[e] Faculty of Law, University of Groningen, The Netherlands

INTRODUCTION

Background

- Explainable AI for applications with legal, ethical, or social consequences.
- AI & law interpretation of training examples as cases.

	AI	law
Decision maker	ML system	Court
Precedent	Training data	Decided cases

- An XAI method was developed on the basis of this interpretation.[1]

Present Work

- Extend theory of precedential constraint with notion of justification.
- Show that the top-level model can be seen as an instance of this.

---

[1]Henry Prakken and Rosa Ratsma. "A top-level model of case-based argumentation for explanation: Formalisation and experiments". In: Argument & Computation 13 (2022), pp. 159–194. DOI: 10.3233/AAC-210009.

Background

· Explainable AI for applications with legal, ethical, or social consequences.

· AI & law interpretation of training examples as cases.

	AI	law
Decision maker	ML system	Court
Precedent	Training data	Decided cases

· An XAI method was developed on the basis of this interpretation.[1]

Present Work

· Extend theory of precedential constraint with notion of justification.

· Show that the top-level model can be seen as an instance of this.

---

[1]Henry Prakken and Rosa Ratsma. "A top-level model of case-based argumentation for explanation: Formalisation and experiments". In: Argument & Computation 13 (2022), pp. 159–194. DOI: 10.3233/AAC-210009.

PRECEDENTIAL CONSTRAINT

The Formalism

- The theory of precedential constraint, formalizing a fortiori reasoning in case law.[2]
- Illustrate through example: judging risk of recidivism.

Formal Definition	Recidivism Example
Outcomes	Low risk (0), High risk (1)
Dimension	`Age`, `Priors`, `Sex`
Dimension order	E.g. "People age out of crime"
Fact situation	$(45, 4, M), (30, 5, M)$
Case	$(45, 4, M) : 1$
Case base	$\{(45, 4, M) : 1, (30, 5, M) : 1\}$
Forcing	$(45, 4, M) \preceq_1 (30, 5, M)$

Interpreting Training Examples as Cases

1. Determine dimension orders.
2. Interpret examples as fact situations, and labels as outcomes.

The Formalism

· The theory of precedential constraint, formalizing a fortiori reasoning in case law.[2]

· Illustrate through example: judging risk of recidivism.

Formal Definition	Recidivism Example
Outcomes	Low risk (0), High risk (1)
Dimension	$\texttt{Age}$, $\texttt{Priors}$, $\texttt{Sex}$
Dimension order	E.g. "People age out of crime"
Fact situation	$(45, 4, M), (30, 5, M)$
Case	$(45, 4, M) : 1$
Case base	$\{(45, 4, M) : 1, (30, 5, M) : 1\}$
Forcing	$(45, 4, M) \preceq_1 (30, 5, M)$

Interpreting Training Examples as Cases

1. Determine dimension orders.

2. Interpret examples as fact situations, and labels as outcomes.

[2]John Horty. "Reasoning with dimensions and magnitudes". In: Artificial Intelligence and Law 27.3 (2019), pp. 309–345. DOI: 10.1007/s10506-019-09245-0.

JUSTIFICATION

Definition of Forcing

Given cases a and b we say a *forces* the outcome of b, denoted $a \preceq b$, if

1. a and b have the same outcome s, and
2. b is at least as good for this outcome as a on all dimensions;

$$a(d) \preceq_s b(d) \text{ for all } d \in D.$$

Definition of Justification

A relation $\sqsubseteq$ on cases is called a *justification* relation if it extends the forcing relation:

$$\preceq \subseteq \sqsubseteq.$$

Definition of Forcing

Given cases a and b we say a *forces* the outcome of b, denoted a $\preceq$ b, if

1. a and b have the same outcome s, and
2. b is at least as good for this outcome as a on all dimensions;

$$a(d) \preceq_s b(d) \text{ for all } d \in D.$$

Definition of Justification

A relation $\sqsubseteq$ on cases is called a *justification* relation if it extends the forcing relation:

$$\preceq \, \subseteq \, \sqsubseteq .$$

Recidivism Example

- In our earlier example we saw:

$$(45, 4, M) : 1 \preceq (30, 5, M) : 1.$$

- If instead we had $(25, 4, M) : 1$, the forcing relation would not hold.
- But a justification relation might still hold:

$$(25, 4, M) : 1 \sqsubseteq (30, 5, M) : 1.$$

Relation to Compensation

- The example suggests we (contextually) value priors above age.
- Alternatively: the convict's priors compensate for the age difference.
- A justification relation induces a notion of compensation and vice versa.

Recidivism Example

· In our earlier example we saw:

$$(45, 4, M) : 1 \preceq (30, 5, M) : 1.$$

· If instead we had $(25, 4, M) : 1$, the forcing relation would not hold.
· But a justification relation might still hold:

$$(25, 4, M) : 1 \sqsubseteq (30, 5, M) : 1.$$

Relation to Compensation

· The example suggests we (contextually) value priors above age.
· Alternatively: the convict's priors compensate for the age difference.
· A justification relation induces a notion of compensation and vice versa.

CONCLUSION AND DISCUSSION

Conclusion

- Introduce concept of justification to the theory of precedential constraint
- Show that the top-level model of Prakken & Ratsma can be equivalently formulated in terms of this notion of justification.

Future Work

- Further develop the relation between compensation and justification.
- Relate justification to similar models using notions of compensation.

Discussion

- Is the word 'justification' appropriate?
- How can justification relations be determined in an automated way?
- In the context of XAI, what is the difference between justification and explanation?

Conclusion

- Introduce concept of justification to the theory of precedential constraint
- Show that the top-level model of Prakken & Ratsma can be equivalently formulated in terms of this notion of justification.

Future Work

- Further develop the relation between compensation and justification.
- Relate justification to similar models using notions of compensation.

Discussion

- Is the word 'justification' appropriate?
- How can justification relations be determined in an automated way?
- In the context of XAI, what is the difference between justification and explanation?

Conclusion

· Introduce concept of justification to the theory of precedential constraint
· Show that the top-level model of Prakken & Ratsma can be equivalently formulated in terms of this notion of justification.

Future Work

· Further develop the relation between compensation and justification.
· Relate justification to similar models using notions of compensation.

Discussion

· Is the word 'justification' appropriate?
· How can justification relations be determined in an automated way?
· In the context of XAI, what is the difference between justification and explanation?