A Peer-to-Peer Advertising Game*

Paolo Avesani and Alessandro Agostini

ITC-IRST,
Via Sommarive 18 - Loc. Pantè, 38050 Povo, Trento, Italy
{avesani,agostini}@itc.it

Abstract. Advertising plays a key role in service oriented recommendation over a peer-to-peer network. The advertising problem can be considered as the problem of finding a common language to denote the peers' capabilities and needs. Up to now the current approaches to the problem of advertising revealed that the proposed solutions either affect the autonomy assumption or do not scale up the size of the network. We explain how an approach based on language games can be effective in dealing with the typical issue of advertising: do not require ex-ante agreement and to be responsive to the evolution of the network as an open system. In the paper we introduce the notion of advertising game, a specific language game designed to deal with the issue of supporting the emergence of a common denotation language over a network of peers. We provide the related computational model and an experimental evaluation. A positive empirical evidence is achieved by sketching a peer-to-peer recommendation service for bookmark exchanging using real data.

1 Introduction

A recent evolution of architectures for distributed systems attempts to overcome the narrow view of client-server approach to promote a fully distributed view, where every host can play both the role of service provider and service consumer at the same time. Napster [18] and Gnutella [10] are only the most well known examples of peer-to-peer architectures, mainly designed to support file sharing. However, this kind of architectures are going to be used in the field of e-learning [8], database [9] and knowledge management [13].

The peer-to-peer view [17] sustains a service oriented approach with the design and deployment of software components. In such a case a service may perform a task on demand but at the same time it may become a consumer of another service to accomplish the original commitment. In this twofold perspective of provider and consumer, a common language is crucial to support the peers interoperability independently from the specific role. Let us consider a very simple scenario of information retrieval: we need a language to support the service advertising, for example to express capability like *"I'm able to deliver*

* This work was funded by Fondo Progetti PAT, EDAMOK (*"Enabling Distributed and Autonomous Management of Knowledge"*), art. 9, Legge Provinciale 3/2000, DGP n. 1060 dd. 04/05/01.

M.E. Orlowska et al. (Eds.): ICSOC 2003, LNCS 2910, pp. 28–42, 2003.

contents on topic x", and we need a language to express the query, for example to formulate needs like *"I'm looking for contents on topic y"*.

More generally the issue of a common language has to deal with the problem of semantic interoperability, that received an increasing attention after the success of the XML-based protocols. Nevertheless, XML-based protocols succeeded to provide an effective standard for the interoperability at the syntactic level, but the related semantics problem remains an open issue. Usually, the semantics of a new XML-protocol has to be agreed in advance. Such a process of negotiation is performed off-line and it doesn't allow to capture the evolutionary dinamics of an open network of peers. The consortia arranged to manage these agreements on semantics of a given protocol, moreover, are very slow in including new extensions.

A solution to the problem of a common language agreement has been proposed by DAML [11] and OIL [5]. The idea is to design a well defined ontology and refer to it to decode the semantics of a given interaction protocol [12,19,20]. We can refer to these approaches as solutions based on an *ex-ante* agreement: first let agree on semantics, then let use it. A drawback of ex-ante approaches is the underlying assumption of a centralized management of knowledge representation. As mentioned before, this solution is not responsive with respect to the ontology evolution and it contradicts the working assumption that aims to see the peers as autonomous (and not only distributed) sources of knowledge.

More recently there is a new kind of approach that aims to preserve the autonomy assumption while supporting an *ex-post* agreement view: first let use a semantics, then let map it to others. The basic idea is to allow the single peers to define its own semantics and then finding a pairwise mapping with other peer's semantics [1,4,7,14,15]. While this way to proceed represents a meaningful enhancement with respect to the ex-ante approaches, a couple of factors are neglected that are very crucial in a peer-to-peer architecture. The first is that, in an open world, peers join and leave the network; a mapping-based solution doesn't provide the opportunity to exploit past mapping efforts when a new peer join the network and a new custom mapping must be defined from scratch. The second critical factor is related to the scalability issue. Since a peer has to mantain a pairwise mapping for each other peer of the network, this solution requires a quadratic effort with respect to the size of the peer network.

We argue that instead of pursuing a pairwise custom language, the interoperability effort should be devoted to achieve a common language shared by all the peers. Language games [21,23], introduced by Steels in robotics [22], can be considered a powerful tool to support the emergence of a common language among a community of peers preserving their autonomy. Naming games, a specific type of language game, allow to achieve a shared denotation language through an iterative process of pairwise interactions. We claim that naming games can be an effective approach to the challenge of delivering service advertisement in an open network of peers.

In the following we introduce an extension of the naming game model, namely *advertising game*, to deal with the issue of achieving an ex-post agreement on an advertising language. A shared advertising language, differently from the mapping approach, requires only one mapping for each peer, therefore this solution is linear with respect to the size of the peer network. Advertising games differ from naming games because they have to deal with *indirect* feedbacks that introduce a component of uncertainty in the interaction process.

In Section 2 we illustrate a reference example that refers to a recommendation service over a peer-to-peer network. Nevertheless we believe that advertising game can have a great impact even in the field of semantic web and mutiagent systems, where the issue of capabilities language plays a similar role to advertising.

After a brief presentation of the general definition of naming game in Section 3, we introduce the advertising game model in Section 4 and the related computational schema. A more formal definition of the advertising game model is illustrated in [2]. Section 5 is devoted to present the results of the experimental evaluation performed using real world data.

2 A P2P Recommendation Service

To better understand how a language game approach can be effective in supporting the service advertising over a network of peers, let us introduce a reference scenario concerned with the delivery of bookmark recommendation services.

We conceive a community of users where each of them organizes his own bookmarks in the usual fashion of folders and subfolders. A folder can be considered representative of a topic of interest and the folder's contents, i.e. the URI, are the goods that can be shared among the users. Users autonomously collect bookmarks and organize them according to their topics of interest using directory path to uniquely refer to a folder and using a mnemonic label, i.e. the directory name, to denote the semantics of folder contents, i.e. a concept or a category. Of course, we assume to have a peer per user.

Once a user joins the network for the first time he has to deal with the following problems: how to share own topics of interest, i.e. bookmark's folders, and how to look for bookmarks according to his topics of interest. The first issue is concerned with service advertising (advertising language), the second issue is concerned with information retrieval (inquiring language). These issues are associated to the two roles that a peer can play over the network: as service provider and as service consumer respectively. The goal is to define a denotation language that may support both the purposes, advertising and inquiry, exploiting the mutual dependency that holds between an advertising language and an inquiring language.

Let us suppose that our user has a folder devoted to advertising topic. To correctly publish over the network the capability to support recommendation, i.e. new bookmarks, on this topic it is needed to assess what is the right denotation. A correct denotation allows to prevent misunderstanding in taking advantage of the recommendation service: if the user publishes an advertising service recommendation (i.e. capability to deliver advertising related bookmarks), how will be

```
<TopicAdvertisement>
  <Name>...</Name>
  <Topic>...</Topic>
</TopicAdvertisement>
```

```
<TopicQuery>
  <Name>...</Name>
  <Topic>...</Topic>
</TopicQuery>
```

Fig. 1. XML Communication Protocol. On the left hand side a sketch of the protocol to support the advertising of a new topic of interest, i.e. a folder to collect and to share bookmarks according to a predefined category. On the right hand side a sketch of the protocol to support an inquiry over the network to receive recommendations on related bookmarks.

interpreted by other peers? Will the label advertising be interpreted like a capability to deliver bookmarks related to companies that offer TV broadcasting advertisements, or like bookmarks on web services and the related techniques to support the advertising step? The answer of course is user dependent because the right interpretation is given by the expectation of the user that performs a query to other peers using the same denotation. If the user asks for *advertising* bookmarks recommendations, a satisfactory suggestion will include bookmarks that can be stored in the local folder with the same mnemonic label, i.e. interpretations of seeker and provider overlap.

Therefore the mutual dependency between the advertising language and the inquiring language can be exploited to assess the better way to publish over the network the own recommendation service. Before to advertising a new topic the user will try to refer to the other peer's recommendation services checking whether the denotation is consistent with the local working hypothesis. Of course the user can take into account the choice performed by other users or he can ignore them. A denotation not compliant with other users choices will affect the reliability of the services delivered over the network.

The open challenge is to arrange an interaction strategy that brings the peers to adopt a denotation language that reduces the misunderstanding between advertisements of recommendation services and queries for recommendation services. The ultimate goal is to promote the emergence of a shared language where all the users adopt the same denotation for the topic that has to be referred to.

In Figure 1 it is sketched a pairwise protocol to support the communications among peers: a topic advertising protocol and a topic inquiring protocol. Both protocols are defined by a pair: a label to denote the topic of interest, and the reference to the original encoding of a topic (see Figure 2 and Figure 5 for an example). In our example a topic is defined by the name of the directory that hosts the bookmark's folder, by the name of the folder and by the bookmarks stored in the folder.

From the point of view of the single peer a denotation language can be conceived as a mapping between a collection of words and a collection of topics. The name of the folder can be considered the topic's denotation choosen locally by a single peer.

```
<topic>
   <directory>/top/home/cooking/soups_and_stews/fish_and_seafood<\directory>
   <name>fish and seafood</name>
   <bookmark>...</bookmark>
   <bookmark>...</bookmark>
   ...
   <bookmark>...</bookmark>
</topic>
...
<topic>
   <directory>/top/home/cooking/soups_and_stews/beef<\directory>
   <name>beef</name>
   <bookmark>...</bookmark>
   <bookmark>...</bookmark>
   ...
   <bookmark>...</bookmark>
</topic>
```

Fig. 2. Topic XML Schema. A couple of examples of directories extracted from Google's web directories; each directory is defined by an identifier (in this case the full path), a name that should provide a mnemonic support to detect the related category associated to the directory (the local denotation), and finally a collection of bookmarks as defined in a following figure.

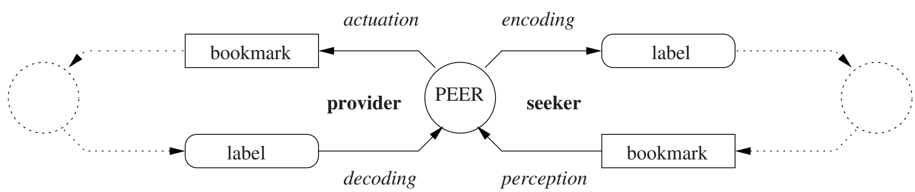

Fig. 3. Twofold Peer's Roles. The schema shows the four basic interactions according to the two roles a peer plays interacting with other peers: provider and seeker.

Figure 3 summarizes what happens in our scenario based on a peer-to-peer bookmark recommendation service. The interaction of a peer with an other peer differs with respect to the role it assumes in the interaction: provider or seeker. When the peer plays the role of provider he receives in input a label that refers to a topic of interest formulated by the sender. In this case the peer performs a *deconding* step to find what kind of topic is denoted by the given label. The next step is an *actuation* step that selects from the related folder the most novel bookmarks and sends them to the inquiring peer. From the other side when a peer plays the role of seeker the first step is an *encoding* step. Given a topic of interest, e.g. a bookmark's folder, the denotation is obtained looking at the lexical relation that binds a word to this topic. The word is sent to the other peers to look for novel bookmarks on related topic. The following step is a *perception* step that is in charge to assess whether the recommended bookmarks, received from the other peers, are compliant with the inquired topic and then can be stored in the associated folder.

It is worthwhile to note that the *encoding* and *decoding* steps are supported by the same common language, i.e. the mapping between words and topics, while the two steps of *actuation* and *perception* allows to be effective in delivering recommendation and to assess the compliance of the language (as by-product of recommendation service invocation).

Let us proceed to show how such a common denotation language can be achieved by an ex-post agreement approach using naming game.

3 A Naming Game Approach

Language games have been introduced to study language formation and evolution interacting with visually grounded robots [22]. A typical kind of language game is "naming," that is, how vocabulary and meanings are learned individually and a shared lexicon eventually emerges in a group of agents. The problem of naming may be expressed in game-theoretical terms, and was extensively studied as naming games in [24]. In short, each "player" (or even agent, peer, ...) has a set of words and a set of objects, and randomly associates a word to an object, called "the topic," to form his local lexicon. It is assumed that all the agents gain a positive payoff in cooperating "but only if they use the same language". A naming game is a coordination game and it is repeatedly played among randomly chosen pairs of players. Thus, a naming game involves a different couple of agents at each repetition of playing. By definition a naming game is adaptive, in the sense that the players in the game change their internal state. A reason for changing is to be more successful in playing future games.

More formally a naming game is defined by a set of peers $\mathcal{P}$, of size $\mathcal{N}_\mathcal{P}$ where each peer $p \in \mathcal{P}$ has a set of objects $\mathcal{O}_p = \{o_1, \ldots, o_n\}$ of size $\mathcal{N}_\mathcal{O}$. The objects are shared among the peers. A lexicon $\mathcal{L}$ is a relation between objects and words, where it is assumed that they are composed using a shared and finite alphabet. Lexicon is extended with a couple of additional information: the number of times the relation has been used and the number of times the relation was in successful use. Each peer $p \in \mathcal{P}$ has his own lexicon drawn from the cartesian product $\mathcal{L}_p = \mathcal{O}_p \times \mathcal{W} \times \mathcal{N} \times \mathcal{N}$, where $\mathcal{W}$ is a set of words and $\mathcal{N}$ the natural numbers to represent the peers' preferences. The lexicon may include synonymous words, two words associated to the same object, and homonymous words, the same word can be associated to two different objects. A peer $p \in \mathcal{P}$ is then defined as a pair $p =< \mathcal{L}_p, \mathcal{O}_p >$.

A naming game is an iterative process where at each step two peers are selected to interact together. Two different roles are given to them: a speaker p_s and a hearer p_h. The interaction proceed as follows. First the speaker p_s randomly selects a topic from his set of objects, then he encodes the topic o_i through a word w_j. The word is choosen accordingly to the current version of the local lexicon $\mathcal{L}_s$ (local to speaker p_s). The denotation of object o_i is obtained looking at the most successful word (a word w_j is more successful than a word w_k iff $< o_i, w_j, u_j, s_j > \in \mathcal{L}_s$, $< o_i, w_k, u_k, s_k > \in \mathcal{L}_s$, $u_j \geq u_k$ and either $s_j/u_j > s_k/u_k$ or $s_j/u_j = s_k/u_k$ and $u_j > u_k$). If the are more successful words a random choice is performed. The hearer p_h decodes the word w_j retrieving the associated object. Whether the object referred by the hearer is the same selected by the speaker both of them give a positive reinforcement to their lexica updating the following relations: $< o_i, w_j, u_j + 1, s_j + 1 >\in \mathcal{L}_s$ and $< o_i, w_j, u_j+1, s_j+1 >\in \mathcal{L}_h$. If the hearer replies with an object $o_l \neq o_i$, it means that the communication failed, the peers' lexicon is updated with a negative reinforcement increasing only the counters of lexical relation (while the counters of successful use of the lexical relation remain the same): $< o_i, w_j, u_j + 1, s_j >\in \mathcal{L}_s$ and $< o_i, w_j, u_j + 1, s_j >\in \mathcal{L}_h$.

Of course a next stage of the game may involve the same pair of peers with inverted roles. After a certain number of iterations, and under given conditions, the game brings the peers to converge to the same lexicon. It means that even though the lexical relation of different peers are not the same, given a topic all peers select the same word as the most successful denotation. In this case the communication between two peers becomes effective because it can't occur a misleading denotation or ambiguous words.

As mentioned before the final result of the game is a common denotation language. It is not stored on a specific server devoted to this purpose but it is encoded in a distributed way. Each peer has his own mapping table, the lexicon, that provides the support for the advertising and inquiring tasks. The distributed representation of the language, although introduces some redundancy, it allows the whole system to be fault tolerant, no one single point of failure, but mostly important it allows the peers to be responsive to the evolution of the language. Language can be evolve because new peers join the network, new words are introduced in the lexicon or new topics have to be delivered over the network. The issues related to the language evolution are deeply analyzed in the spatially distributed game [24] but they are not the goal of this paper.

We prefer to focus our attention to a crucial assumption that underlies the naming games. The key step of the peers interaction is represented by the assessment phase. The assessment step is in charge to check whether, given a word, the two peers refer to the same topic. This test is the precondition of the reinforcement policy because provides a reliable feedback on the lexical relations of the peers. But there is a further inherent condition that has to be satisfied when two peers interact together: the communication channel has to transfer contents that belong to a shared space. It is the case of words but not of the topics. Since topics represent abstract concepts or categories the encoding in terms of bookmarks's folders is local to a single peer. We have already seen in Figure 2 that a peer implicitly defines a topic through a folder pathname and a folder name. It straightforward to notice that in our case it doesn't exist the opportunity for the peers to assess the agreement on lexicon because the topic representations, even according to a common syntax, doesn't refer to a common semantics. Two peers can refer the same topic arranging the bookmark folders in different directories and giving to the folder a mnemonic label that respects their preferences[1].

[1] May be the double denotation of a topic may be misleading. It is important don't confuse the label of the folder, that plays the role of the local denotation of the topic, with the word defined in the lexicon, that globally defines the denotation for the same topic.

A trivial solution to this issue can be arranged looking at the approaches adopted in the semantic web and web services: they suppose the availability of a centralized representation of the topics where their meanings are well defined. It will be in charge of the peers to qualify their topics with respect to such a kind of catalog. The naming game will be accomplished comparing the $c(o_i) = c(o_j)$, where $c : T \rightarrow I$ is a function that takes in input a topic $t \in \mathcal{T}$ and gives in output the index $i \in \mathcal{I}$ of a common referenced representation (where $\mathcal{I}$ is the set of all the meanings indexed by a unique identifier i) . Of course this solution drastically reduces the advantages of a naming game approach.

The challenge is to preserve the language game framework while supporting the assessment step without any further condition of ex-ante agreements.

4 Advertising Game

We have seen in the previous section that if we allow the peers to encode locally their topics of interest, some communication issues may arise. If denotation is matter of negotiation and the representation is autonomously managed the only way to support the assessment of topic meaning is through the exchange of examples.

In our scenario the examples take the form of bookmarks. Bookmarks belong to a common space of the peers and can be shared among them. The link between topics and bookmarks can be defined providing a more detailed view of the two tasks of a peer: actuation and perception.

Actuation can be modeled as a function $f_a : \mathcal{T} \longrightarrow 2^{\mathcal{O}}$ that takes in input a topic and gives in output a subsample of objects. In our scenario bookmarks play the roles of objects and each of them can be considered as an example of a given topic. Actuation function has a stochastic component therefore two subsequent invocations of $f_a(t_k)$ not necessarily produce the same outcome. For example when a peer has to provide a recommendation on a given topic he can sample the related folder selecting the most novel bookmarks; of course the novelty of a bookmark is a time dependent notion therefore the sample may include time by time different bookmarks. From this example it is straightforward to notice that the definition of the actuation function is local to the peer because each of them can have a specific bias in sampling bookmarks for a given topic. Nevertheless, we make the assumption that given a topic $t_k \in \mathcal{T}$ and two peers p_i and p_j, the peers' actuation functions satisfy the following condition:

$$\cup_{n=1}^{\infty} f_{a_{p_i}}^n(t_k) = \cup_{n=1}^{\infty} f_{a_{p_j}}^n(t_k).$$

It means that, independently from the local encoding of the topic, if the meaning selected by two peers is the same, then an infinite iteration of samples produces the same set of bookmarks.

From the other side we model the perception task as a function $f_p : 2^{\mathcal{O}} \longrightarrow \mathcal{T}$ that takes in input a sample of objects, i.e. bookmarks, and gives in output an hypothesis of topic that may subsume such a sample. Of course the hypothesis formulated by the perception function is sensitive of the size of the sample. Given the assumption above on the actuation function, we may conclude that given a sample large enough, virtually infinite, it is possible to assess correctly the topic that underlies the sample generation.

Given the two definitions above, actuation and perception respectively, it s possible now to resume the naming game illustrated in advance and to show how it can be extended in an advertising game. Two are the main variations on the naming game scheme: the first is concerned with the hearer p_h ,the second with the speaker p_s. In the advertising game the hearer once received a word w_k from the speaker, he first decodes as usual w_k in the related topic t_k accordingly with his lexicon, then instead of sending t_k to the speaker, he applies an actuation step communicating $f_{a_h}(t_k)$, i.e. a set of bookmarks representative of the topic t_k. From the other side the speaker, differently from the naming game, doesn't receive a topic t_k but a sample of bookmarks; then he has to perform a perception step, i.e. $f_{p_s}(f_{a_h}(t_k))$, to obtain an hypothesis on the topic selected by the hearer. The assement process can now be carried on easily checking the condition $t_k = f_{p_s}(f_{a_h}(t_k))$.

The schema above introduces the notion of *undirect* feedback because the assessment is inherently uncertain. The uncertainty is related to the reinforcement policy: are we correctly rewarding a positive reinforcement (both denotation and perception hypotheses are correct) or are we erroneously penalizing with a negative reinforcement (drawing wrong conclusions from the error prone perception results)? Of course increasing the amount of examples, i.e. bookmarks, provided each other by the peers, it is possible to reduce the uncertainty virtually to achieve a *direct* feedback.

It is worthwhile to remark that, in the new advertsing game model, peers exchange only words and bookmarks, neither of them affect the assumption of autonomy. Moreover no additional ex-ante agreement is required excepted the communication protocol.

5 Experimental Evaluation

The next step is to put the advertising game model to work to provide some empirical evidence of its effectiveness. Let us resume our scenario concerned with a peer-to-peer bookmark recommendation service. To define a referenced set of topics we looked at the Google web directory [6]. We considered a snapshot of the whole directory, more specifically the Google:Top>Home>Cooking subdirectory. A collection of topics has been encoded accordingly to the XML protocol shown in Figure 2. The single topic has been derived from a node of the Google's structure using the path and the node label as unique identifier. Then accordingly to the bookmark's XML protocol shown in Figure 5 we encode the web pages classified under the given node recording the URI, a web page excerpt and the related preprocessed text that allows to obtain a collection of lemmata as abstract representation of the web page content.

procedure Advertising-Game($\mathcal{P}, \mathcal{L}, \mathcal{F}_a, \mathcal{F}_p$):
Initialize-Lexica($\mathcal{L}_p$)
while not Exit-Condition
$p_s \leftarrow$ Random-Sample($\mathcal{P}$)
$p_h \leftarrow$ Random-Sample($\mathcal{P}$)
$t_s \leftarrow$ Random-Sample($\mathcal{T}_s$)
$w_s \leftarrow$ Lexical-Encoding($\mathcal{L}_s$)
$t_h \leftarrow$ Lexical-Decoding(w_s)
$B_h \leftarrow$ Topic-Actuation(t_h)
$t_h^s \leftarrow$ Topic-Perception(B_h)
if $t_s = t_h^s$
then $\mathcal{L}_s \leftarrow$ Lexical-Reward(t_s, w_s)
else $\mathcal{L}_s \leftarrow$ Lexical-Penalize(t_s, w_s)
endif
endwhile
end Advertising-Game

Fig. 4. Advertising Game Loop. A snapshot of the basic loop of an advertising game. Detailed parameters like the size of the bookmarks samples in the actuation step are omitted.

Figure 4 shows a snapshot of the basic loop designed for advertising game. Actuation functions have been uniformly modelled with a random choice without assigning specific biases to different peers. Perception functions have been modelled by a nearest neighbour classifier based on prototypes [16]. Given a sample of bookmarks, and their related encoding in terms of boolean vector of terms, a prototype is built averaging the different vector representations summarizing a new boolean vector representative of the original sample of bookmarks [3]. The new representation, i.e. the prototype, is compared with the prototypical encoding of the topics through a nearest neighbour rule. In this way it is possible to make an hypothesis on topic given a sample of bookmarks. After the perception step the game follows the same schema illustrated in the naming game model.

The performance of the advertising game is evaluated computing the level of agreement on a common denotation language achieved by the whole set of peers. The denotation agreement is computed looking at a triple $< p_i, p_j, t_k >$ checking whether both p_i and p_j have selected the same word w_h to denote the topic t_k, where $p_i, p_j \in \mathcal{P}$ and $p_i \neq p_j$. The whole agreement is defined as the ratio between the denotation agreements and all the possible communications, i.e. all the triple of the cartesian product $\{\mathcal{P} \times \mathcal{P} \times \mathcal{T}\}$.

The first set of experiments aimed to assess how much the advertising game is effective in supporting the emergence of a common denotation language, even though not all peers play against every other peer. Therefore we arranged an advertising game with 20 peers, 20 topics and 20 words. At each stage a peer playing the role of speaker selected an hearer from a subset of peers, namely his neighbours. We then repeated such a kind of games using a different scope of the peer neighbourhood. Figure 6 shows the results of experiment plotting

```
<bookmark>
   <uri>
     http://www.fish2go.com/rec_0120.htm
   </uri>
   <excerpt>
     Finnan Haddie and Watercress Soup: made with smoked haddock,
     potatoes, watercress, and milk.
   </excerpt>
   <lemmata>
      smoke,watercress,make,haddock,milk,potato,soup
   </lemmata>
</bookmark>
...
<bookmark>
   <uri>
     http://www.bettycrocker.com/default.asp
   </uri>
   <excerpt>
     Crunchy Snacks from Betty Crocker: collection of sweet
     and savory snack recipes which pack a crunch, from healthy
     vegetables to s'mores.
   </excerpt>
   <lemmata>
      snack,collection,recipe,healthy,savoury,vegetable,sweet
   </lemmata>
</bookmark>
```

Fig. 5. Bookmark XML Schema. A couple of bookmark examples extracted from the web directory of Google; each bookmark is defined by its URI, a short description that summarizes the page content, and the result of text processing step that after discarding the stop-words reduces the words to their lemmata.

on the x axis the iterations of a pairwise peer interaction, and in the y axis the percentage of agreement on a common denotation language. Although the restriction of a peer to play only with its neighbours the agreement evaluation was performed considering the hypothesis of full connectivity. The plots shown how the advertising game is effective in supporting the agreement on a common denotation language, although the scope is quite narrow. Even with a set of neighbours based on 30% of the peers the game converges. We didn't explore further smaller subsets of neighbours because in this case it is required to do more precise hypothesis on the network topology.

Thanks to these results we can claim that an advertising game doesn't require a synchronization among the peers to be effective. It means that it is not needed for all the peers to be connected with every other peers. This property is really crucial for the peer-to-peer architectures where the peers join and leave the network asynchronously. But these results are meaningful even from the practical point of view because they enable the opportunity to have an ef-

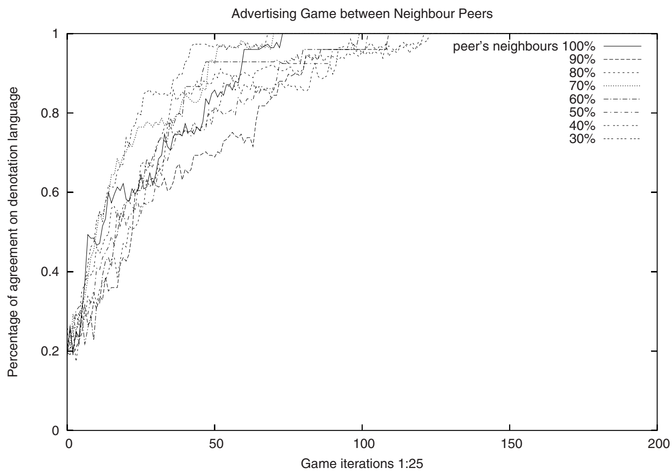

Fig. 6. Advertising Game Evaluation (1). The plot shows the performance of an advertising game where the peers play only with their neighbours. The different curves refer to the scope of the neighbourhood with respect to the size of the network.

fective communication between two peers, i.e. a non ambiguous denotation of a topic, although they have never "meet" before. Looking at our example on bookmark recommendation service it means that if a new user join the network, as a peer, s/he can look for novel bookmark recommendation listening the other peers advertisements without engaging with them a time consuming and bandwidth expensive assessment (of course assuming to have played in advance an advertising game with my neighbours).

The second experiment was concerned with the evaluation of the impact of an indirect feedback on the advertising game. Actuation and perception steps introduced an uncertainty factor because replaced the denoted topics with a sample of their examples. Therefore when the sample of examples is enough large to identify without ambiguity the related topic we fall into the case of naming game because we are facing with a kind of *direct* feedback. We can have different degree of uncertainty depending on the amount of ambiguous examples, i.e. bookmars, may be delivered by the actuation step.

Figure 7 shows the results of such kind of experiment. The different curves in the plot refer to the increasing ambiguity in pairwise peers interaction. It is worthwhile to underline that even though the uncertainty of an *indirect* feedback the advertising games succeed to find an agreement. Still with a 50% of ambiguous examples the network of peers achieves a common denotation language. Differently from the previous case it takes much more interactions to find an agreement, in average four times with respect to a *direct* feedback (pay attention to the different scale of the two plots in Figure 6b). It has to be observed that the performance at the beginning is not meaningful because we didn't ini-

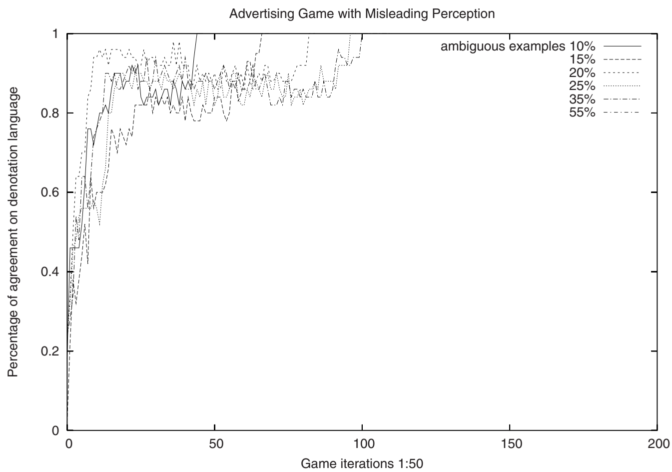

Fig. 7. Advertising Game Evaluation (2). The plot shows the performance of an advertising game when the perception can be misleading. The different curves refer to the percentage of ambiguous objects.

tialize the lexicon of the different peers with a predefined bias, but we simply started with the empty hypotheses.

These results are promising because they provide an evidence that advertising games can be an effective solution to the advertisement issue of distributed systems. Nevertheless, the optimal solution is in finding a trade-off between the bandwidth saving, i.e. small samples of examples, and the ambiguity reduction, i.e. large samples of examples. It would be trivial to enhance the performance of the advertising game without to take into account the scalability issue of bandwidth consumption.

6 Conclusion and Future Work

Advertising plays a key role in delivering service oriented recommendation over a peer-to-peer network. A review of the current approaches to the problem of advertising revealed that the proposed solutions affect the autonomy assumption or don't scale with respect to the size of the network. We explained how an approach based on language games can be effective in dealing with the typical issue of advertising: do not require ex-ante agreement and to be responsive with respect to the evolution of the network scenario both at the level of peers and topics. Nevertheless, naming games are not a satisfactory model for advertising because it doesn't allow to manage a local encoding of topics. We then introduced the notion of advertising game and the related computational model. An empirical evaluation on a real setting data allowed to provide a positive evidence of the proposed model.

As we have already argued, the advertising game is sensitive to the size of the bookmarks samples because if it is increased we overload the network, while if it is decreased we affect the base of induction of the perception. A new challenge arises for the advertising games: how to extend the base of induction without affecting the network overload. The exploitation of the interaction history or strategies based on multicast may help to enhance the current model.

A further critical assumption of the current setting is the homogeneous definition of the perception functions. It is really crucial to investigate what happens when different peers adopt different perception functions or different loss functions.

References

1. Y. Tzitzikas and C. Meghini. *Ostensive Automatic Schema Mapping for Taxonomy-based Peer-to-Peer Systems*. Springer-Verlag LNAI, vol. 2782, 2003.
2. A. Agostini and P. Avesani. Advertising games for web services. In R. Meersman, Z. Tari, and D. Schmit, editors, *Eleventh International Conference on Cooperative Information Systems (CoopIS-03)*, Berlin Heidelberg, 2003. Springer-Verlag LNCS.
3. R. Baeza-Yates and B. Ribeiro-Neto. *Modern Information Retrieval*. Addison Wesley, 1999.
4. Sonia Bergamaschi, Silvana Castano, and Maurizio Vincini. Semantic integration of semistructured and structured data sources. *SIGMOD Record*, 28(1):54–59, 1999.
5. Jeen Broekstra, Michel C. A. Klein, Stefan Decker, Dieter Fensel, Frank van Harmelen, and Ian Horrocks. Enabling knowledge representation on the web by extending RDF schema. In *World Wide Web*, pages 467–478, 2001.
6. Google Web Directory. `http://www.google.com/dirhp`.
7. AnHai Doan, Pedro Domingos, and Alon Y. Halevy. Reconciling schemas of disparate data sources: A machine-learning approach. In *SIGMOD Conference*, 2001.
8. Edutella. `http://edutella.jxta.org`.
9. I. Zaihrayeu F. Giunchiglia. Making peer databases interact, a vision for an architecture supporting data coordination. In *International Workshop on Cooperative Information Agents (CIA-2002)*, volume 2446 of *Lecture Notes in AI*. Springer Verlag, 2002.
10. Gnutella. `http://gnutella.wego.com`.
11. J. Hendler and D.L. McGuinness. Darpa agent markup language. *IEEE Intelligent Systems*, 15(6), 2000.
12. Michel Klein, Dieter Fensel, Atanas Kiryakov, and Damyan Ognyanov. Ontology versioning and change detection on the web. In *International Semantic Web Working Symposium*, 2001.
13. G. Mameli M. Nori M. Bonifacio, P. Bouquet. Kex: a peer-to-peer solution for distributed knowledge management. In *Proceedings of the Fourth International Conference on Practical Aspects of Knowledge Management (PAKM-02)*, volume 2569 of *Lecture Notes in AI*, pages 490–500. Springer Verlag, 2002.
14. Jayant Madhavan, Philip A. Bernstein, and Erhard Rahm. Generic schema matching with cupid. In *The VLDB Journal*, pages 49–58, 2001.
15. Bernardo Magnini, Luciano Serafini, and Manuela Speranza. Linguistic based matching of local ontologies. In *Workshop on Meaning Negotiation (MeaN-02)*, Edmonton, Alberta, Canada, July 2002.

16. Tom M. Mitchell. *Machine Learning*. McGraw-Hill, 1997.
17. G. Moro and M. Koubarakis, editors. *Agents and Peer-to-Peer Computing*, Berlin Heidelberg, 2002. Springer-Verlag LNCS 2530.
18. Napster. `http://www.napster.com`.
19. Wolfgang Nejdl, Boris Wolf, Changtao Qu, Stefan Decker, Michael Sintek Ambjrn Naeve, Mikael Nilsson, Matthias Palmer, and Tore Risch. Edutella: A p2p networking infrastructure based on rdf. *Computer Networks Journal, Special Issue on Semantic Web (to appear)*, 2003.
20. M. Paolucci, T. Kawmura, T. Payne, and K. Sycara. Semantic matching of web services capabilities. In *First International Semantic Web Conference*, 2002.
21. L. Steels. Grounding symbols through evolutionary language games. In A. Cangelosi and D. Parisi, editors, *Simulating the evolution of language*, pages 211–226. Springer Verlag, London, 2001.
22. L. Steels. Language games for autonomous robots. *IEEE Intelligent systems*, pages 17–22, October 2001.
23. L. Steels and F. Kaplan. Bootstrapping grounded word semantics. In T. Briscoe, editor, *Linguistic evolution through language acquisition: formal and computational models*, chapter 3, pages 53–73. Cambridge University Press, Cambridge, 2002.
24. L. Steels and A. McIntyre. Spatially distributed naming games. *Advances in complex systems*, 1(4), January 1999.