To appear in: 5th European Workshop on Multi-Agent Systems (EUMAS 2007). Hammamet, Tunisia, December 13–14, 2007

A Clustering Protocol for Team Formation Based on Concept, Location, and Time

Arne Schuldt and Sven Werner

Centre for Computing Technologies (TZI)
University of Bremen, Am Fallturm 1, D-28359 Bremen

Abstract. This paper presents a clustering protocol for team formation of intelligent agents. Since agents should be autonomous in their decision whether and with whom to cooperate, all decisions are made locally. The approach is not limited to quantitative distance measures, but is also suited for qualitative representations of concept, location, and time.

1 Introduction

Team formation is a challenging task for intelligent agents. Being incapable or unwilling to solve a certain task on their own, agents may form teams in order to jointly reach their goal [8]. This is especially the case if a high number of potential cooperation partners exists. In such cases it is necessary to establish distinct groups of agents based on their properties. For instance, consider shipping containers at a container terminal awaiting to be forwarded to a warehouse. One option for agents representing containers is to request a truck for transport. However, in terms of costs it is often much more desirable to be transported together with other containers, e. g., by train. Sharing a common location and destination at the same time are criteria for successful cooperation in this scenario.

Clustering algorithms separate a given set of objects into distinct groups. The objective is to achieve a maximal distance between different clusters and a minimal distance between the members of each cluster. Previous algorithms, such as k-means [5], take a centralised perspective on the objects to be clustered. Hence, no autonomy regarding clustering is left to the affected entities. By contrast, intelligent agents should be autonomous in their decision whether or not to join forces with other agents. Furthermore, they should be able to deliberately choose with whom to cooperate, which demands local decision-making.

Such distributed clustering is for example applied in wireless sensor networks. A common approach [3] is to cluster spatially proximal sensors in order to reduce communication costs. Cluster members provide their local cluster-head with the data acquired. The cluster-heads then transmit the aggregated data to the base station. However, quantitative distance measures (like spatial proximity) are not always adequate for the decision whether agents should cooperate. By contrast, Sect. 2 discusses why knowledge about qualitative relations (especially regarding concept, location, and time) is often much more suitable. Subsequently, Sect. 3 presents a clustering protocol that is capable of handling such representations and that additionally leaves all decision-making to the participating agents.

2 Conceptual, Spatial, and Temporal Clustering Criteria

Quantitative measures are not always sufficient and adequate in order to represent the location of agents. Imagine that the spatial distance between two shipping containers in a port is 500 metres. Although this is rather short, it does, however, reveal almost nothing about the question whether cooperation is eligible for these containers: They might be located in neighbouring container terminals of different operators which prevents them from being loaded on a common train. By contrast, a meaningful qualitative abstraction that tessellates the port into distinct areas allows to apply, e. g., topological relations [6] for reasoning about team formation. The temporal logic of [1] enables corresponding reasoning tasks to be performed on temporal intervals in order to coordinate agents by time constraints.

In addition to location and time for some tasks it is necessary to regard conceptual properties in order to identify potential partners. For instance, consider shipping container which are supposed to be received in a common warehouse if they carry similar goods. Applying description logics allows to characterise their content by ontological concepts. Based on this foundation it is then possible to determine their similarity by applying the five degrees of match (e. g., subsumption) proposed by [4]. A more extensive discussion of conceptual, spatial, and temporal clustering criteria is given in [7].

3 Clustering Protocol

The distributed clustering protocol (Fig. 1) presented in this section is capable of handling both quantitative data and qualitative relations. Furthermore, it leaves all decision-making regarding clustering to the affected agents. If an agent chooses to be clustered, it queries a catalogue service for existing clusters. Subsequently, it communicates its properties to all cluster-heads. Clusters matching the properties of the agent send positive answers. Having received a positive answer, the requesting agent joins the respective cluster. Otherwise, it registers as a cluster-head itself.

So far, the protocol suffers from a potential problem: concurrency. Querying the catalogue and registering oneself as a cluster-head is not an atomic operation. Hence, other agents with the same properties can register in between. In order to address this issue, the agent has to send its properties to all cluster-heads that have been registered in between as soon as its registration is finished. If the agent finds another cluster-head exhibiting the same properties and an earlier registration time-stamp, it deregisters and joins the earliest cluster found.

The clustering protocol is implemented within the JADE [2] agent framework. In order to show that it overcomes limitations of recent approaches regarding distribution and qualitative relations, a case study with data from about 2,400 shipping containers is conducted. They have to be clustered in accordance to their content (described by concepts of an ontology). A manual inspection reveals that there exist 215 clusters; so this is the expected outcome of the experiment.

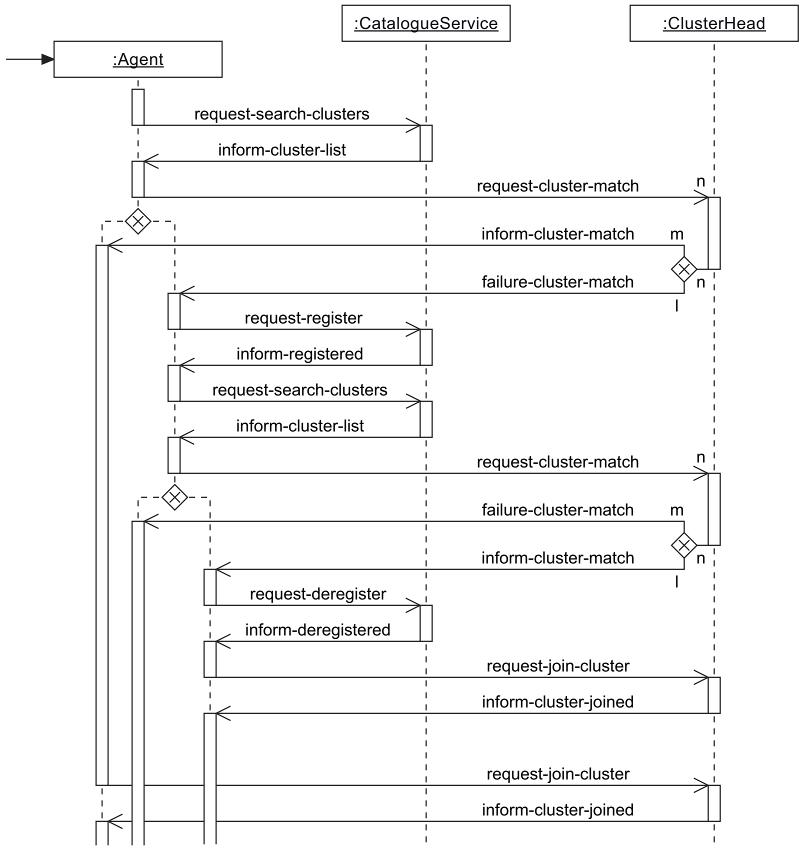

Fig. 1. Distributed clustering approach for autonomous agents described by an AUML sequence diagram (without exceptional messages for the sake of readability)

The average results of 50 test runs are given in Fig. 2. The number of clusters in relation to the total number of agents is depicted on the left hand side. Since the final number of clusters is 215 for all test runs, the algorithm is capable of solving the addressed problem.

The asymptotic communication complexity of the protocol is $O(mn)$ since all n agents contact at most all m clusters, whereby $m \leq n$. For most applications, m is considered to be even much smaller than n, which means $m \ll n$. In the case study, the total number of conversations (Fig. 2 centre) is 341,285. This is below the expected asymptotic complexity. This result can be explained by the fact that not all clusters exist right from the start. On a computer with Windows XP and an Intel Centrino Duo processor with 2.16 Gigahertz, the total time for clustering is 284 seconds (Fig. 2 right). The main benefit of this protocol, however, is the degree of autonomy left to the individual agent, since the catalogue service does not make any decisions for the agents.

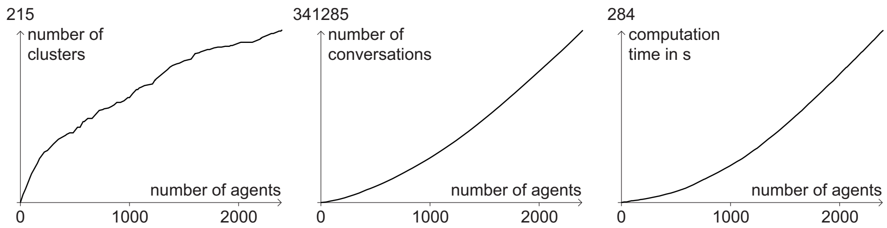

Fig. 2. The number of clusters (left) and conversations (centre) and the time needed for computing the simulation (right) in relation to the number of agents

4 Conclusion

The clustering approach introduced allows agents to form teams by semantically defined properties instead of being limited to quantitative data. Being highly distributed, the clustering protocol leaves all decision-making to the participating agents instead of a central entity. Its general applicability is demonstrated in a logistic scenario. However, the application is not limited to logistics as no special assumptions on the represented objects are made.

References

1. J. F. Allen. Maintaining Knowledge about Temporal Intervals. *Communications of the ACM*, 26(11):832–843, 1983.
2. F. Bellifemine, G. Caire, and D. Greenwood. *Developing Multi-Agent Systems with JADE*. John Wiley & Sons, Chichester, UK, 2007.
3. W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan. Energy-Efficient Communication Protocol for Wireless Microsensor Networks. In *33rd Hawaii International Conference on System Sciences*, volume 8, pages 8020–8029, 2000.
4. L. Li and I. Horrocks. A Software Framework for Matchmaking Based on Semantic Web Technology. *International Journal of Electronic Commerce*, 8(4):39–60, 2004.
5. J. MacQueen. Some Methods for Classification and Analysis of Multivariate Observations. In *Fifth Berkeley Symposium on Mathematical Statistics and Probability*, pages 281–296, 1967.
6. D. A. Randell, Z. Cui, and A. G. Cohn. A Spatial Logic Based on Regions and Connection. In *Third International Conference on Knowledge Representation and Reasoning*, pages 165–176, San Mateo, 1992. Morgan Kaufman.
7. A. Schuldt and S. Werner. Distributed Clustering of Autonomous Shipping Containers by Concept, Location, and Time. In *MATES 2007*, pages 121–132, Leipzig, Germany, 2007. Springer-Verlag.
8. M. Wooldridge and N. R. Jennings. The Cooperative Problem Solving Process. *Journal of Logic & Computation*, 9(4):563–592, 1999.