

# Uma Proposta Preliminar de Síntese Automática de Normas para Sistemas Multiagente

Jhonatan Alves, Jomi Fred Hübner, Jerusa Marchi

<sup>1</sup>Programa de Pós-Graduação em Engenharia de Automação e Sistemas  
Universidade Federal de Santa Catarina  
Caixa Postal 88040-900 – Florianópolis -- SC – Brasil

jhonatan.alves@posgrad.ufsc.br, {jomi.hubner, jerusa.marchi}@ufsc.br

**Resumo.** Normas têm sido usadas em sistemas multiagentes para regular a sociedade dos agentes na forma de regras que, em certos contextos, impõem restrições sobre os seus comportamentos. Sob estas circunstâncias, este trabalho apresenta a proposta preliminar de um framework para a síntese automática de normas. Como resultados esperados pretende-se que as normas obtidas sejam capazes de evitar conflitos no sistema mantendo os seus objetivos alcançáveis e as ações da sociedade coerentes.

**Abstract.** Norms have been used in multiagent systems to regulate the agent society in the form of rules that, in certain contexts, impose restrictions on their behavior. Under these circumstances, this paper presents a preliminary proposal of a framework for the automated synthesis of norms. As expected, it is hoped that the obtained norms will be able to avoid conflicts in the system keeping its objectives achievable and the actions of the society coherent.

## 1. Introdução

Em sistemas multiagente (SMA) a autonomia é a principal característica de um agente, pois lhe confere a capacidade de decidir como e quando interagir com os demais membros da sociedade para contemplar os seus objetivos [Wooldridge and Jennings 1995]. Entretanto, as decisões tomadas pelos agentes podem causar conflitos no sistema diminuindo a sua eficiência e estabilidade [Weiss 1999]. Sob a luz desses fatos, uma questão crucial no desenvolvimento dos SMA é “como garantir a eficiência e preservar propriedades desejáveis do SMA de modo que a autonomia dos agentes seja respeitada?” [Boella et al. 2006].

Para isso, é necessário que o comportamento dos agentes seja regulado através de um mecanismo de coordenação para que ajam de forma coerente no sistema [Huhns and Stephens 1999]. Desde o trabalho precursor de [Shoham and Tennenholtz 1992] as normas têm sido usadas em SMA para regular a sociedade na forma de regras que, em certos contextos, governam os comportamentos dos agentes. O estudo das normas neste campo é extenso e variado. Entre os tópicos de estudo tem-se a síntese que se refere ao processo da criação de um conjunto de normas para o sistema [Frantz and Pigozz 2018]. Neste sentido, este trabalho apresenta a proposta preliminar de um framework que monitora o comportamento dos agentes e, conforme conflitos são detectados, sintetiza normas para garantir a estabilidade e o alcance dos objetivos do sistema.

Este trabalho está organizado da seguinte forma: na seção 2 a síntese de normas em SMA e trabalhos correlatos são apresentados e discutidos; na seção 3 a proposta deste trabalho é apresentada; e, por fim, na seção 4 as conclusões e direções futuras são apresentadas.

## 2. Síntese de Normas e Trabalhos Correlatos

A síntese de normas pode ser classificada, basicamente, como *off-line* ou *on-line*. A síntese *off-line* ocorre antes da execução do sistema e tem como principal vantagem permitir que a sociedade dos agentes seja regulada desde o início das suas interações, uma vez que as normas são obtidas previamente à inicialização do sistema. Ainda, a síntese *off-line* se distingue entre as abordagens manuais e automáticas. A abordagem manual é realizada pelo projetista, ao passo que a abordagem automática é realizada por um mecanismo computacional.

Por outro lado, a síntese *on-line* é automática e ocorre em tempo de execução do sistema. Este tipo de síntese se distingue entre as abordagens por emergência e observação. Na abordagem por emergência o projetista do sistema fornece, em tempo de projeto, um conjunto de ações a partir do qual os agentes decidem através de diversas interações qual ação será adotada como obrigatória por toda a sociedade. Na abordagem por observação a síntese consiste em determinar em quais circunstâncias as normas devem ser criadas. Para isso, faz uso de um mecanismo computacional que, ao identificar conflitos durante a interação dos agentes, sintetiza normas para evitá-los em novas situações.

Dentro desse contexto, buscou-se na literatura trabalhos de síntese de normas em SMA, onde foram analisados 15 trabalhos, dentre os quais 6 são abordagens *off-line* e 9 *on-line*. Dentre os trabalhos *off-line* [Shoham and Tennenholtz 1992, Shoham and Tennenholtz 1995, van der Hoek et al. 2007, Alrawagfeh and Meneguzzi 2015] são manuais. A síntese manual é propensa a erros e complexa de ser realizada para sistemas com grandes espaços de estados. Os trabalhos [Christelis 2011, Morales et al. 2018] são *off-line* e automáticos, porém, ambos compartilham do problema de que certas normas têm a chance de nunca se tornarem ativas, pois em tempo de execução do sistema os agentes podem não apresentar os comportamentos para os quais aquelas normas foram sintetizadas, o que pode implicar em um esforço computacional desnecessário para sintetizar algumas normas.

Os trabalhos [Kittock 1995, Walker and Wooldridge 1995, Onn and Tennenholtz 1997, Epstein 2001, Boella and van der Torre 2007, Savarimuthu et al. 2008, Griffiths and Luck 2011, Mahmoud et al. 2015] realizam a síntese sob a perspectiva da emergência. Nesse tipo de abordagem os agentes devem possuir maquinário para sintetizarem a norma além de se disporem a participar do processo de emergência, o que não pode ser garantido. Além do mais, tais trabalhos se limitam a encontrar apenas uma norma. O trabalho de [Morales et al. 2013] visa obter as normas por observação, deste modo as normas são sintetizadas para o comportamento corrente dos agentes. Entretanto, para que as normas sejam sintetizadas é necessário que ocorram, primeiramente, conflitos no sistema. A cada novo conflito, uma norma é sintetizada. Ao longo da execução do sistema, várias normas são obtidas e inseridas em

uma rede hierárquica de generalização/especialização, onde as normas mais gerais se tornam visíveis aos agentes. Entretanto, para isso, várias normas específicas são geradas.

Visto que a síntese off-line pode gerar normas não úteis ao sistema e que a síntese por emergência exige o engajamento dos agentes no processo de criação, este trabalho tem como objetivo responder a seguinte pergunta: “*Como sintetizar automaticamente normas para SMA que estejam melhor direcionadas ao comportamento corrente dos agentes e que viabilizem o alcance dos objetivos do sistema?*”.

### 3. Proposta

Para responder à pergunta de pesquisa, este trabalho tem como objetivo desenvolver um framework que sintetiza normas em modo on-line para SMA. Como abordado anteriormente, a escolha pela síntese on-line se dá pelo fato de que através deste modo é possível obter um conjunto de normas que regulam o comportamento vigente dos agentes. Em suma, a proposta preliminar do framework funciona em 4 etapas: (i) construção da história corrente dos estados do sistema percorridos pelos agentes conforme são monitorados (onde a história é um subespaço do espaço de estados do sistema); (ii) detecção de conflitos no estado corrente do sistema; (iii) análise da alcançabilidade dos objetivos do sistema antes de uma nova norma ser criada; (v) síntese de uma norma para evitar o conflito detectado. A nova norma é, então, disponibilizada aos agentes. A visão geral do framework está ilustrada na figura 1.

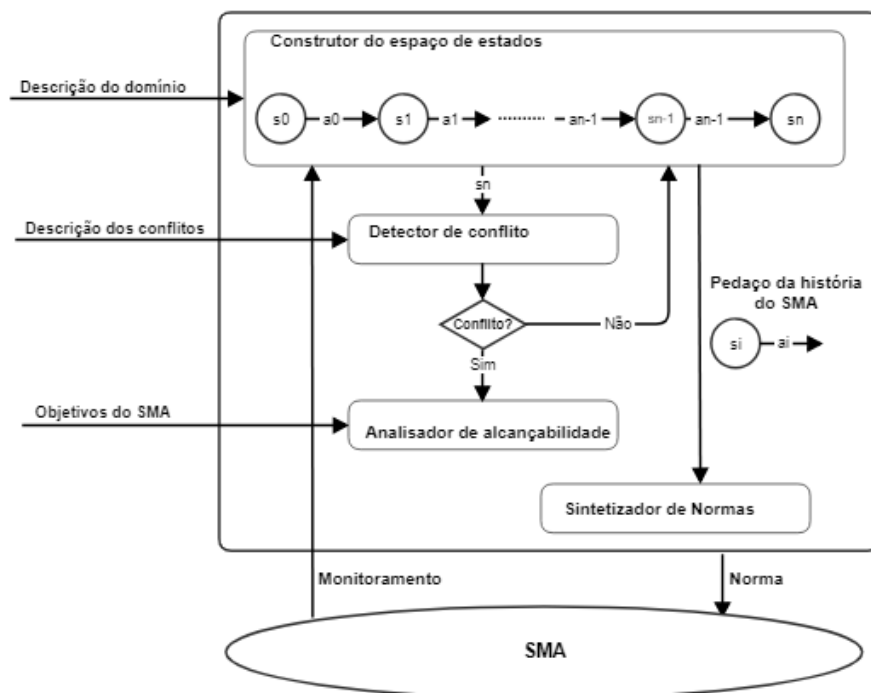


Figura 1. Framework proposto.

#### 3.1. Definição conceitual

A primeira etapa é realizada pelo módulo *construtor de história* que constrói a história corrente do SMA à medida que o monitora. Uma história é uma sequência finita de transições de estados que parte do estado inicial do sistema até o estado onde o sistema

se encontra em equilíbrio. Por estado de equilíbrio entende-se o estado a partir do qual uma estimativa indica que não haverá mais conflitos no SMA. Uma transição que parte de um estado  $s'$  para um estado  $s''$  denota a ação executada pelo agente em  $s'$  e  $s''$  denota o estado do SMA alcançado após a execução de tal ação. Na história, o estado corrente é representado por  $s_n$ .

A segunda etapa é iniciada pelo módulo *detector de conflitos* cuja função é averiguar a existência de algum conflito no estado corrente da história do SMA. Se houver, então há a necessidade de que uma norma seja sintetizada e disponibilizada aos agentes para que eles evitem-no futuramente. Caso contrário o fluxo de execução do framework retorna à primeira etapa.

Uma abordagem ingênua para evitar que o conflito detectado seja evitado pelos agentes em situações futuras é sintetizar uma norma que proíbe a execução da ação que fez com que o sistema transitasse do estado  $s_{n-1}$  para o estado  $s_n$  quando um certo contexto for satisfeito. Entretanto, há a possibilidade de que a norma criada nessas condições impeça o alcance de algum objetivo do sistema. Logo, para evitar a criação de uma norma obstrutora de objetivos o framework dá início à terceira etapa que é realizada pelo módulo *analizador de alcançabilidade*.

O analisador realiza backtracking sobre a história do sistema a fim de determinar o primeiro estado predecessor a  $s_n$  a partir do qual pode-se satisfazer os objetivos. Diferentes técnicas baseadas em planejamento ou grafos podem ser utilizadas pelo verificador em seu processo de backtracking. Uma vez que o estado predecessor foi encontrado dá-se início à quarta e última etapa do framework que é realizada pelo módulo *sintetizador de normas*.

O sintetizador recebe um recorte da história do sistema, mais especificamente o estado predecessor obtido na etapa anterior e a ação executada a partir daquele estado. Com o intuito de que a norma seja genérica (aplicabilidade em diferentes estados), o sintetizador obtém a norma como uma regra abstrata, onde os termos que compõem a norma (isto é, os termos do contexto e da ação) são predicados com variáveis livres. Desta forma, a norma não se aplica a um estado específico e através de um processo de unificação o agente pode verificar a aplicação da norma no estado no qual que se encontra. Em seguida, a norma é disponibilizada em uma base para que os agentes possam acessá-la a fim de raciocinarem sobre restrições comportamentais que elas possam vir aplicar.

### 3.2. Definição Formal

A seguir apresenta-se um conjunto de definições que formalizam os conceitos envolvidos no framework proposto neste trabalho.

**Definição 1** *Seja  $M = \{Ag, Ac, S, P, At, G, T\}$  um sistema multiagente, onde  $Ag = \{ag_1, ag_2, \dots, ag_n\}$  denota um conjunto de agentes,  $Ac = \{ac_1, ac_2, \dots, ac_m\}$  um conjunto de ações que os agentes podem executar,  $P$  um conjunto de predicados,  $At$  um conjunto de átomos proposicionais definidos sobre  $P$ ,  $S$  o conjunto de estados do sistema,  $G = \{g_1, g_2, \dots, g_k\}$  um conjunto de objetivos do sistema com  $g_i \subset At$  e  $T : S \times Ac \rightarrow S$  uma função de transição de estados que indica que a partir do estado  $s \in S$  pela aplicação de uma ação  $a \in Ac$  o sistema transita para um estado  $s' \in S$ .*

Ainda,  $S \subseteq 2^{At}$  e assume-se sobre a construção dos estados a hipótese do mundo fechado, isto é, átomos proposicionais que não estão contidos em um estado  $s \in S$

são considerados como falsos e o mundo evolui de forma determinística. A função de transição  $T$  indica que apenas uma ação pode ser realizada a cada estado.

**Definição 2** Seja  $C = \{c_1, c_2, \dots, c_n\}$  um conjunto de conflitos, onde  $c_i$  é uma expressão lógica cujos termos pertencem a  $P$  e faz uso dos conectivos da lógica de primeira ordem.

Um estado  $s \in S$  é um estado de conflito caso  $s \models c_i$  (a condição  $c_i$  é verdadeira em  $s$ ). Nesse caso, diz-se que  $s$  detrimenta o alcance dos objetivos do sistema  $M$ .

**Definição 3** Uma norma é um par  $\langle \varphi, \theta(ac) \rangle$ , onde  $\varphi \subset P$  denota o contexto no qual a norma se torna ativa,  $\theta$  é o operador deôntico proibido e  $ac \in Ac$  é a ação que é proibida de ser executada no contexto  $\varphi$ .

Sendo os conflitos e as normas definidos sobre o conjunto  $P$  faz-se com que ambos sejam genéricos. Através de um processo de unificação é possível verificar a aplicabilidade destes elementos aos estados do SMA.

#### 4. Conclusão

Este trabalho apresentou a proposta preliminar de um novo framework para a síntese automática de normas. Percebe-se que as seguintes limitações necessitam de uma solução prioritária para que a solução apresentada possa avançar: (i) como estipular um método para estimar o estado a partir do qual o SMA entra em equilíbrio; (ii) em caso de conflito, como determinar a ação responsável ao permitir que os agentes executem ações em paralelo; (iii) como avaliar se uma dada norma é eficiente em evitar o conflito para o qual ela foi sintetizada; (v) como determinar de modo eficiente a alcançabilidade dos objetivos do sistema.

#### Referências

- Alrawagfeh, W. and Meneguzzi, F. (2015). Utilizing permission norms in bdi practical normative reasoning. In *Revised Selected Papers of the International Workshops on Coordination, Organizations, Institutions, and Norms in Agent Systems X - Volume 9372*, pages 1–18, New York, NY, USA. Springer-Verlag New York, Inc.
- Boella, G. and van der Torre, L. (2007). Norm negotiation in multiagent systems. *International Journal of Cooperative Information Systems (IJCIS) Special Issue: Emergent Agent Societies*, 16(2).
- Boella, G., van der Torre, L., and Verhagen, H. (2006). Introduction to normative multiagent systems. *Computational & Mathematical Organization Theory*, 12(2):71–79.
- Christelis, G. D. (2011). *Automated Norm Synthesis in Planning Environments*. PhD thesis, University of Edinburgh.
- Epstein, J. M. (2001). Learning to be thoughtless: Social norms and individual computation. *Computational Economics*, 18(1):9–24.
- Frantz, C. K. and Pigozz, G. (2018). Modeling norm dynamics in multi-agent systems.
- Griffiths, N. and Luck, M. (2011). Norm diversity and emergence in tag-based cooperation. In De Vos, M., Fornara, N., Pitt, J. V., and Vouros, G., editors, *Coordination, Organizations, Institutions, and Norms in Agent Systems VI*, pages 230–249, Berlin, Heidelberg. Springer Berlin Heidelberg.

- Huhns, M. N. and Stephens, L. M. (1999). Multiagent systems. chapter Multiagent Systems and Societies of Agents, pages 79–120. MIT Press, Cambridge, MA, USA.
- Kittock, J. (1995). Emergent conventions and the structure of multi-agent systems. In *Lectures in Complex systems: the proceedings of the 1993 Complex systems summer school, Santa Fe Institute Studies in the Sciences of Complexity Lecture Volume VI, Santa Fe Institute*, pages 507–521. Addison-Wesley.
- Mahmoud, S., Griffiths, N., Keppens, J., Taweel, A., Bench-Capon, T. J. M., and Luck, M. (2015). Establishing norms with metanorms in distributed computational systems. *Artificial Intelligence and Law*, 23(4):367–407.
- Morales, J., Lopez-Sanchez, M., Rodriguez-Aguilar, J. A., Wooldridge, M., and Vasconcelos, W. (2013). Automated synthesis of normative systems. In *Proceedings of the 2013 International Conference on Autonomous Agents and Multi-agent Systems, AAMAS '13*, pages 483–490, Richland, SC. International Foundation for Autonomous Agents and Multiagent Systems.
- Morales, J., Wooldridge, M., Rodriguez-Aguilar, J. A., and Lopez-Sanchez, M. (2018). Off-line synthesis of evolutionarily stable normative systems. *Autonomous Agents and Multi-Agent Systems*, 32(5):635–671.
- Onn, S. and Tennenholtz, M. (1997). Determination of social laws for multi-agent mobilization. *Artificial Intelligence*, 95(1):155 – 167.
- Savarimuthu, B. T. R., Cranefield, S., Purvis, M., and Purvis, M. (2008). Role model based mechanism for norm emergence in artificial agent societies. In Sichman, J. S., Padget, J., Ossowski, S., and Noriega, P., editors, *Coordination, Organizations, Institutions, and Norms in Agent Systems III*, pages 203–217, Berlin, Heidelberg. Springer Berlin Heidelberg.
- Shoham, Y. and Tennenholtz, M. (1992). On the synthesis of useful social laws for artificial agent societies (preliminary report). In *AAAI*, pages 276–281. AAAI Press / The MIT Press.
- Shoham, Y. and Tennenholtz, M. (1995). On social laws for artificial agent societies: off-line design. *Artificial Intelligence*, 73(1):231 – 252. Computational Research on Interaction and Agency, Part 2.
- van der Hoek, W., Roberts, M., and Wooldridge, M. (2007). Social laws in alternating time: effectiveness, feasibility, and synthesis. *Synthese*, 156(1):1–19.
- Walker, A. and Wooldridge, M. (1995). Understanding the emergence of conventions in multi-agent systems. In *International Conference on Multiagent Systems, ICMAS'95*, pages 384–389. MIT Press.
- Weiss, G., editor (1999). *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence*. MIT Press, Cambridge, MA.
- Wooldridge, M. and Jennings, N. R. (1995). Intelligent agents: theory and practice. *The Knowledge Engineering Review*, 10(2):115–152.