

Reconhecimento de Intenções para uma Interação Robusta com Bases de Conhecimento Médicas

Paulo Quaresma*
pq@fct.unl.pt

José Gabriel Lopes
gpl@fct.unl.pt

Centro de Inteligência Artificial, UNINOVA
2825 Monte da Caparica

Resumo

Uma interação robusta em Língua Natural implica, por parte de cada interveniente, a capacidade de inferência das crenças, intenções e planos dos outros agentes. Neste artigo é apresentada uma proposta de uma arquitectura que suporta o reconhecimento de intenções através de inferência abductiva utilizando a informação veiculada pelos diversos agentes e o conhecimento do domínio específico da interacção. A abdução das intenções é efectuada utilizando o formalismo e a semântica proposta por Kakas e Mancarella ([7]) que é uma extensão à semântica dos modelos estáveis de Gelfond e Lifschitz para a programação em lógica ([5]). O artigo descreve o modo como foi utilizado este formalismo para suportar uma interacção robusta com bases de conhecimento médicas. Finalmente, são discutidos alguns dos problemas em aberto e é apontado algum do trabalho a realizar.

1 Introdução

O reconhecimento das diferentes atitudes (crenças, intenções, conhecimentos, dúvidas, desejos, etc.) dos diversos intervenientes em diálogos é fundamental para permitir uma participação apropriada e construtiva

*Este trabalho foi suportado em parte pela JNICT, bolsa de estudo n^o PMCT/BD/1766/91-IA.

nesse mesmo diálogo ([9, 8]). Os sistemas de informação com capacidade de interacção em Língua Natural necessitam, portanto, de modelar as atitudes dos utilizadores para poderem suportar actos de fala indirectos, elipses e fornecer informações mais específicas do que as solicitadas pelos seus utilizadores.

Nos últimos anos diversos trabalhos têm sido feitos sobre o modo como um sistema de informação pode reconhecer e inferir algumas das atitudes dos utilizadores. Entre as abordagens utilizadas é de realçar o trabalho de Litman ([8]) em que os planos dos utilizadores são inferidos utilizando estratégias de planeamento em que as acções são representadas de acordo com o modelo desenvolvido no sistema STRIPS ([4]) e NOAH ([12]). Neste modelo cada acção é constituída por uma cabeça, uma lista de pré-condições, uma lista de efeitos, uma lista de restrições e um conjunto de sub-acções ou sub-objectivos. A inferência dos planos é feita, recorrendo a uma biblioteca pré-existente de planos, a possíveis objectivos do utilizador e a regras heurísticas, de modo a ser construída uma lista ordenada de acções que, assume-se, representam as intenções dos utilizadores.

A abordagem apresentada neste artigo é distinta da abordagem de planeamento clássica, na medida em que a inferência das atitudes é efectuada através de raciocínio não-monotónico, nomeadamente por mecanismos de abdução ([7, 10, 11, 3, 1, 6]). O raciocínio abductivo é efectuada utilizando

o formalismo e a semântica proposta por Kakas e Mancarella que é uma extensão à semântica dos modelos estáveis de Gelfond e Lifschitz para a programação em lógica. Neste formalismo o sistema é descrito por cláusulas Horn e as restrições de integridade são fórmulas fechadas de primeira ordem. O modelo estável obtido incorporará as atitudes dos diversos agentes que foram inferidas.

Na secção 2 é feita uma breve descrição do sistema proposto, com um especial ênfase na definição do modelo de abdução utilizado e nos operadores que descrevem o estado mental dos agentes ([2]). Na secção 3 é apresentada um sistema que utiliza raciocínio abduutivo para inferir as intenções dos utilizadores no acesso a bases de conhecimento médicas. Finalmente, na secção 4 é feita uma análise do trabalho efectuado e dos problemas em aberto, sendo apontado algum trabalho futuro.

2 Inferência abduativa de intenções

A inferência das intenções dos diversos agentes intervenientes numa interacção em Língua Natural é feita a partir de um conjunto de regras e de factos que modelizam o domínio específico da aplicação e de um conjunto de restrições de integridade.

O sistema recorre à semântica formal para a abdução definida por Kakas e Mancarella ([7]) a partir da semântica dos modelos estáveis de Gelfond e Lifschitz ([5]).

Formalmente o sistema é definido por um triplo ordenado $\langle P, A, IC \rangle$ em que:

- P é um conjunto de cláusulas da forma $H \Leftarrow L_1, \dots, L_k$, com $k > 0$, sendo H um átomo e L_i , com $1 \leq i \leq k$, um literal;
- A é um conjunto de predicados abduzíveis;
- IC , as restrições de integridade, são um conjunto de fórmulas fechadas.

O modelo estável obtido a partir deste sistema incorporará as atitudes dos diversos agentes. A modelização das atitudes dos agentes é baseada no trabalho de Cohen e Levesque ([2]) e na simplificação descrita por Appelt e Pollack ([1]). Neste sentido foram definidos quatro operadores:

- $INT(a, \alpha)$: agente a tem como intenção α ;
- $BEL(a, p)$: agente a acredita que p ;
- $ACH(a, p)$: agente a acredita que p será verdadeiro em consequência das suas acções;
- $EXP(a, p)$: $BEL(a, p) \vee ACH(a, p)$; o agente a acredita que p é verdadeiro ou acredita que será verdadeiro em consequência das suas acções;

Foi também definido um operador que permite construir acções mais complexas:

- $TO(\alpha, p)$: o plano que consiste em fazer α de modo a tornar p verdadeiro.

Para além destes operadores, é necessária a definição de algumas regras Horn que definem o comportamento do sistema. No contexto deste artigo é suficiente a regra que diz que, se um agente a acredita que o estado p é um efeito da acção α e se ele tem como intenção fazer a acção α e atingir o estado p , então a sua intenção ao fazer a acção α é atingir o estado p :

$$1. INT(a, TO(\alpha, p)) \Leftarrow BEL(a, TO(\alpha, p)), INT(a, \alpha), ACH(a, p)$$

É também necessário definir uma regra de restrição de integridade:

$$2. \neg(BEL(a, p) \wedge ACH(a, p)).$$

Na secção seguinte será demonstrado a aplicação deste modelo num domínio específico: bases de conhecimento médico.

3 Interação com uma base de conhecimentos médica

Suponhamos uma arquitectura em que um médico interage com uma base de conhecimentos médica de interações medicamentosas que possui também informações sobre os seus utentes. Uma possível interação poderia ser:

- M: Vou receitar Lasix à Maria.
- S: Não deve receitar Lasix à Maria porque ela é diabética!

Para modelar esta interação vamos supor que tanto o médico como o sistema têm conhecimento sobre os diagnósticos e sobre a acção dos medicamentos, mas que somente o sistema conhece as interações medicamentosas.

3. $BEL(a, \text{diagnostico}(maria, hipertensao))$
4. $BEL(a, \text{diagnostico}(maria, diabetes))$
5. $BEL(s, \text{interage}(lasix, diabetes))$
6. $BEL(a, \text{TO}(\text{receitar}(a, lasix, x), saude(x))) \Leftarrow$
 $BEL(a, \text{diagnostico}(x, hipertensao)),$
 $notBEL(a, \text{interacao}(x, lasix))$
7. $BEL(a, \text{TO}(\text{receitar}(a, m, x), doente(x))) \Leftarrow$
 $BEL(a, \text{interacao}(x, m))$
8. $BEL(a, \text{interacao}(x, m)) \Leftarrow$
 $BEL(a, \text{diagnostico}(x, d)),$
 $BEL(a, \text{interage}(m, d))$
9. $INT(a, \text{receitar}(a, x)) \Rightarrow$
 $BEL(a, doente(x))$

As quatro últimas regras indicam que se um agente receitar Lasix a um utente hipertenso, ele ficará de saúde desde que não haja interações medicamentosas negativas e que se um médico receitar um medicamento a alguém é porque ele acredita que esse alguém está doente.

É necessário definir, ainda, algumas restrições de integridade específicas ao domínio:

10. $\neg(\text{saude}(x) \wedge \text{doente}(x))$
11. $\neg(BEL(a, \text{saude}(x)) \wedge BEL(a, \text{doente}(x)))$
12. $BEL(m, \text{saude}(x)) \vee ACH(m, \text{saude}(x))$

Estas restrições indicam que um indivíduo ou está doente ou de saúde; que um agente não pode acreditar que alguém esteja doente e de saúde simultaneamente; e que um médico ou acredita que os seus utentes estão de saúde ou quer que eles venham a ficar bons.

Considerou-se, ainda, que o conjunto dos abduzíveis conteria os predicados que representam as atitudes: *BEL*, *ACH* e *INT*.

Após a primeira frase do médico é adicionado o facto:

13. $INT(m, \text{receitar}(lasix, maria))$

De acordo com o formalismo descrito na secção anterior, é calculado o modelo estável referente a este programa lógico que inclui os seguintes factos:

14. $BEL(m, doente(maria))$ — pela regra 9;
15. $ACH(m, saude(maria))$ — utilizando o facto anterior e as restrições de integridade 11 e 12 o sistema pode abduzir que o médico acredita que a Maria virá a ficar boa;
17. $BEL(m, \text{TO}(\text{receitar}(lasix, maria), saude(maria)))$ — pelas regras 6 e 8 e utilizando o facto 3. Neste ponto é importante focar que é utilizada a negação por falhanço na inferência desta proposição;
18. $INT(m, \text{TO}(\text{receitar}(lasix, maria), saude(maria)))$ — pela regra 1 e utilizando os factos 17, 13 e 15;

19. $BEL(s, TO(receitar(lasix, maria), doente(maria)))$
— pelas regras 7 e 8 e utilizando os factos 3, 4 e 5.

Este conjunto de factos permite que o sistema faça a inferência das intenções do médico mas que mantenha as suas próprias crenças. Nomeadamente, o sistema acredita que a acção do médico não é correcta e informa-o disso.

4 Conclusões e trabalho futuro

Conforme foi referido, um sistema de interacção em Língua Natural necessita de ter a capacidade de inferir sobre a informação veiculada pelos diversos agentes de modo a poder reconhecer e representar as atitudes dos diversos agentes da interacção. A metodologia proposta neste artigo fornece uma semântica formal de suporte a esta inferência que suporta raciocínio não-monotónico abductivo e que permite modelar as diversas atitudes dos diferentes agentes participantes numa interacção.

No entanto, um dos problemas em aberto é que este processo de raciocínio tem uma elevada complexidade computacional.

Por outro lado, o formalismo apresentado não tem em linha de conta a necessidade de suportar a revisão e a visibilidade das crenças dos agentes ao longo das interacções. Estes problemas terão de ser objecto de análises futuras.

Agradecimentos

Gostaríamos de agradecer ao Dr. Gustavo Quaresma pelos exemplos de aplicação à área médica e à Irene Rodrigues pelos seus comentários e sugestões em relação a este trabalho.

Referências

- [1] Douglas E. Appelt and Martha E. Pollack. Weighted abduction for plan ascription. *User Modeling and User-Adapted Interaction*, 2(1), 1992.
- [2] P. Cohen and H. Levesque. Intention is choice with commitment. *Artificial Intelligence*, 42(3), 1990.
- [3] K. Eshghi and R. Kowalski. Abduction compared with negation as failure. In *Proceedings of the 6th International Conference on Logic Programming*. ICOT, 1989.
- [4] R. E. Fikes and Nils J. Nilsson. Strips: A new approach to the application of theorem proving to problem solving. *Artificial Intelligence*, (2):189–208, 1971.
- [5] M. Gelfond and V. Lifshitz. The stable model semantics for logic programming. In *Proceedings of the 5th International Conference On Logic Programming*, 1988.
- [6] J. Hobbs, M. Stickel, P. Martin, and D. Edwards. Interpretation as abduction. In *Proceedings of the 26th Annual Meeting of ACL*, 1988.
- [7] A. Kakas and P. Mancarella. Generalised stable models: A semantics for abduction. In *Proceedings of the 9th ECAI*, 1990.
- [8] Diane J. Litman. *Plan Recognition and Discourse Analysis: An Integrated Approach for Understanding Dialogues*. PhD thesis, Dep. of Computer Science, University of Rochester, 1985.
- [9] Martha E. Pollack. *Inferring Domain Plans in Question-Answering*. PhD thesis, Dep. of Computer and Information Science, University of Pennsylvania, 1986.
- [10] C. Preist and K. Eshghi. Consistency-based and abductive diagnoses as generalised stable models. In *Proceedings*

*of the International Conference on 5th
Generation Computer Systems*, pages
514–521, 1992.

- [11] William Rodi and Stephen Pimentel. A nonmonotonic assumption-based tms using stable models. In *Principles of Knowledge Representation and Reasoning*, pages 485–495, 1991.
- [12] Earl D. Sacerdoti. *A Structure for Plans and Behavior*. American Elsevier, New York, 1977.