ulm university universität ulm

Marvin Schiller & Birte Glimm | 26.02.20

Explaining Ontological Inferences

Outline

Motivation

Knowledge representation essential for

- information management systems,
- intelligent systems,
- cognitive systems,
- Semantic Web,
- …

Logic-Based Knowledge Representation

- Flexible representation of data and schema
- Well-defined semantics

$\rightsquigarrow$ (Automated) deduction support:

  - Derivation of logical consequences, i.e., implicit knowledge is made explicit

- Often via ontologies formalized in Description Logics (DLs)
- DLs underpin the W3C standardized Web Ontology Language OWL

Explanations of Inferences

- Strength of logic-based methods: Logical inferences can be explained

⇝ Justifications: A minimal set of axioms that support a conclusion
  - Problem: Users unfamiliar with logic and inference calculi

⇝ Explanations: *Step-wise* explanations for inferences as *natural language*
  - Problem: Explanations can become very long

  ⇒ Techniques to make explanations more concise

  ⇒ Empirical evidence that verbalizations are (still) understandable

- Related work:
  - [Horridge et al., 2011]: Empirical study on complexity of justifications
  - [Nguyen et al., 2012]: Empirical study on understandability of individual inference steps
  - [Kazakov & Klinov, 2014]: Consequence-based reasoning with tracing

Example

Axioms

$$\begin{aligned} \mathit{EsophagealPathology} &\equiv (\mathit{PathologicalCondition} \sqcap \\ &\quad \exists \mathit{LocativeAttribute}.\mathit{Esophagus}) \\ \mathit{Esophagus} &\sqsubseteq \mathit{GastrointestinalTractBodyPart} \\ \mathit{DigestiveSystemPathology} &\equiv (\mathit{PathologicalCondition} \sqcap \\ &\quad \exists \mathit{LocativeAttribute}.\mathit{GastrointestinalTractBodyPart}) \end{aligned}$$

$$\models$$

Subsumption

$$\mathit{EsophagealPathology} \sqsubseteq \mathit{DigestiveSystemPathology}$$

Verbalization

$$\rightsquigarrow$$

*An esophageal pathology is defined as a pathological condition that is located in the esophagus.
The esophagus is a part of the gastrointestinal tract, thus an esophageal pathology is located in a part of the gastrointestinal tract.
Furthermore, since an esophageal pathology is a pathological condition, an esophageal pathology is a pathological condition that is located in a part of the gastrointestinal tract.
A digestive system pathology is defined as a pathological condition that is located in a part of the gastrointestinal tract. Thus, an esophageal pathology is a digestive system pathology.*

Considered Ontology Language

- Description Logic $\mathcal{EL}$ plus common features underlying the OWL2 EL profile:
  - unsatisfiable concept
  - role inclusion
  - role composition
  - domain axioms
  - disjointness axioms
- OWL2 EL covers a number of large and practically relevant ontologies (e.g. SNOMED CT, GO)
- Considered inferences: Subsumption between concepts
- Challenge: Compute *concise* and *understandable* explanations

Preliminaries – Considered $\mathcal{EL}$ Fragment

Names

Class names: N_C={A, B, C,...}
Individual names: N_I={a, b, c,...}
Role names: N_R={r, s, t,...}

Interpretation

$\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$ where
$\Delta^{\mathcal{I}}$ is a nonempty set of individuals
$\cdot^{\mathcal{I}} : N_C \to \mathcal{P}(\Delta^{\mathcal{I}}), N_I \to \Delta^{\mathcal{I}}, N_R \to \mathcal{P}(\Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}})$

$\mathcal{EL}$ Class Expressions & Axioms

	Syntax	Semantics
Atomic class	C	$C^{\mathcal{I}}$
Universal concept	$\top$	$\Delta^{\mathcal{I}}$
Intersection	$C_1 \sqcap C_2$	$C_1^{\mathcal{I}} \cap C_2^{\mathcal{I}}$
Existential restriction	$\exists r.C_1$	$\{x \in \Delta^{\mathcal{I}} \mid \exists y \in \Delta^{\mathcal{I}} : (x, y) \in r^{\mathcal{I}} \wedge y \in C_1^I\}$
Subsumption axiom	$C_1 \sqsubseteq C_2$	$C_1^{\mathcal{I}} \subseteq C_2^{\mathcal{I}}$

Further Elements from $\mathcal{EL}^{++}$

	Syntax	Semantics
Unsatisfiable concept	$\bot$	$\emptyset$
Nominal	$\{a\}$	$\{a^{\mathcal{I}}\}$
Role composition	$r_1 \circ \ldots \circ r_k \sqsubseteq s$	$r_1^{\mathcal{I}} \circ \ldots \circ r_k^{\mathcal{I}} \subseteq s$
Domain axiom	$dom(r, C_1)$	$r^{\mathcal{I}} \subseteq C_1^{\mathcal{I}} \times \Delta^{\mathcal{I}}$
Disjointness axiom	$disj(C_1, C_2)$	$C_1^{\mathcal{I}} \cap C_2^{\mathcal{I}} = \emptyset$

... where C_1 and C_2 are class expressions

Verbalization Pipeline for a Given Subsumption

1. Compute justifications [Kalyanpur, 2006; Horridge, 2011]

Verbalization Pipeline for a Given Subsumption

1. Compute justifications [Kalyanpur, 2006; Horridge, 2011]
2. Construct rule-based proof
   - Using custom ruleset (includes/modifies rules from ELK and Nguyen et al.)

Verbalization Pipeline for a Given Subsumption

1. Compute justifications [Kalyanpur, 2006; Horridge, 2011]
2. Construct rule-based proof
   - Using custom ruleset (includes/modifies rules from ELK and Nguyen et al.)

For the example:

$$\boldsymbol{R}_{\sqsubseteq\equiv}\,\dfrac{\boldsymbol{R}^{+}_{\sqcap}/\boldsymbol{R}_{5}\,\dfrac{\boldsymbol{R}_{\equiv-}\,\dfrac{\mathrm{EP} \equiv \mathrm{PC} \sqcap \exists \mathrm{loc.E}}{\mathrm{EP} \sqsubseteq \mathrm{PC}} \qquad \boldsymbol{R}_{\exists}/\boldsymbol{R}_{15}\,\dfrac{\boldsymbol{R}_{\equiv-}\,\dfrac{\mathrm{EP} \equiv \mathrm{PC} \sqcap \exists \mathrm{loc.E}}{\mathrm{EP} \sqsubseteq \exists \mathrm{loc.E}} \qquad \mathrm{E} \sqsubseteq \mathrm{GTP}}{\mathrm{EP} \sqsubseteq \exists \mathrm{loc.GTP}}}{\mathrm{EP} \sqsubseteq \mathrm{PC} \sqcap \exists \mathrm{loc.GTP}} \qquad \mathrm{DSP} \equiv \mathrm{PC} \sqcap \exists \mathrm{loc.GTP}}{\mathrm{EP} \sqsubseteq \mathrm{DSP}}$$

(EP: EsophagealPathology, E: Esophagus, PC: PathologicalCondition, GTP: GastrointestinalTractBodyPart, DSP: DigestiveSystemPathology), loc: LocativeAttribute

Verbalization Pipeline for a Given Subsumption

1. Compute justifications [Kalyanpur, 2006; Horridge, 2011]
2. Construct rule-based proof
   - Using custom ruleset (includes/modifies rules from ELK and Nguyen et al.)

For the example:

$$R_{\sqsubseteq\equiv}\,\cfrac{R^{+}_{\sqcap}/R_5\,\cfrac{R_{\equiv-}\,\cfrac{\mathrm{EP} \equiv \mathrm{PC} \sqcap \exists \mathrm{loc}.\mathrm{E}}{\mathrm{EP} \sqsubseteq \mathrm{PC}} \qquad R_{\exists}/R_{15}\,\cfrac{R_{\equiv-}\,\cfrac{\mathrm{EP} \equiv \mathrm{PC} \sqcap \exists \mathrm{loc}.\mathrm{E}}{\mathrm{EP} \sqsubseteq \exists \mathrm{loc}.\mathrm{E}} \qquad \mathrm{E} \sqsubseteq \mathrm{GTP}}{\mathrm{EP} \sqsubseteq \exists \mathrm{loc}.\mathrm{GTP}}}{\mathrm{EP} \sqsubseteq \mathrm{PC} \sqcap \exists \mathrm{loc}.\mathrm{GTP}} \qquad \mathrm{DSP} \equiv \mathrm{PC} \sqcap \exists \mathrm{loc}.\mathrm{GTP}}{\mathrm{EP} \sqsubseteq \mathrm{DSP}}$$

(EP: EsophagealPathology, E: Esophagus, PC: PathologicalCondition, GTP: GastrointestinalTractBodyPart, DSP: DigestiveSystemPathology), loc: LocativeAttribute

3. Convert prooftree to text using patterns

Verbalizing Inference Rules

- Post-order traversal of proof tree
- Rules specify order in which premises are verbalized
- Patterns do not repeat premises verbalized as conclusion of previous steps

Example

$$\boldsymbol{R}_{\sqsubseteq}/\boldsymbol{R}_{12}\ \frac{(1)\ C_1 \sqsubseteq C_2 \qquad (2)\ C_2 \sqsubseteq C_3}{C_1 \sqsubseteq C_3}$$

(1)&(2)	*Since* $\mathrm{v}(C_1 \sqsubseteq C_2)$ *and* $\mathrm{v}(C_2 \sqsubseteq C_3)$ *it follows that* $\mathrm{v}(C_1 \sqsubseteq C_3)$.
(1)	$\mathrm{v}(C_1 \sqsubseteq C_2)$, *therefore being* $\mathrm{v}(C_3)$.
(2)	*Given that* $\mathrm{v}(C_2 \sqsubseteq C_3)$, $\mathrm{v}(C_1 \sqsubseteq C_3)$.
—	*Thus, we have established that* $\mathrm{v}(C_1 \sqsubseteq C_3)$.

Verbalizing Class Expressions and Axioms

Recursively apply verbalization patterns to class expressions. The basic patterns are:

Class Expression	Pattern	Example
A	a[n] A	a mollusc
$A \sqcap B$ [$\sqcap$...]	v(A) that is v(B) [...and...]	a mollusc that is a cephalopod
$\exists r.A$ [$\sqcap$...]	something that r v(A) [... and...]	something that has a shell
$A \sqsubseteq B$	v(A) is v(B)	a clam is a mollusc
$A \equiv B$	According to its definition, v(A) is v(B)	According to its definition, a typical mollusc is [...]

Optimization: Aggregating Class Expressions

- For intersections of class names:
  - $old\_lady \sqsubseteq elderly \sqcap female \sqcap person$
  - ⇝ *"An old lady is an elderly that is a female and a person."*
  - ⇝ *"An old lady is an elderly female person."*
- For existential restrictions on the same role:
  - $Saline \sqsubseteq \exists hasDissolvedWithin.Chlorine \sqcap \exists hasDissolvedWithin.Sodium$
  - ⇝ *"Saline is something that has dissolved within chlorine and something that has dissolved within sodium"*
  - ⇝ *"Saline is something that has dissolved within chlorine and sodium"*
- Furthermore, unnamed concepts are named, if possible (using domain axioms):
  - *"something that owns a dog"*
  - ⇝ *"a person that owns a dog"*

Further Techniques & Heuristics to Promote Conciseness

- Inference rules that represent logical “shortcuts”, e.g.:

Further Techniques & Heuristics to Promote Conciseness

- Inference rules that represent logical “shortcuts”, e.g.:

$$\boldsymbol{R}_{\sqsubseteq\equiv}\ \frac{C_1 \sqsubseteq C_2 \quad C_2 \equiv C_3}{C_1 \sqsubseteq C_3} \quad \text{combines} \quad \boldsymbol{R}_1\ \frac{C_1 \equiv C_2}{C_2 \sqsubseteq C_2} \quad + \quad \boldsymbol{R}_{\sqsubseteq}/\boldsymbol{R}_{12}\ \frac{C_1 \sqsubseteq C_2 \quad C_2 \sqsubseteq C_3}{C_1 \sqsubseteq C_3}$$

Further Techniques & Heuristics to Promote Conciseness

- Inference rules that represent logical “shortcuts”, e.g.:

$$\boldsymbol{R}_{\sqsubseteq\equiv}\ \frac{C_1 \sqsubseteq C_2 \quad C_2 \equiv C_3}{C_1 \sqsubseteq C_3} \quad \text{combines} \quad \boldsymbol{R}_1\ \frac{C_1 \equiv C_2}{C_2 \sqsubseteq C_2} \ + \ \boldsymbol{R}_{\sqsubseteq}/\boldsymbol{R}_{12}\ \frac{C_1 \sqsubseteq C_2 \quad C_2 \sqsubseteq C_3}{C_1 \sqsubseteq C_3}$$

$$\boldsymbol{R}^{+}_{\sqcap}/\boldsymbol{R}_5\ \frac{C_1 \sqsubseteq C_2 \ \dots \ C_1 \sqsubseteq C_{n+1}}{C_1 \sqsubseteq C_2 \sqcap \dots \sqcap C_{n+1}} \qquad \text{(n-ary)}$$

Further Techniques & Heuristics to Promote Conciseness

- Inference rules that represent logical “shortcuts”, e.g.:

$$\mathbf{R}_{\sqsubseteq\equiv}\ \frac{C_1 \sqsubseteq C_2 \quad C_2 \equiv C_3}{C_1 \sqsubseteq C_3} \quad \text{combines} \quad \mathbf{R}_1\ \frac{C_1 \equiv C_2}{C_2 \sqsubseteq C_2} + \mathbf{R}_{\sqsubseteq}/\mathbf{R}_{12}\ \frac{C_1 \sqsubseteq C_2 \quad C_2 \sqsubseteq C_3}{C_1 \sqsubseteq C_3}$$

$$\mathbf{R}^{+}_{\sqcap}/\mathbf{R}_5\ \frac{C_1 \sqsubseteq C_2\ \ldots\ C_1 \sqsubseteq C_{n+1}}{C_1 \sqsubseteq C_2 \sqcap \ldots \sqcap C_{n+1}} \quad \text{(n-ary)}$$

- Skipping of “trivial” inference steps in the verbalization

Further Techniques & Heuristics to Promote Conciseness

- Inference rules that represent logical "shortcuts", e.g.:

$$\mathbf{R}_{\sqsubseteq\equiv}\ \frac{C_1 \sqsubseteq C_2 \quad C_2 \equiv C_3}{C_1 \sqsubseteq C_3} \quad \text{combines} \quad \mathbf{R}_1\ \frac{C_1 \equiv C_2}{C_2 \sqsubseteq C_2} \ + \ \mathbf{R}_{\sqsubseteq}/\mathbf{R}_{12}\ \frac{C_1 \sqsubseteq C_2 \quad C_2 \sqsubseteq C_3}{C_1 \sqsubseteq C_3}$$

$$\mathbf{R}^{+}_{\sqcap}/\mathbf{R}_5\ \frac{C_1 \sqsubseteq C_2 \ \ldots \ C_1 \sqsubseteq C_{n+1}}{C_1 \sqsubseteq C_2 \sqcap \ldots \sqcap C_{n+1}} \qquad \text{(n-ary)}$$

- Skipping of "trivial" inference steps in the verbalization

$$\mathbf{R}^{-}_{\sqcap}/\mathbf{R}_2\ \frac{C_1 \sqsubseteq C_2 \sqcap \ldots \sqcap C_n}{C_1 \sqsubseteq C_i}\ 2 \leq i \leq n$$

(since this would yield, e.g. "A man is a human that is a male."
$\rightarrow$ "A man is a human.")

Further Techniques & Heuristics to Promote Conciseness

- Inference rules that represent logical “shortcuts”, e.g.:

$$\boldsymbol{R}_{\sqsubseteq\equiv} \frac{C_1 \sqsubseteq C_2 \quad C_2 \equiv C_3}{C_1 \sqsubseteq C_3} \quad \text{combines} \quad \boldsymbol{R}_1 \frac{C_1 \equiv C_2}{C_2 \sqsubseteq C_2} \quad + \quad \boldsymbol{R}_{\sqsubseteq}/\boldsymbol{R}_{12} \frac{C_1 \sqsubseteq C_2 \quad C_2 \sqsubseteq C_3}{C_1 \sqsubseteq C_3}$$

$$\boldsymbol{R}^{+}_{\sqcap}/\boldsymbol{R}_5 \frac{C_1 \sqsubseteq C_2 \ \dots \ C_1 \sqsubseteq C_{n+1}}{C_1 \sqsubseteq C_2 \sqcap \ldots \sqcap C_{n+1}} \qquad \text{(n-ary)}$$

- Skipping of “trivial” inference steps in the verbalization

$$\boldsymbol{R}^{-}_{\sqcap}/\boldsymbol{R}_2 \frac{C_1 \sqsubseteq C_2 \sqcap \ldots \sqcap C_n}{C_1 \sqsubseteq C_i} \; 2 \leq i \leq n$$

(since this would yield, e.g. “A man is a human that is a male.” $\rightarrow$ “A man is a human.”)

- Verbalization patterns designed to avoid textual repetitions

Experimental Study: Goals

- Measure adequacy of verbalizations (Is text readable and natural?)
- Measure "understandability"
- Compare two modes of verbalization: proof-based explanation (cf. previous slides) vs. verbalized justifications (axioms verbalized one-by-one), e.g.

A vibrio vaccine is a bacterial vaccine.
A vibrio cholerae vaccine is a vibrio vaccine.
Vaxchora uses a live attenuated pathogen.
Vaxchora is a vibrio cholerae vaccine.
A bacterial vaccine is a vaccine.
According to its definition, a live attenuated vaccine
is a vaccine that uses a live attenuated pathogen.
Thus, vaxchora is a live attenuated vaccine.

Experimental Study: Procedure

- Recruitment of “reliable” *Amazon Mechanical Turk ’Masters’* ($n = 24$)
- Participants judged four verbalized justifications and four explanations each (3–4 inference steps, 3–6 premise axioms)
- Four groups of participants (random assignment), each judged
  - 2 correct and 2 faulty explanations (incl. shortcut rules, skipping)
  - 2 correct and 2 faulty verbalized justifications
- Before the experiment: pre-study questionnaire (English language skills, age, familiarity with used domains), instructions using an example verbalization
- After the experiment: concluding questions (sufficient allocated time)
- Real-world ontologies from diverse domains: food, vaccines, material, (bio)medicine, physiology

You have been endowed with Toby, a cute robot who has his own knowledge base. However, you cannot know if you can trust Toby's reasoning – it can be faulty. Luckily, Toby tries to explain his reasoning, so that you can check up on him. In the following you will be shown examples of how Toby reasons, and you will have to judge his explanations for correctness and appropriateness.

Experimental Study: Instructions

Based on his knowledge base, Toby determines that the following conclusion should hold:

A nautiloid is a typical mollusc.

Toby provides the following explanation for his conclusion:

A nautiloid is a cephalopod.
A cephalopod is a mollusc.
A nautiloid has a shell.
According to its definition, a typical mollusc is a mollusc that has a shell.
Thus, a nautiloid is a typical mollusc.

The statements mentioned by Toby in the underlined part of the explanation are true according to his knowledge base. Toby claims the underlined statements logically entail his conclusion. So, is this the case?

In the example above, Toby's reasoning is logically correct. That is, it is impossible that the statements in Toby's knowledge base are true while the conclusion is false.

However, Toby sometimes makes errors when he reasons. The following example shows what an error could look like.

Conclusion

A sediment is a solid substance.

Explanation

Quartz is a mineral.
A mineral is a solid substance.
It follows that a sediment is a solid substance.

In the example above, Toby's reasoning contains a logical error because his conclusion can be false while the underlined statements from his knowledge base are true. That is, even if quartz is a mineral and even if a mineral is a solid substance, it is still possible that a sediment is not a solid substance.

Experimental Study: Example Explanation

Based on his knowledge base, Toby determines that the following conclusion should hold:

Vaxchora is a live attenuated vaccine.

Toby provides the following **explanation** for his conclusion. Assume that each individual statement mentioned by Toby in the underlined part of the explanation is true according to his knowledge base.

Since vaxchora is a vibrio cholerae vaccine, which is a vibrio vaccine, which is a bacterial vaccine, which is a vaccine, vaxchora is a vaccine.
Furthermore, since vaxchora uses an oral suspension, vaxchora is a vaccine that uses a live attenuated pathogen.
A live attenuated vaccine is defined as a vaccine that uses a live attenuated pathogen. Thus, vaxchora is a live attenuated vaccine.

Experimental Study: Example Questions

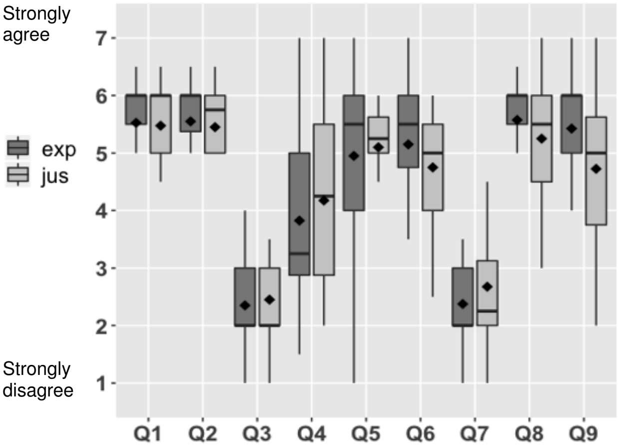
	No	Yes
The presented reasoning is logically correct (i.e. Toby's conclusion is a logical consequence of the available knowledge).	○	○

	Strongly disagree	Disagree	Slightly disagree	Neutral	Slightly agree	Agree	Strongly agree
I am confident with my answer as to whether the reasoning in the explanation is correct.	○	○	○	○	○	○	○
The reasoning presented in the explanation is understandable.	○	○	○	○	○	○	○
The explanation conveys less information than I need to fully understand it.	○	○	○	○	○	○	○
I find that the explanation should be made more concise.	○	○	○	○	○	○	○
The text of the explanation is well-formed (according to writing conventions).	○	○	○	○	○	○	○
I find the explanation easy to read.	○	○	○	○	○	○	○
The explanation or parts of it are ambiguous.	○	○	○	○	○	○	○
It is clear how the statements in the explanation are to be interpreted.	○	○	○	○	○	○	○
The sentences are arranged such that they fit together well.	○	○	○	○	○	○	○

Results

- 71% accuracy for correct vs. incorrect inferences (same for explanations and justifications, mostly incorrect classified as correct $\rightsquigarrow$ not read carefully, 4 participants excluded)
- Results consider logically correct explanations that were accepted as correct

Results: Mean Response

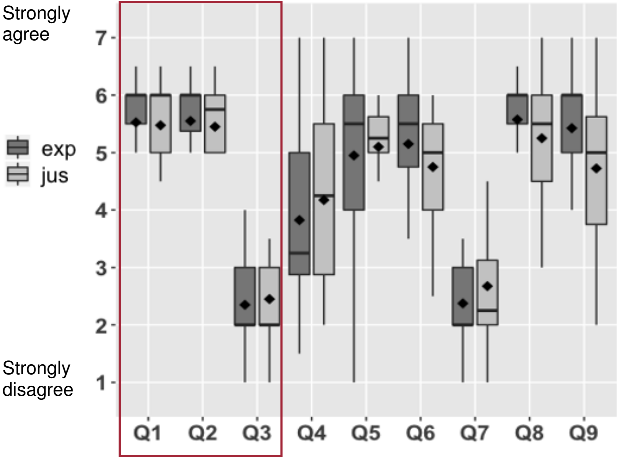


◇ mean values

boxes upper/lower quartile

– median

Results: Mean Response

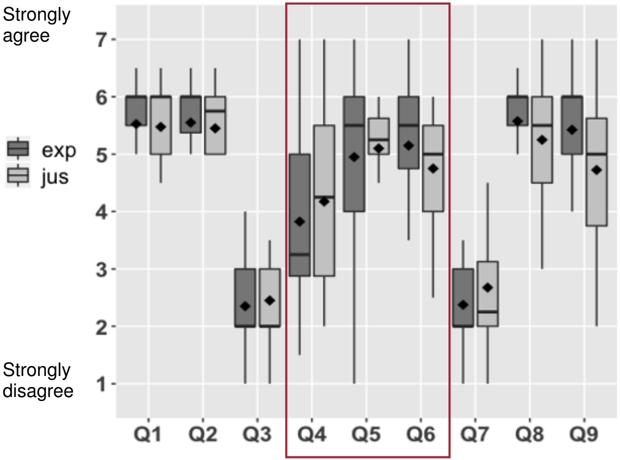


Q1 I am confident with my answer as to whether the reasoning in the explanation is correct.

Q2 The reasoning presented in the explanation is understandable.

Q3 The explanation conveys less information than I need to fully understand it.

Results: Mean Response



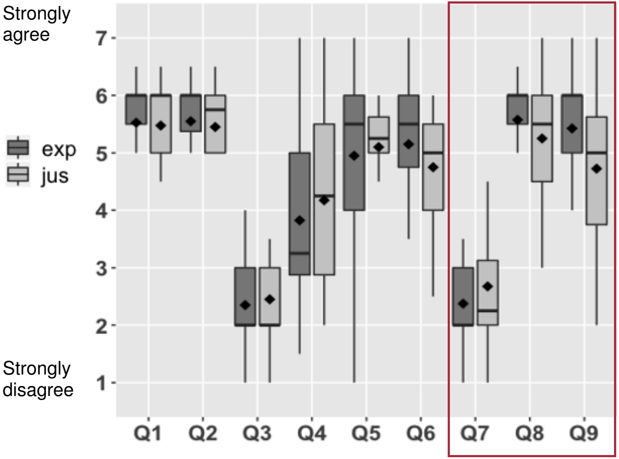
Q4 I find that the explanation should be made more concise.

Q5 The text of the explanation is well-formed (according to writing conventions). (trend* of preference for explanations)

Q6 I find the explanation easy to read. (trend* of preference for explanations)

*In both cases, Wilcoxon signed rank test p-value=0.08

Results: Mean Response



Q7 The explanation or part of it are ambiguous.

Q8 It is clear how the statements in the explanation are to be interpreted.

Q9 The sentences are arranged such that they fit together well. (trend* of preference for explanations)

*Wilcoxon signed rank test p-value=0.08

Conclusion & Discussion

- Shortened explanations are (slightly) preferred
- Still relatively few inferences evaluated with few participants
- Used in a DIY assistant (developed with Robert Bosch GmbH)