

Revisão de Conceitos e Lógica de Descrição

Orientadora: Renata Wassermann

Aluno: Márcio Moretto Ribeiro

13 de Junho de 2005

1 Resumo

A dinâmica dos estados epistêmicos ou teoria da revisão de crenças, como é mais conhecida, foi desenvolvida para descrever como um agente deve mudar suas crenças na presença de novas informações, porém modelos clássicos de revisão de crenças [Gär88] assumem o agente muito idealizado, por exemplo que tendo memória e tempo de raciocínio ilimitados. Agentes reais, porém, claramente não possuem tais propriedades, logo, para se modelar revisão de crenças em agentes com recursos limitados, o modelo de alguma forma deve ser modificado [Was99]. Estudos com bases de crenças [Han97], por exemplo, tratam de conjuntos de crenças finitos tornando a área de revisão de crenças bem mais interessante do ponto de vista computacional.

Lógicas de descrição, por sua vez, tem se destacado como principal formalismo para representação de conhecimento por serem, na prática, eficientes para fazer inferências e ao mesmo tempo suficientemente expressivas para representar conhecimento de forma adequada para uma vasta gama de aplicações. Na literatura tem sido amplamente defendido que o conhecimento deve ser representado *funcionalmente* e não *estruturalmente* [Neb90], para tanto um sistema de inferência deve ser especificado por uma interface do tipo “*Tell & Ask*”.

Em [Was03] foram propostas três formas de se fazer revisão de crenças em agentes menos idealizados, a primeira seria limitando a capacidade de acesso a memória, a segunda limitando o poder de inferência de um agente e por

último tolerando inconsistências. Lógicas de descrição, por serem menos expressivas do que a lógica de primeira ordem, são muitas vezes decidíveis e, na prática, inclusive viáveis computacionalmente [Lev86], logo uma outra forma de se tentar fazer revisão em agentes reais seria diminuindo a expressividade da lógica usada (usando por exemplo alguma lógica de descrição no lugar da lógica de primeira ordem).

Por outro lado seria interessante que a interface “Tell & Ask” dos sistemas de representação de conhecimento pudessem lidar com novas entradas inconsistentes ao seu conjunto de asserções atual ou pudessem remover (interface “Tell, Ask & Forget”) conhecimento de tal conjunto de uma maneira funcional, para tanto é necessário que tais sistemas possam realizar alguma forma de revisão de crenças.

2 Objetivos

O objetivo da iniciação é estudar revisão de crenças, tanto em agentes idealizados como em agentes reais, estudar lógicas de descrição e tentar entender o que significaria fazer revisão de crenças em lógica de descrição, além disso discutir qual é a viabilidade computacional disso, quanto se ganha (em tempo de processamento) e quanto se perde (em expressividade).

3 Cronograma

4 Resumo da Bibliografia

4.1 Lógicas de Descrição

Em vários problemas de inteligência artificial é interessante em algum ponto poder de alguma forma armazenar conhecimento adquirido e extrair conhecimento explícita ou implicitamente armazenado. Para isso é necessário um formalismo que represente simbolicamente o conhecimento e seja capaz de raciocinar sobre esse conhecimento simbolicamente armazenado, ou seja inferir consequências implícitas à base de conhecimento.

As lógicas de descrição tem se tornado o principal formalismo para representação de conhecimento pois possuem uma semântica bem definida (são subconjuntos da lógica de primeira ordem) e além disso são, nos casos computacionalmente interessantes, decidíveis.

O preço pago pela decidibilidade, porém, é uma diminuição na expressividade de tais lógicas [Lev86]. Ultimamente muitos trabalhos tem sido apresentados no sentido de achar as lógicas mais expressivas que ainda se mantêm decidíveis (para mais detalhes consulte [BCM⁺03]).

Uma base de conhecimento em lógica de descrição é tipicamente dividida em duas partes, um *TBox* que contém conhecimento dito “intensional”, ou seja, define as propriedades dos conceitos e um *ABox* que contém conhecimento dito “extensional”, ou conhecimentos específicos dos indivíduos.

Um *TBox* é constituído de axiomas como:

- $homem \sqsubseteq pessoa$
- $pai \equiv homem \sqcap \exists temfilho.pessoa$

Ou seja, todo pai é uma pessoa e um pai é um homem que tem (pelo menos) um filho que é pessoa.

A principal forma de inferência que se pode fazer em uma terminologia (*TBox*) é verificar se um determinado conceito é mais geral do que outro, ou seja se $C \sqsubseteq D$, tal forma de inferência é chamada de “concept subsumption”. Todo sistema gerenciador de conhecimento em lógica de descrição deve ser capaz de fazer tal inferência.

Um *ABox* contém sentenças sobre indivíduos, podem ser afirmações sobre conceitos como:

- Mulher(BIANCA)

ou sobre relações entre indivíduos (“roles”) como:

- irmão(MÁRCIO, CAIO)

A principal inferência que um *ABox* deve ser capaz de realizar é chamada de *instanciação*, ou seja, verificar se determinado indivíduo é ou não instância de um conceito.

4.2 Revisão de Crenças

Revisão de crenças consiste no estudo dos *estados epistêmicos* e sua dinâmica, fornecendo uma representação para os *elementos epistêmicos* (crenças) e um *critério de racionalidade*. Nos focaremos primeiramente na representação baseada em *conjuntos de crenças*, para tal nos baseamos em [Gär88].

4.2.1 Conjuntos de Crenças

Supomos definida uma lógica L com a idéia de *consequência lógica* e *consistência* definidas.

Usaremos letras maiúsculas $A, B, C \dots$ para representar variáveis e e os símbolos \top para uma tautologia e \perp para contradição.

A *atitude epistêmica* relacionada com *conjuntos de crenças* atribui a cada sentença A umas das três possibilidades:

- A é aceita
- A é rejeitada
- A é indeterminada

Definição 1 *Critérios de racionalidade:*

- O conjunto de sentenças aceitas é consistente
- As consequências lógicas de sentenças aceitas devem ser aceitas

Ambas asserções são muito idealizadas, mas servem como bons pontos de partida. Posteriormente assumiremos um critério mais fraco e consequentemente mais palpável quando falarmos de bases de crenças.

Definição 2 *Um conjunto de sentenças K é chamado de conjunto de crenças sse:*

- K for logicamente fechado
- K for consistente

Definição 3 Podemos escrever as atitudes epistêmicas usando a linguagem dos conjuntos de crença. Sejam K nosso conjunto de crenças e A uma sentença qualquer:

- A é aceita em $K \equiv A \in K$
- A é rejeitada em $K \equiv \neg A \in K$
- A é indeterminada em $K \equiv A \notin K$ e $\neg A \notin K$

4.2.2 Expansão, Revisão e Contração

Voltamos agora nossa atenção para a *dinâmica dos estados epistêmicos*. Uma mudança em um *conjunto de crenças* pode ser de três tipos:

(expansão) A era indeterminada e A ou $\neg A$ passa a ser aceita

(contração) A era aceita e passa a ser indeterminada

(revisão) A era aceita e $\neg A$ passa a ser aceita

Nosso objetivo agora é justificar uma escolha de axiomas para *expansão*, *contração* e *revisão* baseados no critério de *mudança mínima* [Har86]:

“Quando trocamos crenças devido a novas evidências, devemos manter o máximo das crenças antigas possível.”

4.2.3 Expansão

A expansão modela a *mudança epistêmica* que representa aprender algo novo (que não contradiz nada em que acreditávamos anteriormente). É o caso em que A (ou $\neg A$) era indeterminado e passa a ser aceito.

Seja K nosso conjunto de crenças, ao adicionar uma sentença nova a ele devemos também adicionar todas as conseqüências dessa nova sentença junto com nossas crenças antigas, ou seja:

$$K + A = Cn(K \cup A) \tag{1}$$

4.2.4 Revisão

Devemos fazer uma revisão no nosso conjunto de crenças toda vez que passamos a acreditar em algo que contradiz com nosso estado presente de crenças, para isso temos que abandonar algumas de nossas crenças antigas (para manter a consistência), ou seja quando revisamos nosso conjunto nem todas as antigas informações são mantidas (a revisão é *não monotônica*).

Segundo o paradigma AGM uma revisão deve satisfazer:

(K*1) $K * A$ é um conjunto de crenças

(K*2) $A \in K * A$

(K*3) $K * A \subseteq K + A$

(K*4) Se $\neg A \notin K$, então $K + A \subseteq K * A$

(K*5) Se $K * A = K_{\perp}$ sse $\vdash \neg A$

(K*6) Se $\vdash A \leftrightarrow B$, então $K * A = K * B$

Esses axiomas são chamados de *axiomas básicos* da revisão. Os próximos postulados mostram a relação entre revisão e o conectivo \wedge :

(K*7) $K * (A \wedge B) \subseteq (K * A) + B$

(K*8) Se $\neg B \notin K * A$, então $(K * A) + B \subseteq K * (A \wedge B)$

4.2.5 Contração

Sempre que desejamos desistir de uma crença devemos fazer uma *contração*. Repare que se desejamos tirar uma sentença A de nosso conjunto de crenças não basta apenas tirar A , pois o conjunto remanescente pode implicar em A e simplesmente tirá-la implicaria que o novo conjunto não seria logicamente fechado.

Esses são os axiomas que,7 segundo o paradigma AGM, uma contração deve obedecer:

(K^-1) $K - A$ é um conjunto de crença

(K^-2) $K - A \subseteq K$

(K^-3) Se $A \notin K$, então $K - A = K$

(K^-4) Se $\not\vdash A$, então $A \notin K - A$

(K^-5) Se $A \in K$, então $K \subseteq (K - A) + A$

(K^-6) Se $A \leftrightarrow B$, então $K - A = K - B$

Como na *revisão* esses foram os axiomas básicos e os próximos são os que dizem respeito a conjunção.

(K^-7) $(K - A) \cap (K - B) \subseteq K - (A \wedge B)$

(K^-8) Se $A \notin K - (A \wedge B)$, então $K - (A \wedge B) \subseteq K - A$

4.2.6 Teoremas de Levi

Mostramos agora dois teoremas que mostram a ligação entre *revisão* e *contração*.

Teorema 1 *Se uma contração satisfaz (K^-1)-(K^-8) então a revisão definida como: $K * A = (K - (\neg A)) + A$ satisfaz (K^*1)-(K^*8).*

Teorema 2 *Se uma contração satisfaz (K^*1)-(K^*8) então a revisão definida como: $K - A = K \cap (K * \neg A)$ satisfaz (K^-1)-(K^-8).*

Repare que só precisamos, então, definir uma das funções (contração ou revisão), pois a outra pode ser definida através de algum desses teoremas.

4.2.7 Partial Meet Contraction

Os postulados descritos tanto para *revisão* quanto para *contração* não caracterizam totalmente uma função, para construirmos uma *função contração* precisamos de mais informações sobre o conjunto. Tentaremos agora mostrar uma possível maneira de se construir uma *função contração* que satisfaça os postulados propostos, primeiramente trataremos da *partial meet contraction*.

Definição 4 *Seja K um conjunto de crenças, K' é chamado de subconjunto maximal que não implica A sse ¹:*

- $K' \subseteq K$
- $A \notin K'$
- Se $K' \subset K'' \subseteq K$ então $K'' \vdash A$

Escolhendo alguns conjuntos de $K \perp A$ e fazendo a intersecção deles, temos o que chamamos de *partial meet*.

Definição 5 *Mais formalmente, seja K um conjunto de sentenças uma função de seleção para K é uma função γ tal que para toda sentença A :*

- Se $K \perp A \neq \emptyset$ então $\gamma(K \perp A) \neq \emptyset$ e $\gamma(K \perp A) \subseteq K \perp A$

Definição 6 *Sejam K um conjunto de sentenças, A uma sentença e γ uma função seleção então chamamos o conjunto $\bigcap \gamma(K \perp A)$ de *partial meet contraction* em K gerada a partir de γ .*

Teorema 3

Uma contração é partial meet se e somente se satisfizer (K^{-1}) - (K^{-6}) .

Definimos agora uma escolha mais específica de conjuntos de $K \perp A$.

Definição 7 $S(K \perp A) = \{K' \in K \perp A : K'' \leq K' \text{ para todo } K'' \in K \perp A\}$

¹Chamaremos K' de $K \perp A$

Onde \leq é simplesmente uma relação definida para união de todos os conjuntos de $K \perp A$, chamamos a contração definida por $\bigcap S(K \perp A)$ de *relational partial meet*. E então podemos mostrar que:

Teorema 4 *Uma contração é relational partial meet se e somente se satisfizer (K^-7) .*

Por fim se exigirmos que \leq seja uma relação transitiva e chamamos essa contração de *transitively relational partial meet*, então temos que:

Teorema 5 *Uma contração é transitively relational partial meet se e somente se satisfizer (K^-8) .*

4.2.8 Sistemas de Esferas

Um outro modelo para a dinâmica dos estados epistêmicos é o chamado *sistemas de esferas* [Gro88]. Nesse modelo ao invés de considerarmos o conjunto das crenças consideramos os mundos possíveis onde nossas crenças são verdadeiras.

Formalmente um mundo possível é um conjunto consistente maximal $(L \perp^\perp)$. Seja K um conjunto de crenças, representaremos por $[K]$ o conjunto dos mundos possíveis onde os elementos de K são verdade e usarem o símbolo $[\alpha]$ como abreviação para $[C(\alpha)]$.

Os sistemas de esferas nos fornecem a vantagem de fornecer uma visualização interessante ao problema de revisão de crenças.

Se $[K] \cap [\alpha] \neq \emptyset$ então existem mundos possíveis com respeito a K e a α ao mesmo tempo, ou seja, $K \cup \{\alpha\}$ é consistente, logo $K * \alpha = K + \alpha$ e isso equivale a $[K] \cap [\alpha]$. Se por outro lado $[K] \cap [\alpha] = \emptyset$ então K e α são inconsistentes, nesse caso fazer a revisão $K * \alpha$ é equivalente a pegar um subconjunto de $[\alpha]$.

O resultado mais interessante sobre sistemas de esferas é que uma revisão definida como acima é equivalente a uma revisão que satisfaça os axiomas $(K * 1)$ - $(K * 6)$ [Han97].

Um refinamento a esse modelo é definir um conjunto de conjuntos de mundos possíveis ordenados com respeito a relação de inclusão (“fallbacks”) e cujo menor elemento é $[K]$. Nesse caso uma revisão seria a intersecção entre $[\alpha]$ e o “fallback” mais interno possível de forma que essa intersecção não seja vazia. Tal revisão é equivalente a uma satisfazendo $(K * 1) - (K * 8)$.

Em [WF99] foi proposta uma forma de interpretar o sistema de esferas com o intuito de se fazer revisão de conceitos. Ao invés de pensar nas esferas como mundos possíveis com respeito a determinado conjunto de crenças elas representariam objetos possíveis de acordo com um determinado conjunto de crenças sobre o objeto.

A idéia seria representar um conceito (as crenças sobre ele) com um par de conjuntos $[P]$ e $[K]$ (complexo) onde $[P] \subset [K]$. O conjunto K representaria os objetos possíveis enquanto P representaria o conjunto dos protótipos possíveis de um objeto (os objetos mais “comuns”). Nesse trabalho foram propostas formas de se fazer revisão de crenças nesse complexo além e foram provados teoremas de representação, porém o trabalho só se preocupou com a semântica do problema, ou seja, nenhuma lógica foi explicitamente usada para representar os conceitos.

Referências

- [BCM⁺03] Franz Baader, Diego Calvanese, Deborah McGuinness, Daniele Nardi, and Peter Patel-Schneider, editors. *The Description Logic Handbook*. Cambridge University Press, 2003.
- [Gro88] A. Grove. Two modellings for theory change. *Journal of Philosophical Logic*, 1988.
- [Gär88] Peter Gärdenfors. Knowledge in flux - modeling the dynamics of epstemic states. *MIT Press*, 1988.
- [Han97] Sven-Ove Hanson. *A Textbook of Belief Dynamics*. Kluwer Academic Publishers, 1997.
- [Har86] G. Harman. Change in view: Principles of reasoning. *MIT Press*, 1986.

- [Lev86] Hector J. Levesque. Knowledge representation and reasoning. *Annual Reviews*, 1986.
- [Neb90] Bernhard Nebel. *Reasoning and revision in hybrid representation systems*. Springer-Verlag New York, Inc., New York, NY, USA, 1990.
- [Was99] Renata Wassermann. Resource bounded belief revision, 1999.
- [Was03] Renata Wassermann. Generalized change and the meaning of rationality postulates. *Studia Logica*, 2003.
- [WF99] Renata Wassermann and Eduardo Fermé. A note on prototype revision. *Spinning Ideas*, 1999.