Improving Email Conversation Efficiency through Semantically Enhanced Email

Simon Scerri, Brian Davis and Siegfried Handschuh, *Digital Enterprise Research Institute, National University of Ireland Galway, IDA Business Park, Galway, Ireland.*

***Abstract*—Despite persisting in popularity email is still plagued with information overload, hindering the workflow of data the user has to handle. While the revolutionization of the Web by the Semantic Web is underway, we aspire to use the same technology to enhance electronic mail with useful semantics. In so doing we will tackle one of the largest flaws of the email communication genre - the lack of shared expectations about the form and content of the interaction. This can be attributed to the lack of explicit semantics covering the context and content of exchanged email messages. Earlier research showed that email content can be captured by applying speech act theory. We will refine and extend this work to develop an ontology for email speech acts and outline non-deterministic models to support the user in deciding the best course of action upon sending or receiving an email.**

***Index Terms*—Email, Speech Act Theory, Metadata Extraction, Semantic Web.**

I. Introduction

EMAIL REMAINS plagued with the problem of information overload. The email user is faced with too much information, and making decisions on how to act upon different messages is immensely time-consuming. As a result, the processing of incoming messages ends up being postponed, sometimes indefinitely, due to different priorities or because the mental effort required would lead to distraction from other tasks[1]. This seriously hampers the workflow of data in email conversations.

Email has been regarded as a new genre[2], since despite being a written form of communication, email texts are closer to spoken than written discourse[3]. We think that the root of the problem is exposed by the definition of a genre - a patterning of communication which structures communication by creating shared expectations about the form and content of the interaction, thus easing the burden of production and interpretation[4]. Email workflow is very inefficient because it lacks these shared expectations on how, and when, exchanged information should be acted upon. Whereas it should be the sender's interest to make any expectations explicit to the recipient, it is usually the latter that ends up consuming more time trying to extract implicit expectations. Apart from being subject to misinterpretation, this stops the recipient from immediately trying to act upon a message. This leads to a breakdown of the email one-touch model[5], where users try to process information at once, but for one reason or another, fail. We believe that the less effort required from the recipient, the greater the chance the sender's expectations are fulfilled in a timely manner.

The rest of the paper is structured as follows: Section 2 provides an overview of previous related work, while in Section 3. we introduce and discuss in detail our own research ideas. In Section 4. we highlight the conceptual achievements of our work via an example scenario, whereas in Section 5 we present our future evaluation plans and provide some concluding remarks.

II. Background And Related Work

There have been a number of attempts directed at solving the problem we presented, inspiring us to pursue this research. We now give a brief description of this previous work.

Speech Act Theory has long been regarded as an important area when it comes to modelling electronic communicative patterns. In the early 1980's, just about a decade after Searle's modern inception of the theory[6], it was already a very influential topic in computer science, particularly in the design of artificial languages for agent-based communication. The theory states that in saying something one is doing something, and is mainly concerned with the difference between the three meanings of utterances or written text: the Locutionary, or literal meaning; the Illocutionary, or the social function the speaker is performing; and the Perlocutionary, or the result or effect produced in the given context. For the speech act 'Could you please close the door', the Illocutionary force is that the speaker is requesting an action, the Perlocutionary force on the hearer means they are expected to close the door, rather than answering a question with a yes or no - which constitutes the Locutionary meaning. The theory was applied to Email a number of times, in particular to classify email based on the sender's intent[2][3][7][8],[9] detect the focus of threaded email discussions[10], predict actions on email messages[11] and ease task management arising through email[12][13] amongst others. Although these provided promising results the almost complete reliance on natural language processing to guess expectations accompanying incoming messages were somewhat inaccurate to be practical. Another point which was frequently disregarded is that an email message is frequently multi-purpose, realizing several purposes at the same time.

Other relevant research work involved the introduction and formalization of Semantic Email processes[14]. Based on the

Manuscript received March 16, 2007 .

Semantic Web paradigm and implemented in the Mangrove project[15], this involved exchanging messages having predefined intents. One drawback is that users have to resort to predefined templates and this lack of flexibility limits the practicality of the approach. Also, average users are not willing to migrate from an email system that works to a different email system, even if the latter provides less ambiguous dialogue and more efficient results.

III. Semantically Enhanced Email

Although email is rife with weaknesses, it also provides a fundamentally right model for a communication system[16], naming asynchronosity, threading and the fact that it is command central system as the major advantages of the model. Therefore we would like to retain this model and extend the functionalities through adding semantic annotations to email content. We believe that rather than guessing expectations of incoming messages, they should be made explicit prior to leaving the sender and integrated within the message. We would like to achieve our vision by:

- Fine-tuning an existing email speech act ontology presented in earlier work[9] and creating our own email speech act ontology;
- Outlining a predictive model for illocutionary and perlocutionary reactions attributed to speech acts in email messages;
- Outlining a model for speech act discourse in emails within a threaded email conversation;
- Using Text Analytics to aid capturing speech acts in outbound messages and Semantic Web technologies to invisibly embed them as content metadata, together with contextual metadata of email in threaded conversations, in the email header;

In so doing we can substantially reduce the occurrence and consequences of the given problems. In this paper we focus on the first three points and we will elaborate on our ideas in the coming sub-sections.

A. *Email Speech Act Ontology*

We base our ontology on an adaptation to an existing taxonomy[9] where hierarchies of Verb and Noun, whose conjunction forms a speech act as the pair (v-n), where defined. We fine-tuned this work to better reflect our main concerns: the intents and expectations accompanying speech acts. This is reflected in our verb hierarchy in Fig. 1, which differs between the two most basic verb roles at the highest level: *Initiative*, initiating a conversational thread; or *Continuative*, continuing an earlier conversation. The roles can be refined into *Requestive*, when something is being requested out of the recipient e.g. 'Can you go to the meeting?'; *Informative*, when the act is not in response to any request and requires no further dialogue e.g. 'I'm going to the meeting'; and *Responsive*, when satisfying a former request e.g. 'Yes I will go to the meeting'. The *Imperative* role is both a requestive and an informative since its behavior corresponds

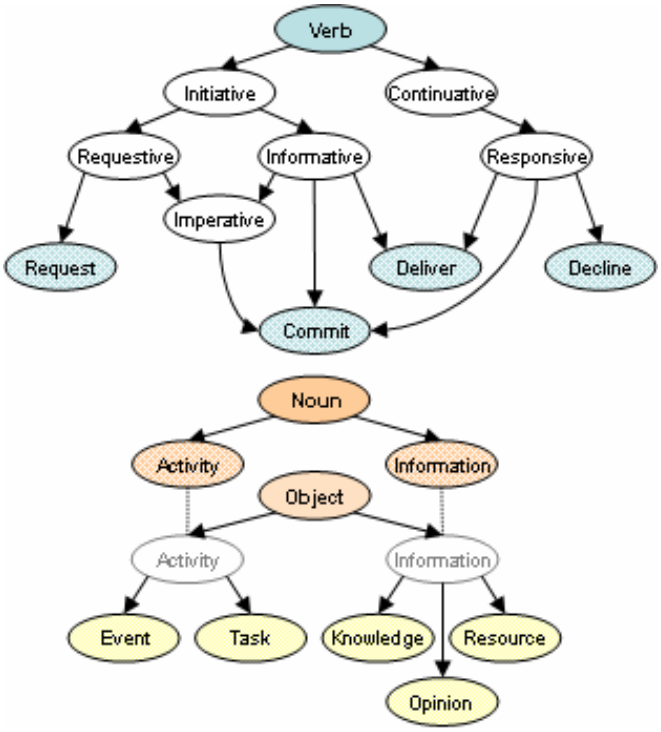

Fig. 1. Speech Act Verb, Noun and Object

to both definitions above, e.g. 'Go to the meeting'. The four end verbs can manifest particular roles in particular situations. Whereas Request and Decline perform a requestive and responsive role respectively, Deliver can double for two roles: 'Here is the requested file' is responsive whereas if the file wasn't requested it is informative. Commit is yet more versatile and can manifest all roles. If the speaker is committing to something the hearer had requested, then it's responsive, whereas if the speaker is informing of a personal commitment that was not requested by the hearer, the role would be informative. If on the other hand the speaker is issuing an order with the aim to commit the hearer, the role of the verb would be imperative since it is both requestive - something is being requested out of the hearer, *and* informative - the speech act is not in response to any request, and no reciprocal communication is required.

In our ontology, we categorize the nouns in two major concepts: *Information*, representing something which occurs strictly within the boundary of email and *Activity*, representing something occurring outside the world of email. We extended our speech act definition to include a *Speech Act Object* representing instances of nouns rather than subclasses. Modeling the workflow and predictions for multiple verb-object pairs can be done by considering the abstract verb-noun pair. *Event* and *Task* are Activity instances and *Knowledge*, *Opinion* and *Resource* are Information instances.

Previous work differed between a speech act requesting permission to attend an event and another requesting someone to attend. We think that these speech acts are fundamentally similar, with the only difference being who is tied to the activity noun in the request: the recipient, the sender or even both. We consider both of them as a Request-Event and differ

Speech Act (Verb-Noun)	Speech Act Object	Activity Subject	Description	Speech Act Role	ER[s]	ER[r]
Request-Activity	Event/Task	Recipient Both Sender	Request recipient to perform activity Request joint activity Request permission for activity	Requestive Requestive Requestive	Expect Expect Expect	Reply Reply Reply
Commit-Activity	Event/Task	Sender Both Recipient	Commit to an activity Commit/Instruct a joint activity Commit/Instruct recipient to activity	Resource/Informative Resource/Imperative Resource/Imperative	Perform Perform None	None Perform Perform
Decline-Activity	Event/Task	Recipient Sender Both	Decline permission for an activity Decline performing an activity Decline performing a joint activity	Responsive Responsive Responsive	None None None	None None None
Request-Information	Knowledge/Opinion/Resource		Request information from recipient	Requestive	Expect	Reply
Deliver-Information	Knowledge/Opinion/Resource		Deliver information	Resource/Informative	None	None
Decline-Information	Knowledge/Opinion/Resource		Decline delivering information	Responsive	None	None

Fig. 2. Realized Speech Acts Combinations and Expected Reactions

instead on the subject of the speech act, i.e, who will attend to the event. Therefore we extended our speech act definition to also include a *Speech Act Subject*, applicable only to speech acts with Activity nouns, where the subject can be the *Sender*, *Recipient*, or *Both*.

Given these two new parameters we define a *Realized Speech Act* (v-(o)[s]); where o denotes possible noun instances and s denotes the subject of activity noun instances, if applicable. Thus, 'Can I go to the show?' becomes a Request-Activity with the sender as the subject and the event as the object, henceforth denoted as (Request-Event[Sender]), 'Can you compile a list of agenda items?' becomes a (Request-Task[Recipient]) and 'Shall we go to the meeting?' becomes a (Request-Event[Both]). On the other hand, speech acts with the information noun, like 'Can you send me the document' becomes (Request-Resource[Ø]), since the speech act subject is not applicable.

B. Predicting Reactions on a Speech Act

The intents and expectations around which our ontology is designed correspond to the illocutionary and perlocutionary forces of the speech acts respectively. We now outline a non-deterministic predictive model to address them. We define the *Illocutionary Expected Reaction* [ERs] as the course of action expected out of the speech act sender on sending, and the *Perlocutionary Expected Reaction* [ERr] as the course of action expected out of the recipient on acknowledgment. We categorize reactions into *Passive* and *Active* reactions. Passive reactions are either *Expect*, where the sender expects a response on sending a speech act; or *None*, where the sender or recipient is expected to do nothing on issuing or receiving the speech act. Active reactions are either *Reply*, when the recipient is expected to reply on getting a speech act; or *Perform*, for speech acts which demand an activity, e.g. a Task, from the sender/recipient on sending/receiving a speech act. We apply this predictive model to our realized speech act definition as (v-(o)[s]) {ERs}→{ERr}, denoting that on sending a speech act specific expected reactions for both sender and recipient are generated.

Not all combinations of the verb-noun pairs in the ontology are relevant. Whereas committing to an event makes sense, committing to a resource does not. Fig.2. is an exhaustive table presenting all relevant speech acts given as the verb-noun pairs, their respective noun instances, and their activity subjects if applicable. A brief description for each realized speech act is given along the verb role and the expected reactions generated for the sender on sending and recipient on acknowledgment. The table highlights the fact that one speech-act can serve more than one role and can thus have more than one predictive force. If a person A requests another person B to attend to an event (Request-Event[Recipient]) then A's speech act has a requestive role. On sending, A expects a response, whereas when reading the email B is expected to reply. On the other hand, if A instructed B to go to the event in the first place (Commit-Event[Recipient]), the role of the speech act is imperative and therefore both informative and requestive. On sending A is expected to do nothing whereas on acknowledging the speech act B is expected to perform.

C. Modeling Speech Act Sequences in Email Conversations

Speech acts in one email may call for further speech acts in consecutive emails. In the previous section we tried to answer the question "What's the expected reaction on sending and receiving a speech act?". In this section the question is "If the email user is in the process of initiating or continuing an email conversation, what speech act/s would most likely be used in the particular circumstance, given previous speech acts in the thread, if any?". In our predictive model, speech acts in a message requiring a responsive speech act are reflected by the 'Reply' perlocutionary force. Given its non-deterministic nature it does not cover all times when further communication will take place. Although on sending 'Please finish the document by tomorrow!' as a (Commit-Task[Recipient]) the sender may expect the recipient to 'Perform' the required activity on acknowledgment, the recipient may act unexpectedly 'Reply' by declining the request.

Standard email threading support provides context retention while getting rid of unnecessary information duplication. Rather then incorporating a thread concept as an additional layer within the email model, we consider the email model to be primarily thread-based. We regard a new email message as

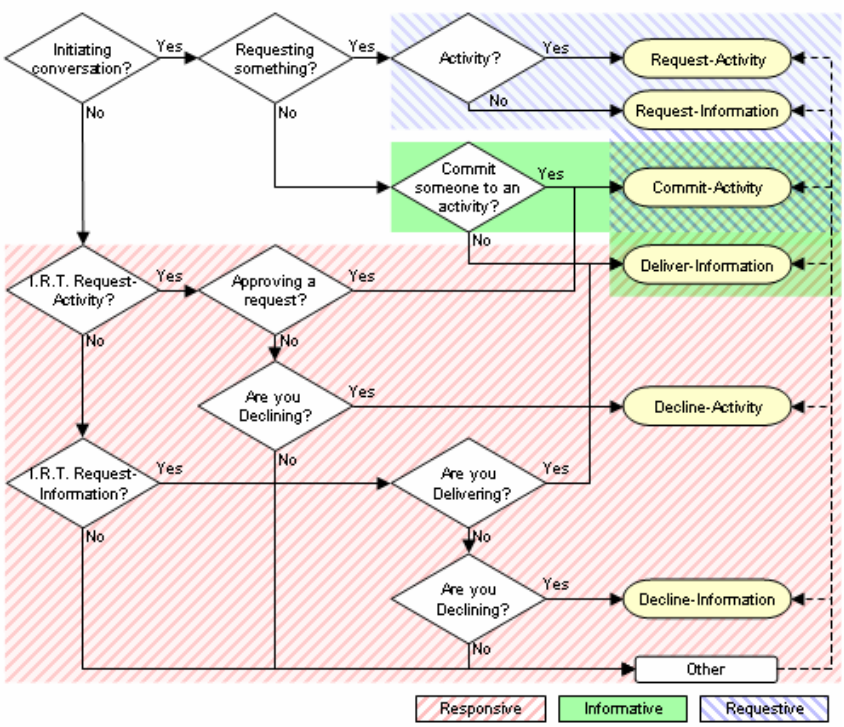

Fig. 3. Speech Act Workflow in Email Conversation

more than simply a new message, but as the first email in an email thread, where a thread is a number of consecutive emails, all stemming from the first email that the thread initiator started. We differ between a global email thread and a personal email thread. A personal thread is equivalent to one participant's personal view of the abstract global thread. In this research we are interested in activities and threads in a personal context, therefore by referring to a thread we imply personal email threads. Given this scenario, the thread initiator is the participant which first introduced a person to an ongoing thread as seen by that person. So if John, after some message exchanges with Paul, adds Mary to the conversation, Mary's (personal) thread starts at that point and her thread's initiator is John, even if the initiator of the conversation between John and Paul, was in fact Paul. The global thread in this case is composed of John's, Paul's and Mary's personal views of the thread, and the global thread initiator is Paul.

Based on the speech act ontology and prediction model, as well as the thread-based approach just described, we try to model sequences of speech acts in new or existing threads. We illustrate this as a flowchart in Fig. 3. The first decision depends on whether the user is initiating a conversation or continuing an initiated conversation. Decisions lead to one particular speech act out of the six verb-noun pairs. The three roles: Requestive, Responsive and Informative, are represented using three different shadings. Since some speech acts can manifest multiple roles, there is an overlap between the shadings, e.g. Commit-Activity is shaded with an overlap of all three roles. When initiating a conversation, the user's choice of speech acts is limited to requestives and informatives, since a responsive speech act implies that a previous thread exists. Although when not initiating a conversation the assumption is that the user requires a responsive speech act (i.e. Deliver, Decline or Commit), the user might choose to reply through another Request, for example. We model this by providing the 'Other' option in the non-initiating part of the flowchart. Therefore, in reply to "What is your colleague's phone extension?" (Request-Information[Ø]) from Paul, John might deliver, e.g. "The number is 03214" or decline, e.g. "I would ask them directly for the number" the information. However John is free to reply with any another speech act, e.g. a request "Sorry but, who are you?". We believe that through this speech act sequence model we correctly capture the non-deterministic flow of speech acts in email conversations.

IV. Achievements: A Case Scenario

Fig.4. gives an insight on how we envision to handle semantically enhanced email dialogues through the terminology and models presented in this paper. The full figure represents the global thread for a particular conversation, consisting of three email threads corresponding to each participant's personal view. The participants have various levels of authority. Paul is a director; Mary is a manager and John an employee. Paul initiates the thread by sending an email to Mary. This email contains a speech act of the form (v-(o)[s]) {ERs}→{ERr} as defined earlier. Paul is requesting whether Mary can perform a required activity (task) for him, hence the activity subject is the recipient. On sending, Paul 'Expect's something back. On acknowledgment, Mary is expected to 'Reply'. Mary eventually responds to John's request, committing to the task. In this case the activity subject is the sender and thus the reaction expected out of Mary on sending is to 'Perform' and her email client might suggest to create a new task in her task list appropriately. On the other hand the recipient, Paul, is not expected to do anything on acknowledging the commitment. Nevertheless Paul goes against this expectation and issues a speech act in a consecutive email to inform Mary and John that Mary will perform a task and should she require extra man-power, John should contribute. John's view of the global thread starts at that point in time, and its initiator is Paul. As both speech acts in this activity deal with delivering information the activity subject is not applicable, and since the purpose was purely informational no reaction is expected neither out of the sender nor the recipients. Some time later Mary decides that the John should help with the task. Therefore she issues another speech act in a consecutive email to instruct him to contribute, and sends a copy to Paul to inform him of the assignment. This results in two speech acts: a Commit-(Task) with the recipient John as the activity subject; and a Deliver-(Information)

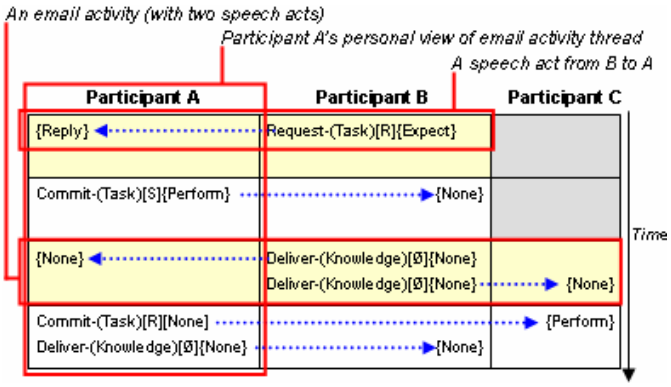

Fig. 4. An email conversation between Mary, Paul and John.

addressed to Paul. Mary is expected to do nothing on sending the activity. Neither does Paul on acknowledging the delivery. However John is expected to perform the task he was committed to, and on acknowledging the commisive speech act his email client can suggest adding the task to his task list.

V. Future Directions and Conclusion

Our future plans include a two-stage evaluation process. The first will consider how well our speech act ontology, predictive model and speech act sequence model fit into real instances of email messages in threaded conversations. We plan to achieve this via statistical means measuring inter-annotator agreement between observed annotations, provided by a number of manual annotators for our Speech Act Ontology using the κ statistic[17]. We intend to make a statistical comparison between our speech act model and the one in [9] by testing them against a common corpus of email threads and comparing the resulting κ values for both models.

After evaluating the first stage results, we plan to implement our models as an extension to a personal information management tool that includes means to synchronously manage the user's emails, events and tasks (e.g. Microsoft Outlook). Inspired by [3], we intend to support the user with semi-automatic content metadata extraction (binding speech acts instances to concepts in our ontology) via text analytics/natural language engineering techniques. Furthermore, we intend to employ Ontology based Information Extraction Techniques and investigate a possible combination of Machine Learning approaches similar to [7].

Eventually we want to enhance email messages semantically with invisible semantic annotations consisting of the acquired content and contextual metadata. We foresee a specific MIME extension allowing for an RDF content-type in the email headers and we have already begun testing this possibility. This ensures that semantic email will be backward compatible to email clients that are not aware of our semantic extension. Once fully implemented, email users will be aided by a smarter email client that eases the data workflow within their email discourse, by predicting their actions on the basis of the semantics accompanying speech acts in email messages. An email sender's expectations will be clear to the recipient on reading the email and the semantically aware email client will support the user by autonomously aiding the workflow of email-generated personal information management. The email client will be able to suggest the most appropriate action for whatever speech act the user is creating in new emails, or getting in incoming emails. This includes full cooperation between Outlook's Mail, Task List and Calendar managers. Rather than going through unread mails, a user will be able to periodically check or even be reminded of speech acts on which they were expected to act upon but never did. Thus using concepts in our ontology, we can semantically annotate individual email messages with information about speech acts contained within. Based on this semantic data, we can predict reactions from both sender and recipient by means of our predictive model. Given our thread-based approach, we can also express the contextual information of messages as semantic annotation. Contextual information for a thread includes the sequence of speech acts within messages in the discourse, and based on this we can aid the user with the selection of speech acts in a successive message, if required.

To summarize, we intend to help the email user decide on the correct action when creating or responding to email messages, and ease the management of personal information generated through email. Finally we believe that the presented ideas in this paper will contribute to improving the efficiency of information workflow in email conversations in general.

References

[1] Whittaker, S., Bellotti, V., Gwizdka, J.: Email and PIM: Problems and Possibilities. CACM.7 (2007).
[2] Goldstein, J., Sabin, R.E.: Using Speech Acts to Categorize Email and Identify Email Genres. System Sciences, HICSS2006 (2006).
[3] Khosravi, H., Wilks, Y.: Routing email automatically by purpose not topic. Natural Language Engineering, Vol. 5 (1999) 237–250.
[4] Erikson, T.: Making Sense of Computer-Mediated Communication (CMC): Conversations as Genres. Hawaiian International Conference on System Services, HICSS2000 (2000).
[5] Whittaker, S., Sidner, C.: Email overload: exploring personal information management of email. CHI '96 Conference Proceedings, Vancouver, BC Canada, April (1996) 276 - 283.
[6] Searle, J.: Speech Acts. Cambridge University Press (1969).
[7] Cohen, W., Carvalho, V., Mitchell, T.: Learning to Classify Email into "Speech Acts". Proceedings of the EMNLP, Barcelona, Spain (2004).
[8] 8. Carvalho, V.R. and Cohen, W.W. On the collective classification of email speech acts. In Proceedings of SIGIR-2005, (2005) 345-352.
[9] Carvalho, V., Cohen, W.: Improving Email Speech Act Analysis via N-gram Selection, ACTS Workshop, New York City, NY (2006).
[10] Feng, D., Shaw, E., Kim, J., Hovy, E.: Learning to detect conversation focus of threaded discussions. Human Language Technology Conference, New York (2006) 208-215.
[11] Dabbish, L., Kraut, R., Fussell, S., Kiesler, S.,: Understanding email use: predicting action on a message. SIGCHI conference on Human factors in computing systems (2005) 691-700.
[12] Khoussainov, R., Kushmerick, N.: Email task management: An iterative relational learning approach. CEAS'2005 (2005).
[13] Corston-Oliver, S., Ringger, E., Gamon, M., Campbell, R.: Task-focused summarization of email. Proceedings of the Text Summarization Branches Out workshop, ACL 2004 (2004).
[14] McDowell, L., Etzioni, O., Halevey, A., Levy, H.:Semantic email.World Wide Web, (2004)
[15] McDowell, L., Etzioni, O., Gribble, S., Halevey, A., Levy, H., Pentney, W., Verma, D., Vlasseva, S.: Evolving the Semantic Web with Mangrove. UW Tech Report, (2003).
[16] Armstrong, E.: What's wrong with email? Bootstrap Alliance (2000)
[17] Carletta, J.: Assessing Agreement on Classification Tasks: The Kappa Statistic. ComputationalLinguistics, Vol. 22, No. 2 (1996). 249-254.