Coalition Formation Among Rational Information Agents

Matthias Klusch[1] and Onn Shehory[2]

[1] Institut für Informatik,
Christian-Albrechts-University Kiel,
24118 Kiel, Germany
e-mail: *mkl@informatik.uni-kiel.d400.de*

[2] Department of Mathematic & Computer Science,
Bar Ilan University,
Ramat Gan, 52900 Israel
e-mail: *shechory@bimacs.cs.biu.ac.il*

Abstract. Information agents can be an important tool for information-gathering and query-answering for the expanding WWW service as well as the large number of existing autonomous databases. Such agents behave like active intelligent front-ends of the stand-alone information systems. Information agents may work as individuals trying to satisfy thier own query-answering, i.e., a set of given information search tasks. However, they must cooperate efficiently with one another in order to gather information in non-local domains. In this paper we present an approach for cooperation and coalition formation among information agents for heterogeneous databases. In order to deal with the required association autonomy of these agents, we have developed a special decentralized coalition formation mechanism. It allows individually rational cooperation among the agents for information search. The semi-automatic creation of a local terminological information model enables each agent to hide local schema data as well as allowing for automated search-task processing. Knowledge-based productions from information search-tasks are sets of terminological interdatabase dependencies. These productions are obtained by classifying the search term of the task according to the local information model. An underlying cooperation convention in this federative agent system is that only members of a fixed coalition are mutually committed to provide the availability of all of the local information that they have used for building this particular coalition.

1 Introduction

Since the last decade there is a growing interest in coupling a set of already existing, heterogeneous and autonomous database systems. Active interoperability of databases in the sense of task-oriented cooperation between respectively attached intelligent front-ends has led to the paradigm of cooperative information systems introduced in [25]. Dependencies between data in different databases are called interdatabase dependencies (IDD). Any declarative specification of IDDs like in [32], [11] or [10] presumes somehow gained knowledge about where to find which kind of semantically related data. The recognition of such relations is needed e.g. for being able to specify queries which span several heterogeneous, autonomous databases or just discover the available distributed non-local information space as it is the case in a multidatabase environment. Most works towards the interoperability of databases, like the well-known approach of federated databases[31], or recent efforts in cooperative information gathering, presumes that this recognition problem is already solved or ignore the required strict autonomy of each database. The first approach for recognizing such interdatabase dependencies in a decentralized and autonomous environment by rationally cooperating information agents was introduced in [17],[18] and [19]. It utilizes methods from both research areas, distributed artificial intelligence (DAI) and terminological knowledge representation and reasoning. No centralized global or partial

schema integration takes place. A federative agent system FCSI[3] is constituted by a set of autonomous information agents which behave like intelligent front-ends for the database each of them is uniquely associated with. In order to be able to search for relevant non-local schema-data while protecting the own local schema structure each of them builds a local terminological information model `LIM` on top of its conceptual database schema. Mutual terminological classification of parts of their information models enables them for an automated search as well as local processing of requested search terms. There is no need for a central mediator or information broker agent. The cooperative information search and selective data sharing is driven by an utilitarian coalition formation process.
The remainder of this paper is organized as follows. Section 2 contains an introductory overview of the related research areas of federated databases and terminological knowledge representation. Section 3 briefly describes the FCSI approach for cooperative information search. A short overview of utilitarian coalition formation in DAI is in Section 4, while the special coalition types and a new decentralized formation algorithm for the FCSI is described in section 5.

2 A Brief Introduction to Some Related Research Areas

2.1 Federated Database Systems

The following short introduction bases on the work of [31].

A *federated database system* (FDBS) is a collection of cooperating but autonomous component databases. In contrast to a distributed database system a FDBS is created in a bottom-up fashion, i.e., by a partial integration of already pre-existing databases while respecting particularly their association autonomy and solving the semantic heterogeneity between them. Semantic heterogeneity occurs when there is a disagreement about the meaning or intended use of the same or related data, like homonyms, synonyms, value or domain conflicts as presented e.g. in [10]. Association autonomy implies that each component database decides itself which kind of its private schema-data in the local conceptual schema it is willing to share. There is no centralized control for data sharing in a federated architecture.
The development of a FDBS includes the equivalent translation of the given local schemas into homogeneous *component schemas* using a canonical data model, the definition of the available parts of these component schemas as *export schemas* and finally their integration into a *federated schema* of a federation leading to a five-level schema architecture. An *external schema* defines a schema or view customized for each user. The *federal data dictionary or directory* is for auxiliary purposes. It includes e.g. information about the translations, mappings and the addresses of the federation participants. Federated databases can be further classified with respect to who creates and manages the federation and how the component databases are integrated. A *tightly coupled* FDBS provides location, replication, and distribution transparency. This is accomplished by developing a federated schema that integrates multiple export schemas. The transparencies are provided by the mappings between the federated and the export schemas such that the user can pose a query against one single or multiple federation schemas without knowing where the requested data is located. A central federation administrator defines and manages the federated schema and the external schemas related to it. There may exist a single or multiple federated schemas within one tightly-coupled FDBS.
In contrast, in *loosely coupled* systems no global or partial static schema integration takes place and no transparency is provided. Instead of a central federation administrator, the user himself must find the appropriate export schemas that can provide the required data,

[3] FCSI is a shortcut for '*F*ederative *C*ell *S*ystem for discovery of *I*nterdatabase dependencies'.

and then to define the respective mapping operations. Systems which base on dynamic integration of export schemas by the use of an extended query language (e.g. `MSQL`[22]), are also called Multidatabase-Language Systems [8].

Following [11], interdependent data are data which are related by an integrity constraint specifying their mutual dependency. Kinds of semantic heterogeneity determine respective interdatabase dependencies. Examples of interdependent data include replicated data, partially replicated data or summary data. Approaches for a declarative or functional specification of such dependencies can be found e.g. in [32], [11] and [10]. The approach of [32] denotes directed data dependencies by a logical predicate using relational algebra while in [10] a SQL-based language is used for this purpose. We omit the description of other approaches here.

To enable the specification of interdatabase dependencies, the locations of various related data must be known. That is, it is necessary to know how to acquire this relevance knowledge and how to automate the underlying decentralized search for such intensional relationships. Some criticism on federated database systems as in [25] state in particular their lack of support with respect to the recognition and maintenance of interdatabase dependencies. In loosely coupled federated databases no global integrity constraints can be expressed across sites, because no mechanism exists to detect the underlying data relationships and then to enforce them. In tightly coupled FDBS, where the local export schemas are integrated into a fixed federated schema by the central administrator, some descriptive information about the local semantics of the available schema-data must be provided in advance. The IDD recognition process shall be performed without any efforts at static schema integration, or necessity for the user to browse through known export schemas, as required in tightly coupled and loosely coupled FDBS, respectively. One main drawback of browsing through export schemas is that the local interpretation of importable, non-local schema structures is heavily burden to the potential importer and relies on the unique mutual understanding of the semantic descriptions attached to these schema-data by the exporter. Besides, all parts of one export schema are visible in the same degree for all participants in the federation as well as for others.

The FCSI provides an approach for cooperatively discovering intensionally related data in a decentralized and autonomous environment by utilizing techniques from terminological knowledge representation and reasoning.

2.2 Terminological Knowledge Representation

Terminological knowledge representation and reasoning deal with investigations of *description logics*. Most of them are decidable fragments of first-order predicate logic and much more expressive than propositional logic. They are based on the work of Brachman and Schmolze, the KL-ONE System[7]. Such terminological or concept languages provide a structured formalism to axiomatically describe the relevant *concepts* of an application domain and the interactions between these concepts using *roles*. Concepts and roles can be seen as unary and binary predicates, respectively. Complex concepts will be inductively build from primitive components or atomic concepts and roles by given *term-forming operators*. A *terminology* (Tbox) is a set of such complete or partial definitions, i.e. *terminological axioms*, of named concepts and roles. Unlike in other conceptual or semantic data models it is possible to describe concepts and roles using intensional descriptions phrased only in terms of necessary but not sufficient properties that must be satisfied by their instances. Common term-forming operators are, among others, conjunction and complement for concepts, and number and value restrictions for roles. In addition, concrete domain objects can be explicitly asserted as concept instances and related to other objects via roles by a set of *assertional axioms* (Abox). This is similar to the classical distinction between the schema and state level in the database area. The terminological and the assertional formalism together constitute a *hybrid knowledge representation language*. *Classification* of some concept into a given terminology bases on the notion of *subsumption*. One concept

subsumes another if, in all possible worlds, its set of instances is contained in that of the other. It is possible to determine subsumption between concepts as well as roles only by considering their given terminological definitions and a structural comparison of respective terms. This leads to a *subsumption hierarchy* of named terms, the concept and the role taxonomy.
Although for almost all expressive description logics term-subsumption is decidable (it is decidable for $\mathcal{ALC}$[34], but not for KL-ONE [27]), its computation is inherently intractable [24]. Thus, for pragmatic reasons most terminological systems use an incomplete but polynomial subsumption algorithm (e.g. CLASSIC[6] or KRIS[1]). For a more details on hybrid terminological knowledge representation and reasoning we refer the reader to [23].

3 FCSI Information Agents

The FCSI approach for recognizing interdatabase dependencies in a decentralized and autonomous environment works without any effort in centralized global or partial schema integration. There is also no need to browse through other export schemas like in loosely-coupled federated databases in order to find some available and possibly intensional relevant schema-data. An agent will be provided only with the necessary descriptions for semantical interpretation of imported schema parts for local processing, but it will not have access to the local schema-data which is linked to this description. The federative agent system FCSI is constituted by a set of autonomous information agents which are each the front-end of the database with which it is uniquely associated (cf. Fig. 1).

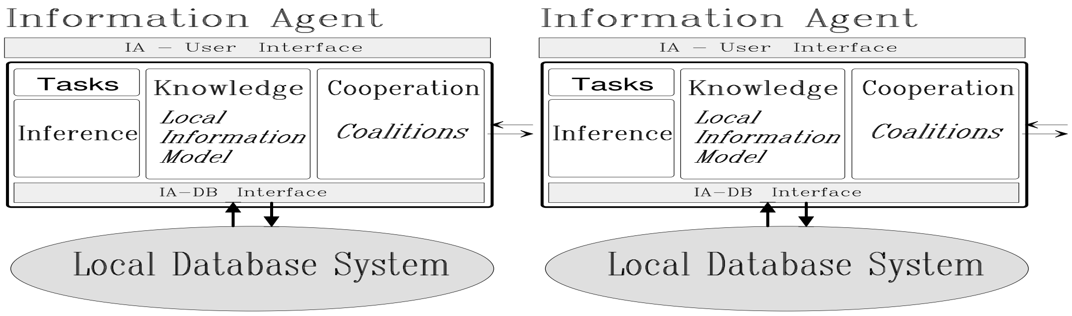

Fig. 1. FCSI Information Agents

To enable automatic search for relevant non-local schema-data, yet protect the own local schema structure, each agent builds a local terminological information model `LIM` on top of its conceptual database schema. A mutual terminological classification of parts of their information-models enables the agents for an automated search as well as local processing of requested search terms. No central mediator or information broker agent is necessary. There are several possibilities for an agent to construct its `LIM`, all rely on terminological knowledge representation methods.
The search for relevant data is achieved by a rational cooperation among the FCSI information agents which is performed via a decentralized, utilitarian coalition formation (cf. Sects. 4, 5). Based on the discovery of some interdatabase dependencies, each agent can then pose directed, intensional data requests to the respective members of the same coalition.

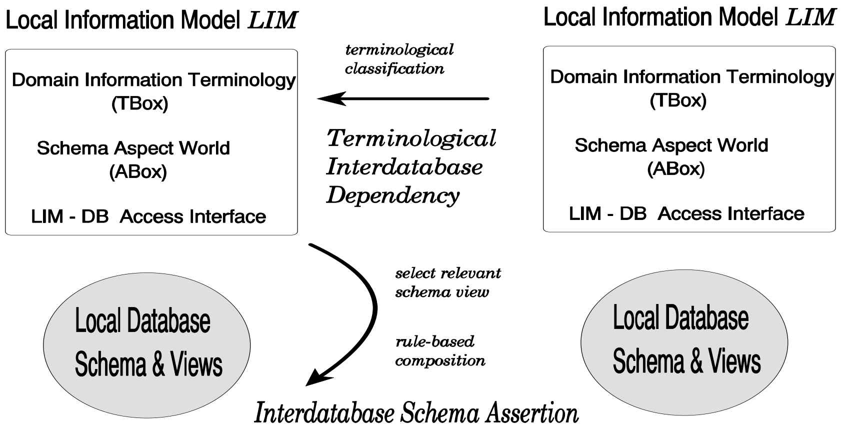

Fig. 2. Recognition of Interdatabase Dependencies

Essentially, a local information model `LIM` entails a more linguistic based, i.e., a terminological description of a given set of views or aspects of database schema-objects. It has the following three components: a local `Domain Information Terminology DIT` (TBox), the `Schema Aspect World W` (ABox) as its instantiation and the internal `LIM-DB Interface` as an DB Access Interface. For construction purposes an *information terminological formalism* ITF as well as a *schema aspect assertional formalism* SAF as its conservative extension[23] is available. ITF provides some most usual term-forming operators such as conjunction of concepts, number, value and existential restriction for roles as well as atomic concept negation and attribute sets. Both formalisms constitute together the hybrid terminological description language $\mathcal{ITL}$. The respective algorithms for hybrid terminological reasoning in this language are used by each FCSI agent to create its own `LIM` automatically. In addition, the agents must be provided with an individual read-access authorization for each available schema part.

The `LIM` construction-methods ([16, 18]) are mostly based on the translation of the given database schema and/or a set of schema-views, i.e. named queries on schema-objects, into a (schema or view) terminology, maybe together with a respective ABox. In principle, the choice of the $\mathcal{ITL}$ does not influence the functionality of the FCSI agents. However, the set of their term-forming operators restricts the set of schema-views which are compilable to database state-equivalent terminological view concepts [4]. Creation of the local information model and linking of available schema data into this terminological knowledge base is done semi-automatically by the information agent. Each local schema-view has a twofold description: first at intensional (i.e., terminological) level, it is described by a compiled *aspect term* and second, at database schema and state level, by the respective view qualification constraint.

Thus, each FCSI information agent is able to detect two different kinds of interdatabase dependencies: first, a *terminological interdatabase dependency (i-IDD)* by a mutual terminological classification of terminological view descriptions into the `LIM` of another information agent and then, projecting down to the local schema; second, an *interdatabase schema assertion (IDSA)* by rule-based composition of both of the underlying view-qualification constraints, formulated in the local DML into a global integrity constraint. This is comparable e.g. to the boolean-valued data dependency predicate in [32].

For example, the terminological interdatabase dependency $p - intsub(o_1, o_2, M, N)$ means

[4] Obviously, there exist several alternatives for hybrid terminological reasoning within each agent like the more expressive $\mathcal{ALCNR}$ [34].

that schema-object o_1 is partially, intensionally subsumed by the schema-object o_2. This bases on the detected term-subsumptions concerning some of the aspect terms of o_1 as given in the set M by those aspect terms of o_2 in the set N. The complete definitions and processes can be found in [16, 18, 20]. A simple example is provided below.

Example 1:

Let FCSI agent a_2 have the following representations of its schema-view 'female person' (vid_2) of the schema-object 'Person' (wrt. its own `LIM`): ($\underline{\texttt{and}}$ $\texttt{Mensch}_\texttt{p}$ $\texttt{Frau}_\texttt{p}$ ($\underline{\texttt{not}}$ $\texttt{Mann}_\texttt{p}$)) and the linked view qualification constraint ($p.WM =' W'$), and let FCSI agent a_1 have the following aspect term for its schema-view 'female student' (vid_1) of the schema-object 'Student' (wrt. its own `LIM`):
($\underline{\texttt{and}}$ $\texttt{Human}_\texttt{p}$ $\texttt{Woman}_\texttt{p}$ $\texttt{Student}_\texttt{p}$ ($\underline{\texttt{not}}$ $\texttt{Man}_\texttt{p}$) ($\underline{\texttt{all}}$ $\texttt{has}-\texttt{child}_\texttt{p}$ $\texttt{Child}_\texttt{p}$)
($\underline{\texttt{atleast}}$ $\texttt{has}-\texttt{child}_\texttt{p}$ 1)($\underline{\texttt{atmost}}$ $\texttt{has}-\texttt{child}_\texttt{p}$ 3)).
Then 'female persons' of agent a_2 terminologically subsumes 'female Student' of agent a_1.

- **cooperative recognition of terminological interdatabase dependencies:**

FCSI Agent a_1:

FIND-Task:
- aspect term for 'female student'
- $p - intsub(Student, ?, \{vid_1\}, \{?\})$

FCSI Agent a_2:

FIND-Task:
- aspect term for 'female person'
- $p - intsub(?, Person, \{?\}, \{vid_2\})$

`exchange and process tasks:` produce task satisfying terminological interdatabase dependencies;
`try to rationally coalesce with other information agent:`
determine utilities and coalition offers (cf. Sect. 5)

Information Availability within a Coalition (IAC Convention):
`All members of a coalition are mutually committed to provide the availability of their local schema information they used for building exactly this coalition.`

- **Intensional Data Request:**

FCSI Agent a_1:

'Names of Persons related to female Students'
request-for $Name$ **from** $Person$ **wrt** vid_2

FCSI Agent a_2:

check ($Person, vid_2$):
get view vid_2 qualification constraint ;
compile into local EER-DML query:
$\underline{\texttt{retrieve}}$ $p.Name$ $\underline{\texttt{from}}$ $Person$
$\underline{\texttt{where}}$ $p.WM =' W'$

A search task (FIND-task) is satisfied iff the result of the included PROLOG-query is not empty, means any terminological interdatabase dependency of the requested type satisfies this task. The required coalition information availability (IAC) has several impacts. One is that an agent has to send an aspect term included in a FIND-task to other agents only if it is a description of an own local schema-view, which will be available for them in cases of cooperation. Other impacts of the IAC assumption on coalition formation will be described in Sect. 5.

4 Coalition Formation in DAI

Distributed artificial intelligence (DAI) is concerned with problem solving in which several agents interact in order to achieve goals. During the past few years, several solutions to the coalition formation problem have been presented by researchers in the field of DAI [28, 30, 26, 15, 38]. These solutions are appropriate for a variety of environments (TOD's, superadditive and general environments). Some are appropriate for individually rational agents while others are designed for cases of bounded rationality. There are models that were designed for Distributed Problem Solvers (DPS) systems while others were developed for Multi-Agent systems (MAS). While some of these solutions are based on concepts from game theory, others are based on operations research, graph theory and algorithmic aspects of combinatorics.

Cooperation among autonomous agents may be mutually beneficial even if the agents are selfish and try to maximize their own expected payoffs [21, 35, 37]. Mutual benefit may arise from resource sharing and task redistribution. Coalition formation is an important method for cooperation in multi-agent environments. Agent membership in a coalition may increase the agent's ability to satisfy goals and maximize either the personal or the system's outcome. This may lead designers of computational agent-systems to adopt such methods. Facing the variety of coalition formation methods, designers of distributed computational agent-systems may be puzzled when trying to incorporate such a mechanism into their system. Game theory, which is most commonly employed by DAI researchers, usually answers the question of what coalition will form, and what reasons and processes will lead the agents to form a particular coalition among all possible coalitions. The designers of agent-systems are mostly interested in the question of which procedure should their agents use to coordinate their actions, cooperate and, if necessary, form a coalition. They must take into consideration the constraints of a multi-agent environment, such as communication costs and limited computation time. This was done by the DAI researchers that have provided coalition formation models. However, when a specific agent-system is been dealt with, the coalition formation methods shall be thoroughly examined before adopted, and well adapted to fit the special properties of the specific system. Otherwise, even the best coalition formation algorithm may lead to a poor performance of the system.

In this paper we adjust a DAI coalition formation method to a set of cooperating, autonomous database systems in which each database is represented by an autonomous information agent. As described in the previous section, these agents are autonomous individually rational agents. That is, they try to increase their own personal benefits, and not necessarily the outcome of the system as a whole. The tasks of the agents are query-answering tasks (cf. Sect.3, Example 1). We assume that the agents (or their owners) receive some kind of payment for the answers to these queries, and that there is an agreed-upon method to assess the monetary or abstract utility value of each query before it was answered. Upon these properties, we seek a coalition formation method for rational agents in a MAS, that will not be limited to superadditive environments. However, we must adjust the method we adopt to the information agents case. In particular, the privacy of information as well as the coalition information availability convention (IAC) mentioned in Sect.3 has to be taken into consideration. Following the IAC assumption, members of a newly-formed coalition have to reveal some of their private information to others in the same coalition. Such a requirement constrains the coalition formation, because it may affect the advisability of the coalition for its members. The details of the adjusted coalition formation procedure are in the next section.

5 Coalitions of FCSI Agents

The main aim of utilitarian coalition building is to increase the ability of the FCSI agents to satisfy their search tasks via cooperation, thus increasing their benefits. The coalition formation in this case requires to find partitions of the agents with respect to their utilities. These utilities are calculated and result from the execution of either the agents' own search tasks or tasks that they receive from other FCSI agents. If the agents are rational, then such partitions, or coalitions, will form iff each member of a coalition will gain more if it joins the coalition than it could gain by itself previously. However, in the special case of FCSI information agents this requirement is not sufficient. The need to respect the association

autonomy of the databases, each represented by a FCSI agent, implies that the coalition formation procedure should not allow coalitions in which there are information-agents that do not explicitly cooperate. This is because the formation of a coalition causes that the members of the potential coalition reveal some of their information to one another (in particular the aspect term and desired type of terminological interdatabase dependency as included in a search task). Thus, they already lose some of their privacy during the coalition formation. In cases of no cooperation, the agents prevent the access to the respective schema information that they would have allowed the committed coalition members. Therefore, an agent that decides to avoid the membership in a potential coalition during the formation process, will have access only to the information that will tell it what do the other agents search for, in terms of the terminological descriptions that it received from them. It will not get their respectively associated local schema-data. Hence, only coalitions in which all of the members have to cooperate are allowed in the FCSI environment.
The IAC convention that all of the FCSI coalitions must respect differentiates them from both the common game theoretic coalitions types and the existing DAI approaches to coalitions. For coalition formation within the FCSI we adapted the production-oriented approach of Shehory and Kraus[28]. The agents' utility functions are calculated with respect to the agents' monetary utility from their own knowledge-based productions, obtained by execution of received or own FIND-tasks (cf. Table 1). The utility that an agent obtains when it satisfies its own tasks exclusively by itself is denoted as the agent's self-value. The coalition value is the sum of these utilities of all its potential coalition members. Considering the marginal contributions of agents to all possible coalitions and computing their averaged sum leads to a fair division of the coalition value to all members using the agents' Shapley values [13]. For bilateral coalition negotiation we obtain a more simple term for its calculation. We assume that the agents know of the coalition value function and agree on the utility division method. The currently used coalition negotiation algorithm for building fair coalition configurations bases on [14] (cf. section 5.2).

Table 1. Let $\mathcal{A}$ set of n FCSI agents, $a_i, a_k, a_j \in \mathcal{A}$, $\mathcal{C} \subseteq \mathcal{A}$, $p(t^{a_k}_{tid_x, a_j})$ knowledge-based production of agent a_k for search task t_{tid_x} it received from agent a_j, $Prod_{a_k, C}$ set of agent a_k's productions for search tasks received from agents of potential coalition $\mathcal{C} \subseteq \mathcal{A}$, $\mathcal{C}_i$ coalition without a_i, $\mathcal{C}^{\pi}_i$ set of agents without agent a_i according to the permutation π on $\mathcal{A}$ before a_i ($\mid S_n \mid = n!$).

Agent Utility Function	$U^{type}_{agent_{id}}(p(t^{a_k}_{tid_x, a_j}))$, with production $p(t^{a_k}_{tid_x, a_j})$
Coalition Value	$v(\mathcal{C})$: $\mathcal{P}(\mathcal{A}) \mapsto \mathbb{R}^+$,
	$v(\mathcal{C}) := \sum_{a_k \in C,\ p \in Prod_{a_k, C}} U_k(p)$
Marginal Contribution of Agent a_i	
to Coalition $\mathcal{C}_i$	$v(\mathcal{C}_i \cup \{a_i\}) - v(\mathcal{C}_i)$
Self-Value of Agent a_i	$v(\{a_i\})$
Shapley-Value of agent a_i	$sv_C(a_i) := \frac{1}{n!} \sum_{\pi} (v(\mathcal{C}^{\pi}_i \cup \{a_i\}) - v(\mathcal{C}^{\pi}_i))$
bilateral Shapley-Value for 2-agent entities	$sv_{\{a_i, a_j\}}(a_i) = \frac{1}{2} v(\{a_i\}) + \frac{1}{2}(v(\{a_j, a_i\}) - v(\{a_j\}))$
Individual Rationality	$sv_C(a_i) \geq v(\{a_i\})$

Different types of agent utility functions lead to respectively different coalition types.

5.1 FCSI Coalition Types

One FCSI coalition type, the so-called *task interaction coalition type* ($\mathtt{C}_{ti}$) bases on the notion of FIND-task *interaction*. This means that one terminological dependency, which satisfies the goal of the received as well as an own FIND-task, was found. The amount of such task interactions determines the agent's utility on its task-oriented productions, i.e. the terminological dependencies, for this special type of coalition (cf. Table 2).

Table 2. Task Interaction Coalition Type $\mathtt{C}_{ti}$

FIND-*task interaction*	$t^{a_i}_{tid_1,a_j} \xrightarrow{<intrel>(o_k,o_j,M,N)} t^{a_i}_{tid_2} :\Leftrightarrow$
	i-IDD $< intrel >$ satisfies <u>both</u> FIND-Tasks
$\mathtt{C}_{ti}$-*utility function* U^{ti}_k	$U^{ti}_k(p(t^{a_k}_{tid_x,a_j})) :=\mid \{t^{a_k}_{tid_x,a_j} \xrightarrow{<ir>(o_k,o_j,M,N)} t^{a_k}_{tid_y}\} \mid\in \mathbb{N}_0$
$\mathtt{C}_{ti}$-*coalition value*	$v_{ti}(\mathcal{C}) = \sum_{a_k\in\mathcal{C},p\in Prod_{a_k,\mathcal{C}}} U^{ti}_k(p)$

Each rational FCSI agent tries to coalesce with other FCSI agents which would maximize its own utility in satisfying its FIND-tasks as much as possible. The detailed algorithm for local determination of task interactions between two FCSI-agents can be found in [16].

Example 2:

Consider the FCSI agents $\{a_1, a_2\} = \mathcal{A}$ as in the examples above, and suppose both aspect terms as part of mutually exchanged FIND-tasks $t^{a_1}_{y,a_2}$, $t^{a_2}_{x,a_1}$. According to their respective local knowledge about each other, both agents are able to determine FIND-task interactions, e.g.
$t^{a_1}_{y,a_2} \xrightarrow{p-intsub(Student,Person,\{vid_1\},\{vid_2\})} t^{a_1}_{x}$ by agent a_1. Further, let their self-values be $v_{ti}(\{a_1\}) = 0$ and $v_{ti}(\{a_2\}) = 3$, i.e. only agent a_2 can satisfy 3 of its own FIND-tasks exclusively by itself through considering all reflexive task-interactions induced by dependencies relating local objects.
This yields $v_{ti}(\{a_1, a_2\}) = v_{ti}(\{a_1\})+v_{ti}(\{a_2\})+U^{ti}_1(\{p-intsub(Student, Person, \{vid_1\}, \{vid_2\})\})$ $+ U^{ti}_2(\{p - intsub(Student, Person, \{vid_1\}, \{vid_2\})\}) = 0 + 3 + 1 + 1 = 5$, thus for the marginal contribution of a_1 to $\{a_2\} : 5 - 3 = 2$ and in turn for a_2: 5 - 0 = 5, which leads to the agents' bilateral Shapley-values $sv_{\{a_1,a_2\}}(a_1) = 1$, $sv_{\{a_1,a_2\}}(a_2) = 4$. Since their fair individual rationality is fulfilled and since no better offer from other agents exists, both agents try to coalesce with each other. They are then mutually committed to get access to all respective schema-views by some now mutually allowed intensional data requests. □

There exist other coalition types within the FCSI which do not restrict the productions utility to such mutual task satisfaction. E.g. the C_{atsat} coalition type relies exclusively on the amount of satisfied own and submitted FIND-tasks, wherein the agent utility for the C_{otsat} type considers the amount of all satisfied own FIND-tasks independent from the fact that they were submitted to the satisfying agent. The 'benevolent' agent utility function for C_{tserv} coalitions considers only the satisfaction of the received tasks of an agent as a server. In addition, we currently investigate agent utility functions which take in addition the *quality of search task satisfactions* into consideration. Such utility bases heavily on uncertain terminological interdatabase dependencies, thus leading to uncertain FIND-task satisfaction. At the physical level one could imagine utility functions which base, for example, on transportation costs, including the amount of visited nodes as well as probable sizes of potentially requested relevant schema-data, and so on.

5.2 Decentralized Coalition Formation Between FCSI Agents

The method for coalition formation between FCSI agents is a modification of the decentralized, bilateral coalition formation algorithm in [14] adapted for the cooperative recognition process within the FCSI [16]. A brief summary of the main steps in the FCSI coalition process is given as follows.

1. **communication**

   `for each agent`:

   ▶ if in the first round:
   send to each agent a set of aspect terms. The local own schema-data which are associated to these terms have to be available for the receiving agent;

▶ if not in the first round: send to each negotiation entity the own self-value.

2. **local calculation** of utilities on coalition participation

   `for each agent`:
   - ▶ produce for each received FIND-task all possible `i-IDDs` (cf. section 3.2)
   - `for each coalition type`:
     - ▶ compute the own utility on these productions and respective coalition values
     - ▶ determine an individually rational *preference list of agents*: ordered list of local agent's 'bilateral Shapley-Values' for particular two-entity coalitions

3. **bilateral negotiation about coalition offer**

   `for each agent`:
   - ▶ consider head of the preference list, means the agent who provides the maximum utility
   - ▶ lookup for an alternative coalition from the tail of the list as follows: calculate an alternative coalition value by successively adding the values from the tail of the preference list until the sum is greater than the particular value of the head[5].

   For the alternative coalition it is required that there are no interdatabase dependencies detected between the head and at least one member of it. Moreover, each pair of members of the alternative coalition should have recognized at least one mutual dependency[6].

   - ▶ request the members of the alternative coalition for the existence of interrelation between them: an interrelation is constituted by the mutual recognition of at least one `i-IDD`;

   Note that this requires just the current value of a boolean flag and not any further information concerning which schema-data the respective interdatabase dependency relates. Each flag reflects the existence of an interrelation between the requested agent and the rest of the candidates of the alternative coalition. It will be false if at least one interrelation does not exist.

   - ▶ Upon this information choose between the head and the alternative coalition: having received all the flags from the requested agents of the alternative coalition if all of them are true choose the alternative coalition, otherwise choose the head.
   - ▶ Send coalition offers: if the head was chosen then send its 'Shapley-Value' for the mutual two-entity coalition as a coalition offer, otherwise send the respective values to all the candidates of the alternative coalition.

4. **coalition commitment**

   `for each agent`:
   - `for each coalition type`:
   - ▶ receive offers for coalition formation
   - ▶ select the maximum value of these offers
   - ▶ if and only if the agent who sends this maximum is also preferred

[5] This calculation is limited by 2, at most the local agent's preference value for the head plus 1.
[6] These conditions must hold in order to enable the agents to satisfy the information availability within a coalition (IAC) as mentioned in section 3.

by the (receiving) agent they commit to coalesce with each other
in a proto-coalition.

▶ if all of the candidates of the alternative coalition or atleast one of them have rejected the respective offer, means that this coalition proposal is not valid for the agent anymore: Get the information about the other proto-coalitions formed by these candidates and then erase all the members of these announced proto-coalitions from the preference list[7].

▶ consider such a proto-coalition as one negotiation entity;
repeat the process until no new proto-coalition is formed and then fix them.

The most calculation consuming step of this algorithm is the local calculation one. There, each agent a_i performs $o(s_i \cdot m_i)$ computational operations [8]. All other steps require less than this amount of computations. There are $o(n^2)$ bilateral inter-agent communications. The number of rounds is limited by $n-1$. Therefore the overall communication complexity is cubic, the computational complexity in terms of classification operations is $o((\sum_{a_i \in \mathcal{A}} s_i \cdot m_i) \cdot n - 1)$.
Each agent can participate at several coalitions of different types by concurrent coalition negotiations for each different type of a FCSI coalition utility. This enables a fine-grained mutual utilitarian assessment between the information agents. Since proto-coalitions are treated as an negotiation entity in the following round during the coalition formation, as an impact of the above mentioned information availability assumption (IAC) all values have to be recomputed. In particular, any coalition entity has to compute its self-value with respect to the agents currently staying outside their coalition and vice versa. This means that only the so-called *contributable amount* of the self-value of an 1- or n-agent coalition entity will be brought in the next calculation step[9]. We can distinguish between two kinds agent's rational behaviour: being fairly rational in the sense that it compares the received coalition offers, means the bilateral Shapley-values as expected participation utilities, with the respectively contributable amounts of its self-value; being selfish rational if these comparisons are done with respect to its whole self-value which is not necessarily known by the other agents. Since the second type may lead to coalitions of deceiving information agents, hence violate the IAC assumption, FCSI agents are committed to the first type of individually rational behaviour.

6 Conclusion

The shortly proposed FCSI realizes a cooperative, decentralized search for semantically related data in heterogeneous and autonomous databases by the use of:

- *Local, Terminological Information Models*
  on top of local conceptual database schema;
- Recognition of *terminological dependencies* (i-IDD) and
  respectively induced *interdatabase schema assertions* (IDSA);
- *Utilitarian Coalition Formation* between FCSI Agents;

[7] Note that this deletion prevents considering the possible number of partitions which would result in an overall exponential complexity of coalition formation. The same is valid in [14].

[8] Note that $\mid \mathcal{A} \mid = n$, let $s_i := \sum_{a_j \in \mathcal{A} \setminus \{a_i\}} \mid RecT_{a_j}^{a_i} \mid \cdot m$, with m_i number of local aspect terms from a_i's `LIM` which have to be classified against search (aspect) terms (included in the set of search tasks of a_i from a_j: $RecT_{a_j}^{a_i} := \{t_{tid_x, a_j}^{a_i}\}$) received from other agents in $\mathcal{A} \setminus \{a_i\}$.

[9] Based on the recomputation of contributable amounts of coalition values it is possible that such a value for a potential coalition is *less* than the sum of both entities self-values like in a non-superadditive environment.

– *Directed Intensional Requests* for semantically related data.

Thus it enables the user to discover some intensionally relevant data without the need to browse through all available schema structures first without any help. In particular it is even not possible to get access to the local schema or state level before any utilitarian coalition commitment with other rational agents is fixed. Data Dependency recognition is exclusively done by the FCSI agents at information type level, i.e. by formal terminological classification of some considered aspect term independent from its actually attached and sofar protected data structures. This is similar to the idea of incrementally building and using some shared ontology for contextual interchange[4] respecting association autonomy[31]. Thus, possible data sharing as well as proposing global integrity constraints (IDSAs) is determined by the necessarily prior success and kind of such mutual aspect term classification and restricted on the respective aspect valuations specified by user. There has been only little related research on using formal terminological classification for the object discovery problem like in [9]. Recent works on system approaches which have influenced the FCSI approach are in particular [4] and [12]. Other partially related works are [2] and [3]. In contrast to our work, [33] aims for schema integration. Since cooperation as well as the acquisition and propagation of knowledge about data relevance between autonomous FCSI agents rely on decentralized, utilitarian coalition building there exists no global information agent [3] or central mediator agent [4]. Terminological representation and corresponding classification enables in particular an automated, local processing.
The main benefit of the proposed algorithm for coalition formation is that in the case of non-superadditive environments it prevents building a coalition which after its formation needs an internal, complex commitment agreement about data sharing among its members. This is because in contrast to the original application of the coalition algorithm of [14] as it is presented in [18] each agent receives more information about existing interrelations among its preferred agents. Exchange of this additional information is done without violating the demanded autonomy of the respective agents. Towards an implementation of FCSI agents an *Interactive Development Environment for the specification and simulation of Agent Systems* IDEAS [20] has been recently implemented on a network of SUN-workstations. Ongoing research on the FCSI includes the formal description of an FCSI agent, its possible implementation in IDEAS and further investigations on coalition formation in the FCSI for cooperative information search.

Acknowledgements:
We would like to thank Prof. Dieter Klusch, Prof. Peter Kandzia and Prof. Sarit Kraus for supporting this work by giving many helpfully hints and advices[10].

References

1. Baader,F., Hollunder,B., 1991, "A Terminological Knowledge Representation System with complete inference algorithms", LNAI 567, Springer
2. Beck, H.W., et al., 1989,"Classification as a query processing technique in the CANDIDE SDM",IEEE Computer
3. Barbuceanu,M., Fox,M.S., 1994,"The information agent: an infrastructure for collaboration in the integrated enterprise", Proc. CKBS-94, Keele(UK)
4. Behrendt,W., et al., 1993,"Using an intelligent agent to mediate multibase information access",Proc. CKBS-93, Keele(UK)
5. Blanco,J.M. et al., 1994, "Building a federated relational database systems: an approach using a knowledge-based system", Intern. Journal on Intelligent Coop. Inform. Syst. 3(4)
6. Borgida, A. et al., 1989,"CLASSIC: a structural data model for objects",ACM SIGMOD

[10] This material also bases on work supported in part by the NSF under grant No. IRI-9123967 and the Israeli Science Ministry grant No. 6288.

7. Brachman,R.J., Schmolze,J.G., 1985, "An overview of the KL-ONE knowledge representation system", Cognitive Science
8. Bright,M.W., Hurson,A.R., 1991, "Multidatabasesystems: an advanced concept in handling distributed data", Advances in Computers
9. Catarci,T., Lenzerini,M., 1993,"Representing and using interschema knowledge in cooperative Information Systems", in IJICIS 2(4)
10. Ceri, S., Widom, J., 1992,"Managing semantic heterogeneity with production rules and persistent queues", Politecnico Milano TR 92-078
11. Elmagarmid, E., Zhang, A., 1992,"Enforceable interdatabase constraints in combining multiple autonomous databases", Purdue Tech. Rep. CSD-TR-92-008
12. Hammer,J., et al., 1993,"Object discovery and unification in FDBS",IEEE RIDE-93,Wien
13. Kahan/Rapoport, 1984, *Theories of coalition formation*, Lawrence Erlbaum, London
14. Ketchpel, S., 1993,"Coalition formation among autonomous agents", Proc. MAAMAW-93
15. Ketchpel, S., 1994, "Forming coalitions in the face of uncertain rewards", Proc. AAAI-94, Seattle, Washington
16. Klusch, M., 1996, *Rational kooperative Erkennung von Interdatenbankabhängigkeiten*, Dissertation, Computer Science Dept., University of Kiel, (in preparation)
17. Klusch, M., 1994,"Using a cooperative agent system for a context-based recognition of interdatabase dependencies", Proc. CIKM-94 Workshop on 'Intelligent Information Agents', Gaithersburg (USA)
18. Klusch, M., 1995, "Cooperative Recognition of Interdatabase Dependencies", ACM SIGMOD Proc. 2. Intern. Workshop on Advances in Databases and Information Systems, Moscow
19. Klusch, M., 1995, "Coalition-based cooperation between intelligent agents for a contextual recognition of interdatabase dependencies", Proc. ICMAS-95, San Francisco
20. Klusch, M., 1996, "Utilitarian coalition formation between information agents for a cooperative discovery of interdatabase dependencies", appears in: S. Kirn / G. O'Hare (Eds.), *Cooperative Knowledge Processing*, 1996, Springer Verlag, London
21. Kraus, S., 1993, "Agents contracting tasks in non-collaborative environments", Proc. AAAI-93, Washington D.C.
22. Litwin,W., 1985,"An overview of the Multidatabase System MRDSM",ACM Nation.Conf.
23. Nebel, B., 1990, *Reasoning and revision in hybrid representation systems*, LNAI 422, Springer
24. Nebel, B., 1990, "Terminological reasoning is inherently intractable", AI 43
25. Papazoglou et al.,1992,"An organizational framework for intelligent cooperative IS",IJICIS-1(1)
26. Sandholm,T., Lesser,V., 1995, "Coalition formation among bounded rational agents", Proc.IJCAI-95, Montrèal
27. Schmidt-Schauss,M., 1989, "Subsumption in KL-ONE is undecidable", Proc. 1. Int. Conf. on Principles of Knowledge Repr. & Reason., Toronto
28. Shehory,O., Kraus,S., 1993, "Coalition formation among autonomous agents: Strategies and complexity", Proc. MAAMAW-93,Neuchâtel
29. Shehory,O. Kraus,S., 1994, "Feasible formation of stable coalitions in general environments", Technical Report, Institute for Advanced Computer Studies, University of Maryland
30. Shehory,O. Kraus,S., 1995, "Task allocation via coalition formation among autonomous agents",Proc. IJCAI-95, Montreal
31. Sheth,A., Larson,J.A., 1990,"Federated database systems for managing distributed, heterogeneous and autonomous DBS", ACM CS 22(3)
32. Sheth, A., et al., 1991,"Specifying interdatabase dependencies in a MDB environment", IEEE Computer (see also Bellcore TM-STS-018609/1)
33. Sheth,A. et al., 1993,"On automatic reasoning for schema integration",Int.Journ.ICIS,2(1)
34. Smolka/Schmidt-Schauß,1991,"Attributive concept description with complements",AI-48
35. Sycara,K., 1990, "Persuasive argumentation in negotiation", Theory and Decision, 28:203–242
36. Wooldridge,M., Jennings,N., 1995, "Intelligent Agents: Theory and Practice", Knowledge Engin. Review
37. Zlotkin,G., Rosenschein,J.S., 1991, "Cooperation and conflict resolution via negotiation among autonomous agents in noncooperative domains", IEEE Transactions SMC, 21(6):1317–1324
38. Zlotkin,G., Rosenschein,J.S., 1994, "Coalition, cryptography, and stability: Mechanisms for coalition formation in task oriented domains", Proc. of AAAI-94, Seattle, Washington

This article was processed using the LaTeX macro package with LLNCS style