Evaluating the Ontology underlying sMail – the Conceptual Framework for Semantic Email Communication

Simon Scerri, Myriam Mencke, Brian Davis, Siegfried Handschuh

Digital Enterprise Research Institute,
National University of Ireland Galway
IDA Business Park, Galway, Ireland.
E-mail: simon.scerri@deri.org, myriam.mencke@deri.org, brian.davis@deri.org, siegfried.handschuh@deri.org

Abstract

The lack of structure in the content of email messages makes it very hard for data channelled between the sender and the recipient to be correctly interpreted and acted upon. As a result, the purposes of messages frequently end up not being fulfilled, prompting prolonged communication and stalling the disconnected workflow that is characteristic of email. This problem could be partially solved by extending the current email model to support light-weight semantics pertaining to the intents of the sender and the expectations from the recipient(s), thus leaving no room for ambiguity. Semantically-aware email clients will then be able to support the user with the workflow of email-generated tasks. In line with this thinking, we present the sMail Conceptual Framework. At its core, this framework has an Email Speech Act Model. Given this model, email content can be categorized into a set of speech acts, each carrying specific expectations. In this paper we present and discuss the methodology and results of this model's statistical evaluation. By performing the same evaluation on another existing model, we demonstrate our model's higher sophistication. After careful observations, we perform changes to the model and subsequently accommodate the changes in the revised sMail Conceptual Framework.

1. Introduction

Although email has now entered the fifth decade of its existence, it is still an essential feature of the internet. However, over time email became plagued with the problem of information overload (Whittaker & Sidner, 1996) - the email user is faced with too much information, and making decisions on how to act upon different messages is immensely time-consuming. Thus, due to different priorities or because the mental effort required to do so would lead to distraction from other tasks (Khosravi & Wilks, 1999), processing of incoming messages ends up being postponed, sometimes indefinitely, seriously hampering the data workflow in email conversations. Email has been defined as a new genre (Goldstein & Sabin, 2006) and we believe that the root of the problem is exposed by the very 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 (Erikson, 2000). The exchange of information in email's disconnected workflow is very inefficient because it lacks these shared expectations on how, and when, the information should be acted upon.

We attribute this problem to the total lack of clear semantics defining an email message's purposes. Devoid of any structure, the content of an email message is rife with ambiguities and the successful fulfilment of its purposes (if any) is subject to the recipient's interpretation. Given the high amount of task-related emails, this ends up hindering the workflow between individuals in typical business environments. Whereas it would not be practical to incorporate heavy semantics within email messages, we believe that by extending the current email model to support light-weight semantics pertaining to the purposes of email messages, we can substantially reduce the occurrence and consequences of these problems. Semantically-aware email clients will then be able to support the user with the workflow of email-generated tasks. In an outline of this work (Scerri & Davis & Handschuh, 2007) we discussed how we envision to ease this problem. In this paper we refined earlier work into the novel *sMail Conceptual Framework,* which is based on three entities: the *Email Speech Act Model*; the *Email Speech Act Process Model* (previously referred to as the Speech Act Prediction Model); and the *Email Speech Act Workflow.* The sMail Ontology[1] represents the knowledge in the first two models. In this paper we will re-introduce our conceptual framework and discuss the methodology and results of the evaluation of the Speech Act Model as represented in the sMail Ontology.

2. Background

In the sMail approach, we combine earlier ideas concerning Speech Act Theory (Searle, 1969) and semantically-enhanced email (McDowell et al, 2004). This is done with the aim to target and ease the infamous problems of email overload and the resulting personal information mismanagement (Whittaker & Bellotti & Gwizdka 2007).

Speech Act Theory has been very influential in modeling electronic communicative patterns. The core idea behind the theory is that every explicit 'speech' has one or more associated implicit 'acts'. It highlights the difference between three different meanings of utterances, being the Locutionary (literal meaning), the Illocutionary (social function the speaker is performing) and the Perlocutionary (the result or effect on the hearer in the

---

[1] http://smile.deri.ie/ontologies/smail.rdfs

given context). The utterance "*Could you please close the door?*" can thus be associated with three different speech act forces. While the Locutionary force would see the speaker asking a 'yes or no' question, the Illocutionary force sees the speaker requesting an action, while the Perlocutionary force places an expectation on the hearer to close the door.

The second approach that is highly relevant to our work involved the introduction and formalization of Semantic Email processes (McDowell et al, 2004). These were the basis of an email client application implemented within the Mangrove project (McDowell et al, 2003), which exchanged messages having predefined, explicit intents, through semantic web technologies. Thus these semantic messages left no room for ambiguity. The main problem we identified is that given this approach, although the processes are given clear semantics, the email system is not very flexible and users would have to resort to fixed templates. In our vision, such a system should also offer the possibility of capturing the semantics and categorizing email process as they are created on the fly. This would be more inline with the flexible, ad-hoc nature of email communication.
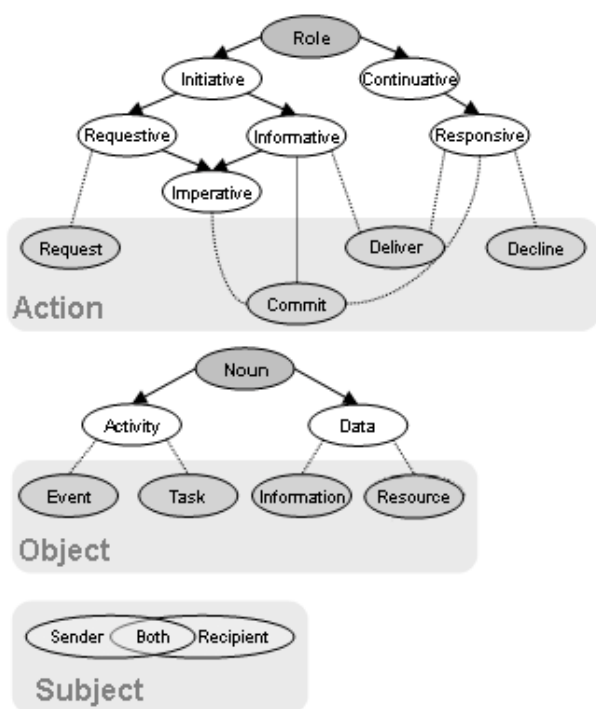

Figure 1 : Speech Act Action, Noun and Object

The sMail Conceptual Framework combines the best ideas from both approaches. Email content will be annotated with instances of the sMail Speech Act Model (henceforth referred to as the sMail model) which is itself based on a succession of previous work in the area, most notably by Carvalho & Cohen (2006).

In brief, the model (Fig.1) represents hierarchies of *Actions* (previously referred to as *Verbs*) and *Objects*, whose conjunction forms a speech act as a pair (*a-o*). Our action hierarchy is more discourse oriented, and differs between different *Discourse Roles*. There are two basic

roles at the highest level: *Initiative*, initiating discourse; and *Continuative*, otherwise. These are refined into *Requestive*, when something is being requested from 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 behaviour corresponds to both definitions above, e.g. "*Go to the meeting*". Some of the action instances serve particular roles in different situations, e.g. *Deliver* can double for two roles: as a response to a request or as an informative. Objects are categorized in two major concepts, or *Nouns*: *Data*, representing something which occurs strictly within the boundary of email (e.g. information) and *Activity*, representing something occurring outside the world of email (e.g. an external action resulting from email). We extended the sMail speech act definition to include a *Speech Act Object* that represents instances of the nouns. *Event* and *Task* are activity objects whereas *Information* and *Resource* are data objects. We then introduced a very important parameter to our speech act definition. Previous taxonomies differed between a speech act requesting permission from the recipient to attend an event and another requesting the recipient to attend. We think that these speech acts are fundamentally similar - the only difference being whether the recipient or the sender is tied to the activity in the request. Speech acts can also have both sender and recipient tied to the activity, e.g. "*Can we have a meeting today?*" We therefore extended our speech act definition to also include a *Speech Act Subject*, where the subject can be the *Sender*, *Recipient*, or *Both*. The subject is only applicable to speech acts with activity noun objects. Given the new parameters we defined the sMail Speech Act as the triple (*a,o,s*); where *a* denotes the action, *o* the possible objects and *s* the subject of activity objects (if applicable).

The second model in the sMail Conceptual Framework is the Email Speech Act Process Model. This model considers each speech act as a separate process. Speech act theory highlights the three forces of utterances - the Locutionary (literal meaning); the Illocutionary (social function the speaker is performing); and the Perlocutionary (the result or effect on the hearer in the given context). The sMail Speech Act Process Model models the latter two forces for all combinations of speech acts given in our model. Given that this model is dependent on the sMail Speech Act Model, it will be re-introduced following the improvement of the latter.

Finally, the sMail Email Speech Act Workflow will support the flow of speech acts in email discourse (threads) based on the two previous models. It will support the user in making decisions on which speech act most likely applies to what they are writing in the email, irrespective of whether they are initiating a conversation or reacting to an incoming email. The workflow model is outside the scope of this work, and thus further information has been excluded from this paper.

3. Evaluating the sMail Speech Act Model

In this section we first present our methodology for the evaluation of the sMail model, before proceeding to an elaborate examination of the results. The findings and observations will be taken into consideration in the next section where we improve the sMail model and consequently all the Conceptual Framework.

3.1 Methodology

In order to perform our evaluation we required an appropriate statistical methodology that serves two purposes:

1. To measure the sMail model's goodness of fit when applied to real data.
2. To compare this measure with that for the Carvalho&Cohen model.

This measure can be obtained by calculating the inter-annotator agreement between human annotators annotating segments of a corpus of emails with one or more speech acts. Of the available methodologies we chose the Kappa statistic, which may be computed as:

$$\kappa = \frac{P(A) - P(E)}{1 - P(E)}$$

where κ is the Kappa coefficient, P(A) is the total agreement, and P(E) is the percentage of agreement which occurs by chance alone. The value of Kappa ranges from -1 to +1, 1 being complete agreement. The use of this statistic in linguistics to measure inter-annotator agreement for linguistic annotation was proposed by (Carletta, 1996).
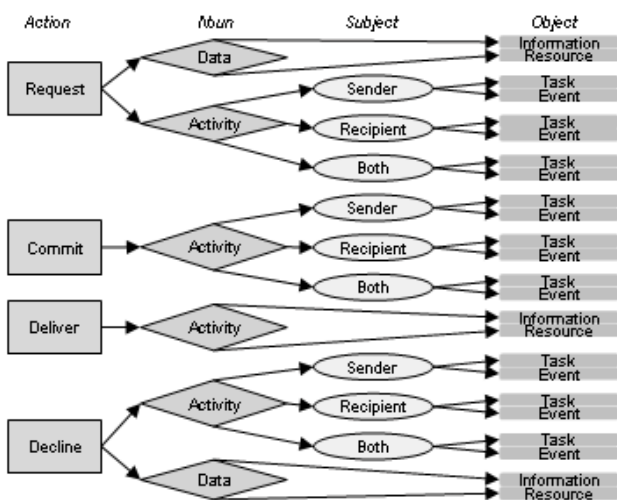

Figure 2 : Valid Speech Act Combinations in sMail Model

Despite that the same statistic was used to evaluate the Carvalho & Cohen model (henceforth referred to as the CC model), the statistic they obtained in their work could not be directly compared to the statistic which we were measuring, given that the email corpus they used is different. Therefore we instructed two annotators to carry out two separate annotation experiments using the same

corpus to calculate κ for both models. Therefore we instructed the annotators to carry out two separate annotation experiments using the same corpus to calculate κ for both models. Two subjects were selected to carry out the annotation tasks, one male and one female postgraduate student, with a computer science and linguistics backgrounds respectively. The corpus for the experiment was comprised of a random selection of 50 email threads from the Enron corpus[2] which discussed social, academic, and corporate issues. The random selection was subject to one constraint - the number of emails in a thread should not exceed 8. The number of emails in the resulting 50 threads totalled 174 (~3.5 emails per thread). Each individual email was between 1 and 500 words in length.

The annotation tasks for the sMail model required annotating multiple text segments within an email. These segments were not pre-agreed upon by the annotators. Where necessary, it was agreed to assign more than one speech act to a single segment. These choices did not compromise the relevance of the inter-annotator agreement, since if one annotator assigned one speech act to a sentence whereas the other assigns none or multiple speech acts, the extra annotations where considered as a disagreement. There were 24 valid speech act combinations in the sMail model (Fig.2) and 16 for the C&C model. These were used as categories for the κ statistic for the two experiments. The experiments were carried out in sequence and not in parallel, starting with the sMail model. The annotators were instructed to select text segments for annotation exclusively according to the model being evaluated, to avoid any undue influence of one model on the other. Given our emphasis on context retention, the annotation task was thread-oriented in order to facilitate the appropriate assignment of speech acts. Without context, a text segment may be assigned an entirely different speech act than the one the sender originally intended. For example, the text "*No problem!*" out of context, would probably be assigned the speech act (Deliver,Information,Ø). However given that the previous email in the thread contained the text "*Lets you and me try and talk today*", it is actually a (Commit,Event,Both).

| 1: "Here is C.'s resume." |
| 2: "We would appreciate any help you could give him." |
| 3: "He is available to come by and meet anyone you would think appropriate in the intern process." |
| 4: "Please advise us what we should do next." |
| 5: "He's here for experience but very interested in any prospects you might have at Enron." |

Segment	Annotator	SA Action	SA Object	SA Subject
1	A,B	Deliver	Resource	Ø
2	A,B	Commit	Task	Recipient
3	A,B	Deliver	Information	Ø
4	A	Request	Information	Ø
	B	Request	Task	Recipient
5	A,B	Deliver	Information	Ø

Table 1 : Email annotated according to sMail model

[2] http://www.cs.cmu.edu/~enron/

Table 1 shows an example of an email annotation (given the sMail model) after the greeting and signature where removed. In this example, both annotators broke down the email in five separate segments. They agreed on the assignment of speech acts in all cases, excluding the fourth, which shows a disagreement between a (Request,Information,Ø) and a (Request,Task,Recipient). This data was compiled for all emails in order to calculate the inter-annotator agreement and to get an insight into the main causes for disagreement.

3.2 Results and Observations

Given the text segments for annotation were selected independently by each annotator, the total number of annotations produced by each annotator differed slightly for both models (six in both cases). In cases where one annotator produced an annotation for a segment, while the other produced none, this was nevertheless considered an annotated text segment – annotated by one annotator with a speech act, and by nil with the second annotator. This is true for other text segment disagreements, e.g. two annotations for a sentence by one annotator as opposed to one by the second. Thus, the joint final number of unique annotations for the sMail model was 419, whereas that for the C&C model was 444. The difference between this number for the two models resulted since the selection of text segments to be annotated was based on the model being considered. The larger amount of annotations for the C&C model can be attributed to the fact that its Deliver-Opinion and Deliver-Data categories are equivalent to the single (Deliver,Information,Ø) speech act in the sMail model. Thus two sentences delivering some information plus an opinion were regarded as one segment in our model as opposed to two in the other model.

Model	Annotations	Agreement	Kappa
C&C (Full Model)	444	336	0.756
sMail (Full Model)	419	340	0.811
C&C (Deliver Merged)	444	369	0.830
sMail (Object Merged)	419	351	0.836
sMail (Subject Merged)	419	342	0.814
C&C (Relevant)	265	13	0.511
sMail (Relevant)	210	131	0.623

Table 2 : Results of experiments on the two models

We calculated the value of Kappa for each model with regards to the full number of categories, as well as with some categories merged or omitted, in order to be able to make more accurate observations. The κ for the full sMail model was calculated at 0.811 as opposed to 0.75 for the full C&C model. This compares well with their earlier inter-annotator agreement experiment (Carvalho & Cohen, 2005) that gave a value between 0.72 and 0.85, and is substantially lower than the one we achieved. To gather more insight into the causes for disagreement, and make better comparisons with the C&C model, we decided to calculate further κ's for both models. The κ where recomputed after the following considerations (the results are summarized in Table 2):

1. C&C (Deliver Merged): Given that the sMail model's (Deliver,Information,Ø) speech act is equivalent to the C&C Deliver-Data and Deliver-Opinion speech acts, comparing this value with the one we obtained might be considered unfair (since their model would have two categories equivalent to our single category). Thus we merged these categories for the C&C model. Although the resulting κ is now much closer to ours, we are satisfied with our result since even though we introduced a third parameter in our speech act definition to provide more structure to the knowledge representation, we virtually achieved the same inter-annotator agreement.

2. sMail (Object Merged): Disregarding the speech act object parameter for the sMail model – in order to examine the effect of this parameter on κ.

3. sMail (Subject Merged): Disregarding the speech act subject parameter for the sMail model – for the same reason. In this case κ improved over the initial κ by a mere 0.004 of a fraction, suggesting that the introduction of the speech act subject parameter was successful.

4. C&C (Relevant): All speech acts in the C&C model have a clear expected action for the email user (to reply, to schedule a meeting etc) - except for Deliver-Data and Deliver-Opinion, which simply require acknowledgment. These constitute 57.1% of the total category assignments for 57.2% of the total category assignments. After removing these categories, we obtained the relevant κ.

5. sMail (Relevant): Similarly, (Deliver,Information,Ø), which accounts for 57.2% of the total category assignments was disregard for our model, obtained a relevant κ which is significantly higher than the C&C relevant κ.

Following these results, we wanted to study the causes for disagreements in order to improve the sMail model. We were able to do this be creating and scrutinizing a confusion matrix (Fig.3) for the disagreements between speech act assignments. The figure highlights the most significant counts (Outlined: 5), disagreements attributed to the speech act subject (Dark Gray: 2), and disagreements attributed to the speech act object. These are classified into disagreement on the type of activity, i.e. task versus event (Gray: 4) and disagreement on the type of data, i.e. information versus resource (Shaded Gray: 7)
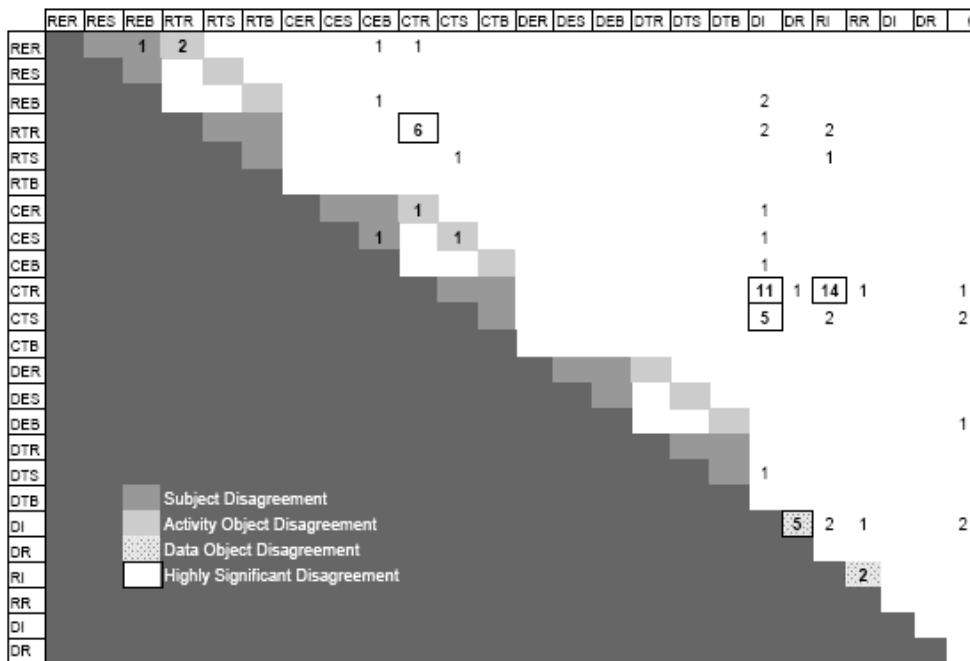
Figure 3 : Confusion matrix for annotation disagreements given the sMail model

Legend:
- Subject Disagreement
- Activity Object Disagreement
- Data Object Disagreement
- Highly Significant Disagreement

personal commitment by the sender.

After disambiguating some definitions with the annotators, we supported them with re-annotating the disagreements for the five highest ranking disagreeing category pairs (numbering 41). The resulting κ for these annotations was 0.609. Although this is a big improvement from the total disagreement in the first round, it shows that some ambiguities could not be resolved even upon discussion. The annotators brought forward to us some other issues that were not apparent in the results and these will also be taken into account for the improvement of our sMail model and consequently the sMail Ontology.

noun objects. Disagreement for one pair in the latter group, (Deliver,Information,Ø) and (Deliver,Resource,Ø), was significant with 5 counts (Light Gray and Outlined), and was attributed to an unclear definition for what constitutes a Resource - something accessed via an email but external to email, e.g. an attached file or a URL. (Commit,Task,Recipient) and (Request,Information,Ø) tops the list of pairs in high disagreement with 14 counts. The rough distinction between these two is that the first expects the recipient to perform something, whereas the second expects the recipient to send information back to the sender. However there are overlapping cases where both apply e.g., "*Please go through the list and determine if you want the following information*". The second highest disagreeing pair, (Commit,Task,Recipient) versus (Deliver,Information,Ø), with 11 counts, is attributed to conditional statements e.g. "*If we can hold off until next week to set the new offices up I'm sure that D. and D. would be very grateful!*" and non-binding statements e.g. "*I recommend that you visit with M.T. on this.*" In both cases, the commisive force is too weak for a (Commit,Task,Recipient) and resulted in one annotator considering it as a (Deliver,Information,Ø). (Commit,Task,Recipient) has also the third most significant disagreement, at 6, with (Request,Task,Recipient) and this is also attributed to cases where both speech acts apply, e.g. "*Please look at the attachment and inform me if you believe we may be perceived in an advisory style role*". The fourth highest disagreeing pair is also related to the Commit verb, making the Commit verb the most ambiguous in our ontology. There are 5 disagreements for (Commit,Task,Sender) versus (Deliver,Information,Ø). Most are also attributed to conditional or non-binding statements, e.g. "*I don't mind helping out - let me know if you need a hand.*". This example falls short of being a

4. Improving the sMail Model

After considering the evaluation results from the previous section, we decided on some modifications to the sMail model (Fig.4). The most important change was to include the verb Suggest, to cater for all weak-commisive, conditional and non-binding statements, which were the cause of a large proportion of the disagreements.
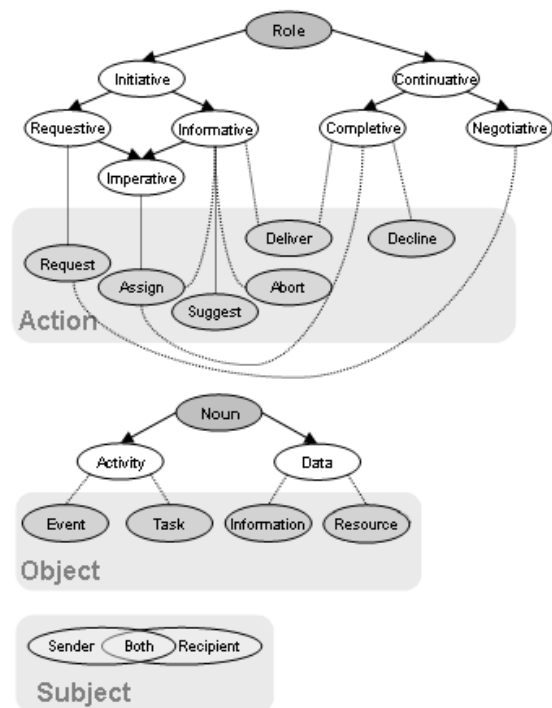

Figure 4 : The improved sMail Speech Act Model

Action	Noun	Subject	Description	Role	IEA	PER
Request	Activity	Recipient	Request/Amend recipients' activity performance	Requestive/Negotiative	Expect	Reply
		Sender	Request/Amend performance of personal activity	Requestive/Negotiative	Expect	Reply
		Both	Request/Amend performing joint activity	Requestive/Negotiative	Expect	Reply
Assign	Activity	Recipient	Permit/Assign recipient's activity	Completive/Imperative	None	Perform
		Sender	Approve/Announce personal activity	Completive/Informative	Perform	None
		Both	Approve/Assign joint activity	Completive/Imperative	Perform	Perform
Suggest	Activity	Recipient	Suggest an activity for indirect consideration	Informative	None	None
		Sender	Suggest a personal activity for indirect consideration	Informative	None	None
		Both	Suggest a joint activity for indirect consideration	Informative	None	None
Decline	Activity	Recipient	Decline permission for recipients activity	Completive	None	None
		Sender	Decline performing personal activity	Completive	None	None
		Both	Decline performing a joint activity	Completive	None	None
Abort	Activity	Recipient	Abort recipient's preset activity	Informative	None	Perform
		Sender	Abort own preset activity	Informative	Perform	None
		Both	Abort preset joint activity	Informative	Perform	Perform
Request	Data		Request data from recipient	Requestive	Expect	Reply
Deliver	Data		Deliver data	Completive/Informative	None	None
Decline	Data		Decline delivering data	Completive	None	None

Table 3 : The adjusted sMail Speech Act Process Model

"I recommend that you visit with M.T. on this." can now be treated as a (Suggest,Task,Recipient), *"ok. I don't mind helping out - let me know if you need a hand."* as a (Suggest,Task,Sender) and *"If we can hold off until next week to set the new offices up I'm sure that D. and D. would be very grateful!"* as a (Suggest,Task,Both). To cater for negotiative requests, we introduced a Negotiative role alongside the Completive role under the Continuative role. The Request verb can now double as the amend verb, which was not physically included because its predictive behaviour is exactly like that for Request. The third addition was the verb Abort. Although occurrence for this verb was very low, we believe that it should be included since we would like a smart email client to also be able to support the user with cancelling pre-established activities. Other superficial changes including renaming the *Commit* action into *Assign* – this proved more intuitive to the human annotators.

After fine-tuning the sMail model, we wanted to immediately evaluate any added benefits on real data. The annotators where again instructed to reconsider the disagreements for the five highest-disagreeing categories whilst keeping the new model in mind. The main result of this experiment is a higher kappa value of 0.732, significantly higher then the 0.609 achieved when we assisted the annotators with their annotation.

The changes in the sMail model have been reflected in the Speech Act Process Model and consequently in the sMail Ontology[1] which represents the knowledge in both models. Very briefly we will now re-introduce the adjusted sMail Email Speech Act Process Model. In essence, it outlines the expected reaction from both initiator and participant of a speech act, on sending it and on receiving it respectively. It assigns the *Initiator Expected Action* [IEA] and the *Participant Expected Reaction* [PER] to each speech act combination in the sMail model, and is applied over it as:

$$(a,o,s) \; [IEA] \rightarrow [PER]$$

where IEA refers to the status or action of the speaker, or in this case, the initiator on sending a speech act (*Expect* or *None*); and the PER refers to the reaction expected from the hearer, or in this case the participant upon receiving and acknowledging a speech act (*Reply*, *Perform* or *None*). *Expect* denotes that the communicator is expecting further communication. *Reply* denotes that the communicator is expected to further the communication, whereas *Perform* denotes that on sending or receiving a speech act the communicator is expected to perform an external action as a direct result of the speech act (e.g. attend an event/perform a task). Table 3 shows all combinations (Action, Noun – i.e. generalization of the Object, and Subject) of speech acts in the sMail speech act model with a brief description, the discourse roles which each can exhibit alongside the associated IEA and PER.

5. Conclusions and Future Directions

Our evaluation supports the belief that our sMail Email Speech Act Model constitutes an improvement over previous models, taxonomies and ontologies tackling the representation of speech acts in the email conversation domain. The modifications in this model have been reflected in the complete sMail Conceptual Framework, including the sMail Ontology. In the future, we will focus and elaborate on the third model of the framework - the sMail Email Speech Act Workflow. Through this workflow, we will be demonstrating how the two models presented in this paper – the Speech Act Model and the Speech Act Process Model can be used in practice to model ad-hoc workflows occurring in email conversation, and how this will impact the email user.

After the results of this evaluation, we started investigating text analytics and Ontology-based Information Extraction techniques that can be employed within the extended email clients to support the user with semi-automatic content metadata extraction, where speech acts occurring in the text are anchored to instances of the sMail model in the sMail ontology.

We have begun to use the GATE (Cunningham et al, 2002) framework to engineer and test speech act extraction algorithms. Initially we implemented a simple rule based approach, whereby speech acts were extracted based on JAPE[3] (Cunningham et al, 2000) pattern action rules. The majority of the speech identification will be done here whereby JAPE patterns benefit from previous linguistic

---

[3] Java Annotations Pattern Engine

annotations to perform speech act annotation at the sentential level. The work is inspired by Khosravi & Wilks (1999) which implemented similar rule based approaches using earlier versions of GATE. Whereas the aforementioned Knowledge Based (KB) approaches to Information Extraction (IE) can be time consuming; Machine Learning (ML) approaches require a large amount of training data, which is not readily available for speech acts. On the other hand, ML systems have high precision and can maintain high recall depending on the volume of training data available, whilst a KB system maintains high precision at the expense of low recall. IE literature does not attempt to restrict the developer to either approach, but it emphasizes the appropriate conditions needed for one or the other, or the possibilities of employing both (Apellt & Israel, 1999). We intend to investigate the third option i.e. a set of hand-coded JAPE rules may serve as input for the training algorithm or as a backup strategy where a trained system fails.

Furthermore, we are undertaking the actual implementation of light-weight extensions to popular email clients[4] to handle the creation and interpretation of speech acts within semantic email. This is supported by the text analytics technologies just described, and algorithms based on the sMail Conceptual Framework models. Through these extensions we want to semantically enhance email messages with invisible semantic annotations (via a specific MIME extension allowing for an RDF content-type in the email headers) that encompasses the acquired message's content metadata as well as its contextual (pertaining to threads) metadata. Once fully implemented, email users will be supported by smarter email clients that predict their actions on the basis of the semantics accompanying text in email messages. Rather than going through unread emails, the user will be able to check, or even be reminded of, speech acts that still require action. The sender's expectations will be clear to the recipient on reading the email and the smart email client will be able to aid the user by easing the workflow of email-generated personal information management.

6. References

Apellt, D., Israel, D., J. (1999) Tutorial on information extraction, *Applied Natural Language Processing*, Washington, D.C, Stockholm, Sweden.

Carletta, J. (1996) Assessing Agreement on Classification Tasks: The Kappa Statistic. *Computational Linguistics*, 22(2), pp. 249--254.

Carvalho, V., Cohen, W. (2006) Improving Email Speech Act Analysis via N-gram Selection. *In Proceedings of the Human Language Technology Conference / North American chapter of the Association for Computational Linguistics,* ACTS Workshop, New York City, NY.

Cunningham, H., Maynard, D., Bontcheva, K., Tablan, V. (2002) GATE: A Framework and Graphical Development Environment for Robust NLP Tools and Applications. *In Proceedings of the 40th Anniversary Meeting of the Association for Computational Linguistics,* ACL'02, Philadelphia.

Cunningham, H., Maynard, D., Tablan, K. (2000) JAPE: a Java Annotation Patterns Engine (Second Edition). *Technical Report CS--00--10*, University of Sheffield, Department of Computer Science.

Erikson, T. (2000) Making Sense of Computer-Mediated Communication (CMC): Conversations as Genres. *In Proceedings of the Hawaiian International Conference on System Services 2000*, HICSS2000.

Goldstein, J., Sabin, R.E. (2006) Using Speech Acts to Categorize Email and Identify Email Genres. *In Proceedings of System Sciences 2006*.

Khosravi, H., Wilks, Y. (1999) Routing email automatically by purpose not topic. *Natural Language Engineering,* 5, pp. 237--250.

McDowell, L., Etzioni, O., Gribble, S., Halevey, A., Levy, H., Pentney, W., Verma, D., Vlasseva, S. (2003) Evolving the Semantic Web with Mangrove. *UW Technical Report*.

McDowell, L., Etzioni, O., Halevey, A., Levy, H. (2004) Semantic Email. *World Wide Web 2004*.

Scerri, S., Davis, B., Handschuh, S. (2007) Improving Email Conversation Efficiency through Semantically Enhanced Email. *In Proceedings of the WebS Workshop*. 18th International Conference on Database and Expert Systems Applications, Regensburg, Germany.

Searle, J. (1969) Speech Acts. *Cambridge University Press*.

Whittaker, S., Bellotti, V., Gwizdka, J. (2007) Email and PIM: Problems and Possibilities. *Communications of the ACM*.

Whittaker, S., Sidner, C. (1996) Email overload: exploring personal information management of email. *In Proceedings of CHI '96 Conference*, Vancouver, BC Canada, April, pp. 276--283.

---

[4] Although we are currently considering only Microsoft Outlook and Mozilla Thunderbird we might target other popular clients in the future.