Lógicas de Descrição com Operadores de Tipicalidade

Igor de Camargo e Souza Câmara

20 de maio de 2020

IME - USP

Introdução

- Primeiro trabalho foi publicado em 2007 [*Giordano, Gliozzi, Olivetti, Pozzato. Preferential Description Logics.*]
- Ainda há trabalhos sendo publicados. [Ex. *Giordano, Gliozzi, Lieto, Olivetti, Pozzato. Reasoning about Typicality and Probabilities in Preferential Description Logics.*]
- Ideias aplicadas a diversas DLs, de $\mathcal{EL}$ a $\mathcal{SHIQ}$.
- Outros grupos de pesquisadores também desenvolveram essa ideias.

Introdução

- Ideia central: incorporar à linguagem um operador de tipicalidade, T.
- Lemos $T(C)$ como "os elementos mais típicos de C".
- Isso permite que façamos inclusões do tipo: $T(C) \sqsubseteq D$, que se lê "*Geralmente, Cs são Ds*" ou "*Os Cs normais são Ds*".
  Exemplo: $T(Passaro) \sqsubseteq Voa$.
- Posteriormente, alguns sistemas incluíram a possibilidade de usar o operador T do lado direito das inclusões: $C \sqsubseteq T(D)$, que se lê "*Os Cs são Ds normais/típicos*".
  Exemplo: $\{maria\} \sqsubseteq T(Estudante)$.

O comportamento de T

- O comportamento de T se espelha nos postulados KLM da lógica condicional *preferencial*, mas pode ser reforçado com o postulado de racionalidade.
- A hierarquia KLM tem as seguintes lógicas: Cumulativa (C), Laço-Cumulativa (CL), Preferencial (P) e Racional (R).
- Cada um desses sistemas fortalece o anterior com um postulado novo.

Postulados KLM

REF.	$A \mid\!\sim A$
LLE.	If $\vdash_{PC} A \leftrightarrow B$, then $\vdash (A \mid\!\sim C) \rightarrow (B \mid\!\sim C)$
RW.	If $\vdash_{PC} A \rightarrow B$, then $\vdash (C \mid\!\sim A) \rightarrow (C \mid\!\sim B)$
CM.	$((A \mid\!\sim B) \wedge (A \mid\!\sim C)) \rightarrow (A \wedge B \mid\!\sim C)$
AND.	$((A \mid\!\sim B) \wedge (A \mid\!\sim C)) \rightarrow (A \mid\!\sim B \wedge C)$
OR.	$((A \mid\!\sim C) \wedge (B \mid\!\sim C)) \rightarrow (A \vee B \mid\!\sim C)$
RM.	$((A \mid\!\sim B) \wedge \neg(A \mid\!\sim \neg C)) \rightarrow ((A \wedge C) \mid\!\sim B)$
LOOP.	$(A_0 \mid\!\sim A_1) \wedge (A_1 \mid\!\sim A_2) ... (A_{n-1} \mid\!\sim A_n) \wedge (A_n \mid\!\sim A_0) \rightarrow (A_0 \mid\!\sim A_n)$
CUT.	$((A \mid\!\sim B) \wedge (A \wedge B \mid\!\sim C)) \rightarrow (A \mid\!\sim C)$

O comportamento de T

- O comportamento de T vai ser determinado por um acréscimo nos modelos usuais para KBs.
- A primeira proposta é uma **função de seleção** $f_T : Pow(\Delta) \rightarrow Pow(\Delta)$, onde Δ é o domínio da interpretação e Pow, o conjunto de suas partes.
- a intuição é que $f_T(C)$ captura os elementos mais típicos de C.
- Ela deve obedecer às seguintes propriedades:
  $(f_T - 1)$ $f_T(S) \subseteq S$
  $(f_T - 2)$ Se $S \neq \emptyset$, então $f_T(S) \neq \emptyset$
  $(f_T - 3)$ [*cautious monotonicity]* Se $f_T(S) \subseteq R$, então $f_T(S) = f_T(S \cap R)$
  $(f_T - 4)$ $f_T(\bigcup S_i) \subseteq \bigcup f_T(S_i)$
  $(f_T - 5)$ $\bigcap f_T(S_i) \subseteq f_T(\bigcup S_i)$

Modelos com Ordens

- Outra maneira de modelar o operador é imbuir o modelo de uma ordem estrita, parcial que satisfaça a *Smoothnes Condition* tal que para todo $S \subseteq \Delta$, $f_T(S) = \min_<(S)$.
  Smoothnes Condition: para todo $S \subseteq \Delta$, há pelo menos um elemento mínimo (i.e. não há cadeias descendentes infinitas).[1]
- Há um *teorema de representação* mostrando que há uma equivalência entre selecionar os elementos mínimos em uma ordem com essas condições e os elementos selecionados pela função seletora descrita anteriormente.

[1] Essa condição equivale à *Limit Assumption* de lógicas de contrafatuais.

Interpretando T através de uma modalidade

- Encarando $<$ como uma relação de acessibilidade, podemos usar $\Box$ da lógica de provabilidade G, de Gödel-Löb, para definir o comportamento do operador T.
- Definimos:
  $(\Box C)^I = \{x \in \Delta | \text{para todo } y \in \Delta \text{ se } y < x \text{ então } y \in C^I\}$
  Ou seja, $\Box C$ designa os elementos que só não são preferenciais com relação a elementos de um determinado conceito.
  Ou seja: $\Box\neg C$ são os elementos que só são preferidos por elementos *fora* de C.
- Assim, designa-se $T(C) := \Box\neg C \sqcap C$
  Elementos de C que só são menores que todos os outros de C (os elementos menores estão todos fora de C).
- Também, para fins de minimização, é possível designar os elementos *anormais* de C: $\neg\Box\neg C \sqcap C$.

Tableaux

- Há sistemas de *tableaux* para $ALC + T$, adaptados dos *tableaux* de ALC.
- Para os modelos minimais (introduzidos posteriormente), também há uma solução de *tableaux*.
- A estratégia geral é combinar o sistema para modelos não minimizados com um segundo *tableaux*, que checa se os modelos encontrados pelo primeiro são minimais.

Monotonicidade

- Essa caracterização torna o operator T não monotônico: $T(C) \sqsubseteq D \not\models T(C \sqcap E) \sqsubseteq D$.
- No entanto, essa relação de consequência ainda não é forte o suficiente para a relação pretendida: não vamos concluir que um pássaro qualquer voa se não for dito, explicitamente, que ele é um pássaro típico.
- Outro problema que essa lógica enfrenta é o **problema da irrelevância**. A não monotonicidade do operador T impede que se mantenham conclusões desejáveis à luz de novas informações - mesmo que elas sejam irrelevantes para o problema.
  Exemplo: $T(Estudante) \sqsubseteq \neg Trabalha$, mas não temos que $T(Estudante \sqcap Alto) \sqsubseteq \neg Trabalha$, que seria desejável.

Monotonicidade

- Para isso, os autores desenvolveram uma **semântica de modelos (canônicos) mínimos**.
  A ideia geral é modelar a relação de consequência apenas pelo que acontece nos modelos que minimizam algumas coisas, abordagem inspirada nos **circumscription patterns**.
  Os primeiros trabalhos minimizavam a **extensão de alguns conceitos**. Uma abordagem posterior, que leva a melhorias de complexidade e torna a operação independente da linguagem, é minimizar o ***rank* dos elementos** do domínio (i.e. o tamanho das cadeias $<$).

Minimização e preferência entre modelos

- A minimização baseada em conceitos consiste na consideração única de modelos mínimos com relação a uma ordem de preferência definida por: $\mathcal{M} < \mathcal{M}'$ sse $\Delta_{\mathcal{M}} = \Delta_{\mathcal{M}'}$ e $\mathcal{M}^{\Box^-}_{\mathcal{L}_T} \subset \mathcal{M}'^{\Box^-}_{\mathcal{L}_T}$ onde $\mathcal{M}^{\Box^-}_{\mathcal{L}_T}$ define os indivíduos anormais para os conceitos da KB.
- Já a minimização baseada no conceito de *rank* dos indivíduos é semelhante.
- **Rank**: $k_{\mathcal{M}}(x) =$ o tamanho da *maior* cadeia $x_0 < \cdots < x$, onde x_0 é elemento mínimo.
- Então, a preferência entre dois modelos se dá quando compartilham domínio e interpretação e:
  (a) Para todo $x \in \Delta$, $k_{\mathcal{M}}(x) \leq k_{\mathcal{M}'}(x)$ **e**
  (b) Para algum $y \in \Delta$, $k_{\mathcal{M}}(x) < k_{\mathcal{M}'}(x)$.
- A minimização define uma lógica mais forte, como, por exemplo $ALC + T_{\text{min}}$.

Rational Closure

- Há outra maneira de caracterizar as conclusões não monotônicas desejáveis: uma adaptação da *Rational Closure*, de KLM.
- Vamos considerar $\overline{TBox}$ a *rational closure* de uma *TBox*, i.e. a extensão da *TBox* com suas conclusões não monotônicas desejadas.
- Essa construção é baseada no conceito de **excepcionalidade**.
  Um conceito C é excepcional para uma base K sse:
  $K \models_{DL+T} T(\top) \sqsubseteq \neg C$.
  O conceito é estendido naturalmente para T-inclusões como:
  $T(C) \sqsubseteq D$ é excepcional para K sse C é excepcional para K.
- Por último, definimos o operador $\mathcal{E}$ como
  $\mathcal{E}(K) := \{\text{conjunto das T-inclusões excepcionais de K}\}$.

Rational Closure

- Com isso, é possível definir uma sequência finita de conjuntos que atribui um *rank* para cada uma das inclusões, revelando uma ordem implícita entre as inclusões não monotônicas:
- $\mathcal{T} = T_0 \supseteq \mathcal{T}_1 \supseteq \ldots$, onde $\mathcal{T}_i = \mathcal{E}(T_{i-1}) \cup \{C \sqsubseteq D | K \models_{DL+T} C \sqsubseteq D\}$.
- Nesse caso poderemos definir o *rank* de uma inclusão como o menor i para o qual ela não é excepcional em $\mathcal{T}_i$.
- Se ela for excepcional para todos os conjuntos, definimos $rank(C) = \infty$.
- Finalmente, a *rational closure* é:
  $\overline{\mathcal{T}} = \{T(C) \sqsubseteq D | rank(C) < rank(C \sqcap \neg D) \text{ ou } rank(C) = \infty\}$
  $\cup \{C \sqsubseteq D : K \models_{DL+T} C \sqsubseteq D\}$

Exemplo de Rational Closure

- Consideremos uma KB que represente o problema clássico do pinguim:
  $KB = \{(1) Pinguim \sqsubseteq Passaro,$
  $(2) T(Passaro) \sqsubseteq Voa,$
  $(3) T(Pinguim) \sqsubseteq \neg Voa, \}$
- *Pinguim* é excepcional. Informalmente: se $x \in Pinguim$ fosse elemento mínimo em Δ, por (1), teríamos $x \in Passaro$ e, por (2) e (3), $x \in Voa$ e $x \in \neg Voa$. $\lightning$
- Disso: $\mathcal{T}_1 = \mathcal{E}(\mathcal{T}_0) = \{Pinguim \sqsubseteq Passaro, T(Pinguim) \sqsubseteq \neg Voa\}$, e $T_2 = \{Pinguim \sqsubseteq Passaro\}$, e termina por aqui.
- Consideremos o conceito adicional e irrelevante: PenasPretas. Nesse caso, $T(Pinguim \sqcap PenasPretas) \sqsubseteq \neg Voa \in \overline{TBox}$.
  Isso porque $Pinguim \sqcap PenasPretas$ tem *rank* 1, ao passo que $Pinguim \sqcap PenasPretas \sqcap Voa$, tem rank 2.

Algumas características dessa abordagem

- A abordagem com o operador T foi aplicada a diversas DLs, tanto leves quanto expressivas.
- Em algumas DLs mais expressivas - ex. *SHIQ* - é possível codificar o operador T dentro da própria DL original, mantendo todos os resultados da DL.
- Há equivalências possíveis entre o operador T e, por exemplo, inclusões *defeasible*, geralmente simbolizadas por $\underset{\sim}{\sqsubset}$.
- Permitem caracterizar explicitamente os elementos típicos, o que não é possível em todas as outras abordagens. Em algumas, é possível através de alguns ajustes.

Algumas características negativas dessa abordagem

- Problema da **herança defeasible/*all-or-nothing***: em geral, é desejável que um indivíduo só não herde as características típicas incompatíveis com ele. Isso significa que *mesmo que um indivíduo não tenha uma ou mais características típicas, ele deveria ter as outras*. Um pinguim não é um pássaro típico, porque não voa, mas, como os demais pássaros, tem penas.
  Como o conceito de tipicalidade é absoluto em $DL + T$, ou um indivíduo herda todas as propriedades típicas, ou não herda nenhuma.
  - Soluções para isso, dentro da tradição do operador T, seriam fechos mais fortes que a rational closure. Há uma proposta baseada na **lexicographic closure**, uma proposta de Lehmann (L do KLM).
  - Abordagens com *defaults*, porque, nesse caso, os únicos *defaults* rejeitados seriam os incompatíveis.

Algumas características negativas dessa abordagem

- Relação de preferência é **global**: se $x < y$, isso vale para *todos* conceitos. Isso significa que não é possível ter um elemento x mais típico que y em C, mas é menos típico em D.
  Um exemplo de porque isso pode ser desejável: dois estudantes atletas. Um é um estudante típico, mas joga xadrez, então é um atleta atípico. O outro é um estudante atípico, porque estuda no período noturno, mas é um atleta típico, porque joga futebol. Esse cenário só pode ser modelado por interpretações que tenham outros elementos no domínio.
    - Soluções possíveis com o operador T: múltiplas relações de preferência, indexadas por algo (podem ser conceitos, como em trabalhos recentes de **Gliozzi**, papéis, ou outros.
      A solução de Gliozzi está em estágio inicial de desenvolvimento e se baseia em uma relação global $<$ que é definida a partir de outras relações $<_C$, indexadas por conceitos.

Algumas características negativas dessa abordagem

- **Problemas com quantificação**: o operador T é baseado em uma teoria proposicional. Por isso, algumas características que vão além da lógica proposicional encontram problemas. Isso pode ser visto no problema de fechamento da $ABox$:
  - Exemplo de [32]: disciplinas de Ciência da Computação geralmente são ministradas por acadêmicos. $(T(CS) \sqsubseteq \forall taught.A)$. Disciplinas de Negócios geralmente são ministradas por consultores $(T(B) \sqsubseteq \forall taught.N)$. Acadêmicos e consultores são classes disjuntas $(A \sqcap N \sqsubseteq \bot)$. Na ABox: *joe* dá um curso de Ciência da Computação $(C1)$ e um curso de Negócios $(C2)$. Há dois modelos incomparáveis e incompatíveis: em um, $C1$ tem *rank* 0; em outro $C2$ tem *rank* 0. Mas não é possível que os dois tenham, simultaneamente, *rank* 0.