Artificial Intelligence 229 (2015) 139–174

Contents lists available at ScienceDirect

Artificial Intelligence

www.elsevier.com/locate/artint

Coalition structure generation: A survey ☆

Talal Rahwan [a,*,1], Tomasz P. Michalak [b,c,1], Michael Wooldridge [b], Nicholas R. Jennings [d,e]

[a] Masdar Institute of Science and Technology, United Arab Emirates
[b] Department of Computer Science, University of Oxford, UK
[c] Institute of Informatics, University of Warsaw, Poland
[d] School of Electronics & Computer Science, University of Southampton, UK
[e] Department of Computing & Information Technology, King Abdulaziz University, Saudi Arabia

ARTICLE INFO

Article history:
Received 3 October 2014
Received in revised form 30 July 2015
Accepted 13 August 2015
Available online 21 August 2015

Keywords:
Coalitional games
Set partitioning
Coalition structure generation

ABSTRACT

The coalition structure generation problem is a natural abstraction of one of the most important challenges in multi-agent systems: How can a number of agents divide themselves into groups in order to improve their performance? More precisely, the coalition structure generation problem focuses on partitioning the set of agents into mutually disjoint coalitions so that the total reward from the resulting coalitions is maximized. This problem is computationally challenging, even under quite restrictive assumptions. This has prompted researchers to develop a range of algorithms and heuristic approaches for solving the problem efficiently. This article presents a survey of these approaches. In particular, it surveys the main dynamic-programming approaches and anytime algorithms developed for coalition structure generation, and considers techniques specifically developed for a range of compact representation schemes for coalitional games. It also considers settings where there are constraints on the coalitions that are allowed to form, as well as settings where the formation of one coalition could influence the performance of other co-existing coalitions.

1. Introduction

The multi-agent systems research field is concerned with understanding and building systems containing multiple autonomous software entities (called *agents*) that may have different preferences, goals, beliefs, and capabilities [85]. One of the key objectives of the multi-agent systems domain is to build agents that can take joint, coordinated actions, for example to improve their performance, or to achieve goals that are beyond the capabilities of individual agents. Such interaction may be useful both in cases where the agents are *cooperative* (i.e., their goal is to maximize some overarching system-wide objective) as well as cases where they are *selfish* (i.e., each agent acts in its own best interests, regardless of the consequences on other agents).

☆ Tomasz P. Michalak and Michael Wooldridge were supported by the European Research Council under Advanced Grant 291528 ("RACE"). Nicholas R. Jennings was supported by the ORCHID Project, funded by an EPSRC (Engineering and Physical Research Council) under the grant EP/I011587/1.

* Corresponding author.
E-mail address: trahwan@gmail.com (T. Rahwan).

[1] Talal Rahwan and Tomasz Michalak contributed equally to this article.

http://dx.doi.org/10.1016/j.artint.2015.08.004

The way the agents are organized in a system influences, or even governs, their interaction. For example, if the agents are organized hierarchically, then an agent would only be able to coordinate with its parent and/or children in the hierarchy. Clearly, there are many alternative organizational paradigms (not just hierarchies), each with their own strengths and weaknesses [32]. One such paradigm, which has received much attention in the literature, involves the formation of *coalitions*, i.e., groups of agents that typically exhibit the following characteristics: Firstly, they are goal-directed and short-lived; they are formed with a purpose in mind and dissolve when that purpose no longer exists, or when they cease to suit that purpose. Secondly, coordination occurs among members of the same coalition, but not among members of different coalitions. Thirdly, the organizational structure within each coalition is usually flat (rather than hierarchical, for example).

A wide range of potential applications of coalition formation have been considered in the literature. For example, by forming coalitions: autonomous sensors can improve their surveillance of certain areas [31]; virtual power plants can reduce the uncertainty of their expected energy output [13]; cognitive radio networks can increase their throughput [38]; and buyers can obtain lower prices through bulk purchasing [39]. It should be noted, however, that the coalition structure generation problem, being an abstraction, only serves as a first step towards understanding and building real-world solutions (as is the case with many research topics in cooperative game theory, where the focus on abstraction means that the solutions disregard much domain-specific information that may be critical when solving real-world problems).

Generally speaking, the coalition-formation process involves three main activities [78]:

- *Forming a coalition structure*. This activity involves each agent joining a coalition.[2] This is done either endogenously (by the agents deciding autonomously among themselves using some bargaining procedure), or exogenously (e.g., by a system designer). The resulting set of coalitions is called a *coalition structure*. Typically, we are interested in the coalition structure that maximizes social welfare, or minimizes the agents' incentive to deviate from their coalitions.
- *Solving the optimization problem of each coalition*. This activity addresses the following question: How should the members of a coalition coordinate their activities such that the performance of the coalition is maximized?
- *Dividing the reward of each coalition among its members*. If the benefits of cooperative action are accrued to a coalition as a whole, then the members of that coalition will need to agree on how to divide these benefits amongst themselves. Typically the goal is to do this in such a way as to satisfy certain desirable criteria, such as *fairness* (where each agent's reward reflects its contribution to the game), or *stability* (where no group of agents can selfishly benefit by forming their own coalition). In this context, a *solution concept* specifies (i) which coalitions to form, and (ii) how the payoff of every formed coalition is divided among its members.

In this article, we focus on the problem of identifying coalition structures that maximize social welfare. While the relevance of this problem is clear when the agents are cooperative, it is perhaps less so when the agents are selfish. The main relevance in the latter setting arises when we need to compute solution concepts that inherently require the agents to be partitioned optimally. This is the case, for instance, with the *core*—one of the key solution concepts in coalitional game theory [24,29]. According to this scheme, for any division of rewards to be stable, a necessary condition is that the agents form a social welfare-maximizing coalition structure (see, e.g., [16], Proposition 2.21). In other words, computing a stable outcome implies solving the coalition structure generation problem. Other fundamental game theoretic contexts where the solution to this problem may be useful are the *Price of Anarchy* and the *Price of Stability*. Specifically, in the coalitional game context, the Price of Anarchy is defined as the ratio between the *worst* stable[3] coalition structure and the welfare-maximizing coalition structure [35]. Similarly, the Price of Stability is defined as the ratio between the *best* stable coalition structure and the welfare-maximizing coalition structure [3].

Arguably, the first attempts to study the algorithmic aspects of coalition formation in the multi-agent community were those made by Shehory and Kraus [81], by Ketchpel [37], and by Zlotkin and Rosenschein [96]. Prior to these works, the primary focus in the literature was on the theoretical analysis of the properties of various solution concepts, rather than focusing on the development of coalition formation *algorithms*. Several heuristic algorithms were later on proposed by Shehory and Kraus for various settings [82–84]. Continuing this line of research, the seminal work by Sandholm et al. [78] studied the computational aspects of identifying coalition structures with worst-case guarantees on solution quality. Since then, numerous algorithms have been proposed to solve this problem, using a range of different techniques. In what follows, we present a comprehensive survey of this literature, and discuss a variety of directions from which the coalition structure generation problem has been approached. Careful attention has been given to ensure that the intuitions behind the different algorithms and their underlying theorems are presented in an accessible and clear manner. As such, the reader is not assumed to have expertise in combinatorial optimization or game theory.

The remainder of the paper is structured as follows.

- **Section 2** (*page 141*) introduces our key definitions and notational conventions.

[2] Games with overlapping coalitions, where an agent can join multiple coalitions simultaneously, have also been considered. See, e.g., the work by Shehory and Kraus [83] who first studied these settings among cooperative agents, and the work by Chalkiadakis et al. [15] who considered selfish agents.

[3] The word "stable" is used here in its broader sense; it is not restricted to the *core*, but could refer to the Nash equilibrium, for example.