

Utilizing iALC to Formalize the Brazilian OAB Exam

Bernardo Alkmim ¹

Alexandre Rademaker ²

Edward Hermann Haeusler ¹

¹ Dep. Informática, PUC-Rio, Brazil

² IBM Research and FGV/EMAp, Brazil

Xaila Workshop, JURIX, Groningen, 2018

Historical Scenario

- ▶ Gentzen, G., 1934/1935, Untersuchungen über das logische Schliessen (Investigation into Logical Inference), Ph.D. thesis, Universität Göttingen.
- ▶ H Kelsen. Pure theory of law, 1934 (2ed 1960). English Edition 1967.
- ▶ KR (Semantic Web) and Proof Theory.
- ▶ How Logic is as important as Ontology in Knowledge Representation.

What is an Ontology?

- ▶ A declarative description of a domain. A set of logical statements that aims to describe a domain completely. A Knowledge Base.
- ▶ Ontology consistency is mandatory, that is, absence of contradictions.
- ▶ Negation is an essential operator.

What does it mean the term “law”?

- ▶ What does count as the unit of law? Open question, a.k.a. The individuation problem.
- ▶ Joseph Raz. *The Concept of a Legal System*, 1970.
- ▶ What is to count as one complete law, Naturally justified law versus Positive Law.

Positive Law

- ▶ According to positivism, law is a matter of what has been posited (ordered, decided, practiced, tolerated, etc.);
- ▶ In a more modern idiom, positivism is the view that law is a social construction
- ▶ The fact that it might be unjust, unwise, inefficient or imprudent is never sufficient reason for doubting its legality
- ▶ Joseph Raz: validity of a law can never depend on its morality

The more corrupt the state, the more laws *Cooruptissima re publica plurimae leges* (Tacito, Annals, Book III, 27)

Natural Law

- ▶ Laws are immanent in nature; that is, they can be discovered or found but not created
- ▶ Law can emerge by the natural process of resolving conflicts, as embodied by the evolutionary process of the common law
- ▶ Whereas legal positivism would say that a law can be unjust without it being any less a law, a natural law jurisprudence would say that there is something legally deficient about an unjust law.

A good judge decides according to justice and right, and prefers equity to strict law. Bonum iudex secundum aequum et bonum iudicat, et aequitatem stricto juri praefert. (Co. Litt, 24.)

Two distinct approaches to the individuation problem

1. Taking all valid statements as in conformance with a declarative statement of an ideal Legally perfect world. This totality is called the law
2. Taking into account all individually legal valid statement as individual laws positively stated and the law is this set

(2) Facilitates the analysis of structural relationship between laws, viz. Primary and Secondary Rules and explicit Grundnorms. Quite adequate to Legal AI.

Why we do not consider Deontic Modal Logic?

- ▶ Deontic Logic does not properly distinguish between the normative status of a situation from the normative status of a norm (rule) (Valente 1995)
- ▶ Norms should not have truth-value, they are not propositions. (General Theory of Norms, Kelsen 1979/1991)
- ▶ An individual law is not a deontic statement, it is not even a proposition. (Kelsen, Alchourrón etc)
- ▶ Deontic logic approach to legal knowledge representation brings us paradoxes

Contrary-to-Duty (or Chisholm's 1963) Paradox

1. It ought to be that Jones goes to the assistance of his neighbors.
 2. It ought to be that if Jones does go then he tells them he is coming.
 3. If Jones doesn't go, then he ought not tell them he is coming.
 4. Jones doesn't go.
- ▶ This certainly appears to describe a possible situation. 1-4 constitute a mutually consistent and logically independent set of sentences.
 - ▶ (1) is a primary obligation, what Jones ought to do unconditionally. (2) is a compatible-with-duty obligation, appearing to say (in the context of 1) what else Jones ought to do on the condition that Jones fulfills his primary obligation. (3) is a contrary-to-duty obligation (CTD) appearing to say (in the context of 1) what Jones ought to do conditional on his violating his primary obligation. (4) is a factual claim, which conjoined with (1), implies that Jones violates his primary obligation.

Standard Deontic Logic (SDL), von Wright 1951

The axioms of SDL:

TAUT all tautologies wffs of the language

OB-K $O(p \rightarrow q) \rightarrow (Op \rightarrow Oq)$

OB-D $Op \rightarrow \neg O\neg p$

MP if $\vdash p$ and $\vdash p \rightarrow q$ then $\vdash q$

OB-NEC if $\vdash p$ then $\vdash Op$

SDL is just the normal modal logic D or KD, with a suggestive notation expressing the intended interpretation.

From these, we can prove the principle that obligations cannot conflict, NC of SDL, $\neg(Op \wedge O\neg p)$.

Contrary-to-Duty Paradox in SDL

1. O_p
2. $O(p \rightarrow q)$
3. $\neg p \rightarrow O\neg q$
4. $\neg p$

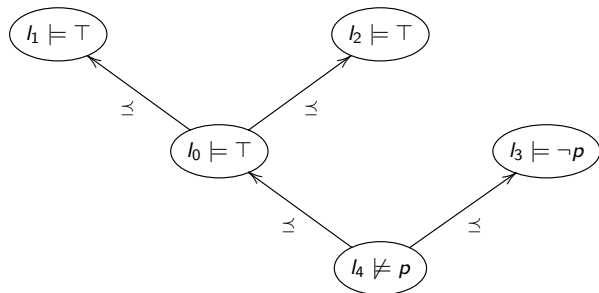
But Chisholm points out

- ▶ from (2) by principle OB-K we get $O_p \rightarrow Oq$,
- ▶ and then from (1) by MP, we get Oq ;
- ▶ but by MP alone we get $O\neg q$ from (3) and (4).
- ▶ From these two conclusions, by PC, we get $Oq \wedge O\neg q$,
contradicting NC of SDL.

Thus 1-4 leads to inconsistency per SDL. But 1-4 do not seem inconsistent at all, so the representation cannot be a faithful one.

An $i\mathcal{ALC}$ model for the Chisholm (ex) paradox

1. The law l_1 , originally Op
2. The law l_2 , originally $O(p \rightarrow q)$
3. From (3), $\neg p \rightarrow O\neg q$, we have $l_3 : \neg p$. If we had $O\neg q \rightarrow \neg p$ the translation would be the same. That is, l_3 is $O\neg q$.
4. The law l_0 that represents the infimum of l_1 and l_2 .



Remember that if $x : A$ then $\forall x' \geq x, x' : A$.

Description Logics

FOL \mapsto Semantic-Network \mapsto Conceptual-Graphs \mapsto DLs

- ▶ Among the best logical frameworks to represent knowledge
- ▶ Binary (Roles) and unary (Concepts) predicate symbols, $R(x, y)$ and $C(y)$.
- ▶ Prenex Guarded formulas ($\forall y(R(x, y) \rightarrow C(y))$, $\exists y(R(x, y) \wedge C(y))$) (decidable fragment of FOL).
- ▶ Non-trivial extensions (transitive Closure R^*).
- ▶ Essentially propositional (Tboxes), but may involve reasoning on individuals (Aboxes).
- ▶ *ALC* can be interpreted as a multi-modal logic \mathcal{K} .

\mathcal{ALC} is the core of DLs

► Syntax :

$$C ::= \perp \mid A \mid \neg C \mid C \sqcap C \mid C \sqcup C \mid \exists R.C \mid \forall R.C$$

$$F ::= C \sqsubseteq C \mid C \equiv C$$

► Semantics :

$$\begin{aligned} \top^{\mathcal{I}} &= \Delta^{\mathcal{I}} \\ \perp^{\mathcal{I}} &= \emptyset \\ (\neg C)^{\mathcal{I}} &= \Delta^{\mathcal{I}} \setminus C^{\mathcal{I}} \\ (C \sqcap D)^{\mathcal{I}} &= C^{\mathcal{I}} \cap D^{\mathcal{I}} \\ (C \sqcup D)^{\mathcal{I}} &= C^{\mathcal{I}} \cup D^{\mathcal{I}} \\ (\exists R.C)^{\mathcal{I}} &= \{a \in \Delta^{\mathcal{I}} \mid \exists b.(a, b) \in R^{\mathcal{I}} \wedge b \in C^{\mathcal{I}}\} \\ (\forall R.C)^{\mathcal{I}} &= \{a \in \Delta^{\mathcal{I}} \mid \forall b.(a, b) \in R^{\mathcal{I}} \rightarrow b \in C^{\mathcal{I}}\} \\ A \sqsubseteq B^{\mathcal{I}} &= A^{\mathcal{I}} \subseteq B^{\mathcal{I}} \end{aligned}$$

Reasoning Algorithms

- ▶ Known proof-procedures (including some industrial Theorem Provers) are based on a specialized FOL Tableaux. Strongly based on individuals even if no ABox is present. (Baader 2003, Horrocks 1998).
- ▶ (McGuinees 96) Presented a Sequent Calculus defined from a standard way from the Tableaux. It has been shown to be not so good for explanation extracting.
- ▶ Proof Theory for Description Logics, (Rademaker 2010).

A T-Box on Family Relationships using *ALCQ*

<i>Woman</i>	\equiv	$Person \sqcap Female$
<i>Man</i>	\equiv	$Person \sqcap \neg Woman$
<i>Mother</i>	\equiv	$\exists hasChild. Person \sqcap Woman$
<i>Father</i>	\equiv	$\exists hasChild. Person \sqcap Man$
<i>Parent</i>	\equiv	$Father \sqcup Mother$
<i>Grandmother</i>	\equiv	$Mother \sqcap \exists hasChild. Parent$
<i>MotherWithoutDaughter</i>	\equiv	$Mother \sqcap \forall hasChild. \neg Woman$
<i>MotherInTrouble</i>	\equiv	$Mother \sqcap (\geq 10 hasChild). \top$

The static part of the trial

- ▶ Considering a jurisprudence basis, classical ALC is not adequate to our approach. We use an intuitionistic version, $iALC$
- ▶ Dealing with the common (deontic) paradoxes
- ▶ A proof-theoretical basis to legal reasoning and explanation
- ▶ laws are inhabitants of a universe that must be formalized
- ▶ Propositions are about laws and not the laws themselves
- ▶ $iALC$ was designed to logically support reasoning on Legal Ontologies based on Kelsen jurisprudence
- ▶ Default $iALC$ is the non-monotonic extension of $iALC$ to deal with the dynamics of legal processes (We will not talk about it today!)

Haeusler, De Paiva, Rademaker (2010-14). See

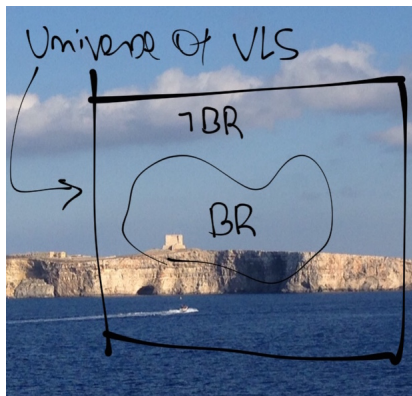
<http://arademaker.github.io/publications/>

Formalization of a Legal System

- ▶ The first-class citizens of any Legal System are VLS. Only VLS inhabit the legal world
- ▶ There can be concepts (collections of laws, VLS) and relationships between VLS. For example: PIL (Private International Law), CIVIL, FAMILY etc, can be concepts. LexDomicilium can be a relationship, a.k.a. a legal connection
- ▶ The relationships between concepts facilitates the analysis of structural relationships between laws
- ▶ The a natural precedence between VLS, e.g. Peter is liable precedes Peter has a renting contract, is modeled as a special relationships between VLS

Intuitionistic vs. Classical Logic (1)

- ▶ The extension of an ALC concept is a set
- ▶ In Brazil, 18 years-old is a legal age. Let BR contains all VLS in Brazil
- ▶ Peter is 17 so Peter is liable is not on BR iff Peter is liable is in the complement of BR
- ▶ Classical negation forces the VLS Peter is liable be valid in some legal system outside Brazil



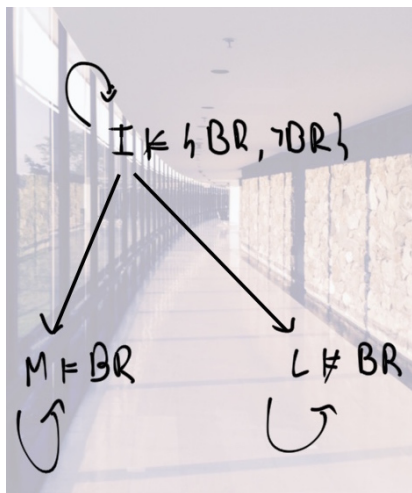
That is, $\phi \sqcup \neg\phi$ is the universe for all ϕ .

Intuitionistic vs. Classical Logic (2)

- ▶ We can have neither Peter is liable $\in BR$ nor Peter is liable $\in \neg BR$. Where $pl \in \neg BR$ means
 - ▶ $pl : \neg BR$
 - ▶ $\mathcal{I}, pl \models \neg BR$
 - ▶ $\forall z. z \geq pl$ we have that $z \not\models BR$
- ▶ There is no z with $z \geq pl$ such that $\mathcal{I}, z \models BR$. There is no VLS in BR dominating Peter is liable

$\models_i \neg A$, iff, for all j , if $i \leq j$ then $\not\models_j A$

$\not\models_i \neg\neg A \rightarrow A$ and $\not\models_i A \vee \neg A$



Comparing with the deontic logic approach

Deontic approach **Laws** must be taken as **propositions?**, or
iALC/Kelsenian approach **Laws** are inhabitants of a universe that
must be formalized, i.e:

Main question

Propositions are about laws or they are the laws themselves?

iALC: a logic for legal theories formalization

- ▶ It can reasoning on individuals (Aboxes), expressed as $i : C$.
- ▶ It is not First-order Intuitionistic Logic. It is a genuine Hybrid logic.

$$C, D ::= A \mid \perp \mid \top \mid \neg C \mid C \sqcap D \mid C \sqcup D \mid C \sqsubseteq D \mid \exists R.C \mid \forall R.C$$

A are general assertions and N nominal assertions for ABOX reasoning. Formulas (F) also includes subsumption of concepts interpreted as propositional statements.

$$N ::= x : C \mid x : N \quad A ::= N \mid xRy \mid x \leq y \quad F ::= A \mid C \sqsubseteq D$$

where x and y are nominals, R is a role symbol and C, D are concepts. In particular, this allows $x : (y : C)$, which is a perfectly valid nominal assertion with x begin its the outer nominal.

iALC Semantics

- ▶ Semantics is Provided by a structure $\mathcal{I} = (\Delta^{\mathcal{I}}, \preceq^{\mathcal{I}}, \cdot^{\mathcal{I}})$ closed under refinement, i.e., $y \in A^{\mathcal{I}}$ and $x \preceq^{\mathcal{I}} y$ implies $x \in A^{\mathcal{I}}$.
- ▶ The interpretation \mathcal{I} is lifted from atomic concepts to arbitrary concepts via:

$$\top^{\mathcal{I}} =_{df} \Delta^{\mathcal{I}}$$

$$\perp^{\mathcal{I}} =_{df} \emptyset$$

$$(\neg C)^{\mathcal{I}} =_{df} \{x \mid \forall y \in \Delta^{\mathcal{I}}. x \preceq y \Rightarrow y \notin C^{\mathcal{I}}\}$$

$$(C \sqcap D)^{\mathcal{I}} =_{df} C^{\mathcal{I}} \cap D^{\mathcal{I}}$$

$$(C \sqcup D)^{\mathcal{I}} =_{df} C^{\mathcal{I}} \cup D^{\mathcal{I}}$$

$$(C \sqsubseteq D)^{\mathcal{I}} =_{df} \{x \mid \forall y \in \Delta^{\mathcal{I}}. (x \preceq y \text{ and } y \in C^{\mathcal{I}}) \Rightarrow y \in D^{\mathcal{I}}\}$$

$$(\exists R.C)^{\mathcal{I}} =_{df} \{x \mid \exists y \in \Delta^{\mathcal{I}}. (x, y) \in R^{\mathcal{I}} \text{ and } y \in C^{\mathcal{I}}\}$$

$$(\forall R.C)^{\mathcal{I}} =_{df} \{x \mid \forall y \in \Delta^{\mathcal{I}}. x \preceq y \Rightarrow \forall z \in \Delta^{\mathcal{I}}. (y, z) \in R^{\mathcal{I}} \Rightarrow z \in C^{\mathcal{I}}\}$$

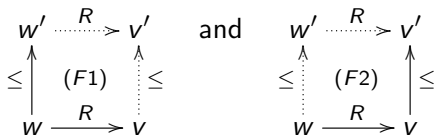
Restrictions on the Interpretations

The structures \mathcal{I} are models for $i\mathcal{ALC}$ they satisfy two frame conditions:

F1 if $w \leq w'$ and wRv then $\exists v'.w'Rv'$ and $v \leq v'$

F2 if $v \leq v'$ and wRv then $\exists w'.w'Rv'$ and $w \leq w'$

The above conditions are diagrammatically expressed as:



OAB Exams

- ▶ The OAB (Ordem dos Advogados do Brasil) Exam is the BAR Exam in Brazil
- ▶ The OAB exams provide an excellent benchmark for the performance of legal information systems, passing the exam would signal capacity of legal reasoning comparable to human lawyers.
- ▶ Interesting problem for explore NLU (Natural Language Understanding) techniques.
- ▶ Initial (shallow) processing, JURIX 2017, was just an starting point for further “deep” language processing.

OAB Exams

- ▶ Only in 2010 were the exams nationally unified.
- ▶ Two stages. We are working on the first stage, multiple choice questions. It has 80 multiple choice questions and each question has 4 options.
- ▶ In order to be approved, candidates need at least a 50% performance.
- ▶ Every year, there are 3 applications of the exam in the country.
- ▶ The exam has a global 80% failure rate. The most recent exam, July 2017, had the highest failure rate, 86% of the candidates failed.

OAB Exams

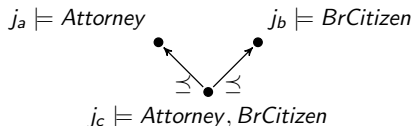
Questions per subject area and their performance rates:

area	#	(%)	area	#	(%)
Ethics	10	65	Constitutional Law	7	42
Consumer's Law	2	56	Civil Procedures	6	40
Children's Law	2	54	Philosophy	2	40
Criminal Procedures	5	47	Labor's Law Proc.	6	40
Regulatory Law	6	47	Criminal Law	6	38
Human Rights	3	47	International Law	2	37
Civil Law	7	44	Business Law	5	33
Environmental	2	43	Taxes	4	42
Labor's Law	5	42			

Data at <http://github.com/oab-exams>: 27 exams with 2220 questions from 2010 to 2018.

Using iALC

- ▶ A concept *Attorney* to represent every VLS for attorneys.
- ▶ Individuals are not real/physical attorneys. There is no world (legal individual) that represents the physical individual *John Doe*, since all that we deal with are laws and legal statements.
- ▶ Take *John Doe* has passed the OAB Exam. We can represent this as $j_a : \textit{Attorney}$ (j_a representing our abstraction over the fact that *John Doe* is a lawyer).
- ▶ As a Brazilian citizen, we can also conclude that we can have $j_b : \textit{BrCitizen}$ (j_b being a **different** legal statement, representing the fact that he is Brazilian).
- ▶ From the intuitionistic aspect of the logic, a Kripke model in it is a Heyting algebra, every pair of worlds has a finite *meet*.



Example 1

Question 50 from 2012 about trademarks.

$Law_{9279} = Art_1 \sqcap \dots \sqcap Art_{244}$. The same happens to the articles with their respective paragraphs.

Regarding trademarks, which option is the correct one?

- A A renowned trademark is a synonym of notoriously known trademark.
- B The period of validity of the register of a trademark is 5 years, renewable by equal and successive periods of time.
- C (correct) The assignment of trademark registration request is allowed, if the transferee meets the legal requirements.
- D The trademark of product or service is the one utilized to identify products or services provided by member of an entity.

Example 1

We want to show that a trademark can be notoriously known and not renowned (a counterexample to option A). The concepts are not synonyms.

- ▶ $HiRe \sqsubseteq BrBrand$ (a trademark to be considered of high renown has to be registered in Brazil) and
- ▶ $NoKn \sqsubseteq (BrBrand \sqcup \neg BrBrand)$ (a trademark can be considered notoriously known having been registered in Brazil or not).

Example 1

$$\frac{\frac{b : \neg BrBrand \quad \neg BrBrand \sqsubseteq \neg HiRe}{b : \neg HiRe} \quad b : NoKn}{b : (\neg HiRe \sqcap NoKn)}$$

$\neg BrBrand \sqsubseteq \neg HiRe$ comes from the counter-positive of what is in article 125.

(?) Art. 125. The trademark registered in Brazil considered of high renown will be guaranteed special protection in all branches of activity.

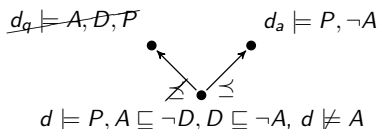
Example 2

Question 6 from the OAB Exam of 2015:

Deise is an Attorney and [she] was elected State Representative. Due to her skills, she was then elected to be part of the Directory of the Legislative Assembly of her state. She willed to conciliate this work with her law practice. According to the Legal Profession Bylaws:

- A Deise's parliamentary activity is incompatible with the [her] practice of the law.
- B (correct) Her participation in the Directory is incompatible with the [her] practice of the law.
- C [her] Being part of the Directory of the Legislative Assembly is compatible with the [her] practice of the law.
- D Deise's parliamentary activity at the Directory can be compatible with the [her] practice of the law on behalf of those in need.

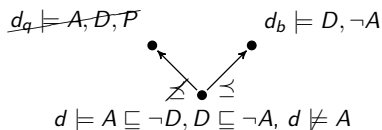
Example 2: A



d_q is our partial knowledge about Deise from the preamble (invalid). But $d \models A, P$ which is compatible with the Law 8906 that says $Attorney \subseteq \neg Directory$ and $Directory \subseteq \neg Attorney$.

The existence of d contradicts (A). If $d \models A$, then d_a would not exist.

Example 2: B



d_q is our partial knowledge about Deise from the preamble (invalid) and d_b is also partial info from item (b). But d satisfy Law 8906 that says *Attorney* $\subseteq \neg$ *Directory* and *Directory* $\subseteq \neg$ *Attorney*.

The existence of d_b validate (B).

Example 3

Luana, Bernardo and Bruno graduated in Law. Luana, 35, already had a management role in a bank when she graduated. Leonardo, 30, is the mayor of Pontal. Bruno, 28, is a police officer at the same County. All three intend to practice private advocacy. Considering the incompatibilities and impediments to the exercise of the lawyers profession:

- A Luana is not prohibited from practicing law because she is employed private institution, and there are no impediments or incompatibilities. . . .
- C (correct) The three graduates, Luana, Leonardo and Bruno, functions incompatible with advocacy, and the prohibition total exercise of private lawyer activities. . . .

Law 8906 article 28.

Example 3

- ▶ $Law_{8906} \equiv Art_1 \sqcap \dots \sqcap Art_{28} \sqcap \dots \sqcap Art_{87}$
- ▶ $Art_{28} \equiv P_1 \sqcap P_2 \dots$
- ▶ The *Lawyer* concept can be read as the set of valid legal statements (VLS) about lawyers.
- ▶ Paragraph I by *Lawyer* $\sqsubseteq \neg ChiefCouncil$ and $ChiefCouncil \sqsubseteq \neg Lawyer$.
- ▶ Paragraph V is formalized by *Lawyer* $\sqsubseteq \neg Police$ and $Police \sqsubseteq Lawyer$.
- ▶ Paragraph VIII is formalized by the two concepts $Lawyer \sqsubseteq \neg Financial$ and $Financial \sqsubseteq \neg Lawyer$.

From the question, we have the hypotheses *lual*: *Lawyer* (Luana acts as lawyer) ...

$$\frac{\frac{\frac{bruno: Police \quad Police \sqsubseteq \neg Lawyer}{bruno: \neg Lawyer}}{[bruno: Lawyer]}}{bruno: \perp}}{\neg(bruno: Lawyer)}$$

THANK YOU !