Aiding the Workflow of Email Conversations by Enhancing Email with Semantics

Simon Scerri[1]

[1] Digital Enterprise Research Institute,
National University of Ireland Galway
IDA Business Park, Galway, Ireland.
{Simon.Scerri}@deri.org

Abstract. Despite persisting in popularity, email is still plagued with information overload, hindering the workflow of data handled by the user. Just as Semantic Web technologies promise to revolutionize the Web, we aspire to use the same technology to enhance electronic mail with useful semantics. Thus 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, which can be attributed to the lack of explicit semantics covering context and content of exchanged 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 email speech act ontology and outline a non-deterministic model that predicts the user's best course of action upon sending or receiving an email.

Keywords: Email, Speech Act Theory, Metadata Extraction, Semantic Web.

1 Introduction

Email persists as on of the top features of the internet. Studies on lexical densities of email discourse showed [1] that despite being a written form of communication, email texts are closer to spoken rather than written discourse. Email has been regarded as a new genre [2], where a genre is 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 [3]. Email workflow is very inefficient because it lacks these shared expectations on how and when the exchanged information is to be acted upon. Processing incoming messages is frequently postponed, sometimes indefinitely, due to different priorities [4] or because the mental effort required would lead to distraction from other tasks. Whereas it should be the email sender's interest to make any expectations explicit to the recipient, the latter frequently ends up having to invest more time to extract and act upon implicit expectations. Apart from being subject to misinterpretation, this process puts off the recipient from immediately trying to act upon a message. In a nutshell, the lesser the effort required out of the recipient, the greater the chance that the sender's expectations are fulfilled in a timely manner.

2 Background and Related Work

Speech Act Theory [5] 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 would be the Locutionary meaning. The theory was applied to Email a number of times, in particular for email classification based on the sender's intent [1][2][6], focus detection of threaded email conversations [7], predicting actions on email messages [8] and easing task management arising through email [9] amongst others. Although these provided promising results, they had a serious limitation since the expectations accompanying messages were only guessed on arrival, and thus never confirmed by the sender. An email message is frequently multi-purpose, realizing several purposes at the same time. Therefore our approach goes beyond simple email classification, since we consider specific segments within an email and not the email as a whole. Other relevant research work involved the introduction and formalization of Semantic Email processes [10]. Based on the Semantic Web paradigm 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.

3 Semantically Enhanced Email

Although email has many weaknesses, it also provides a fundamentally right model for a communication system [11]; the major advantages of the model being asynchronosity, threading and the fact that it is a command central system. Therefore we would like to retain the basic email model, but extend its functionalities by adding a semantic layer to the model. In particular, this will clearly state the otherwise implicit intents and expectations associated with speech acts in a message. We believe that by making this information explicit, the user is aided with the exchange of information. As a result email's disconnected workflow becomes more efficient. By:

- Fine-tuning existing email speech act taxonomies presented in earlier work [6] and creating our own email speech act ontology;
- Outlining a predictive model for illocutionary and perlocutionary reactions attributed to speech acts in email messages;
- Applying the results within extensions to popular email clients capable of capturing and embedding semantic information in exchanged messages ;

we aspire to achieve this scenario and thus substantially reduce the occurrence and consequences of the given problems. In this paper, we are mainly concerned with the first two steps and we will elaborate on our ideas in the coming sections.

3.1 Email Speech Act Ontology

Our ontology is a refinement and extension to an existing taxonomy [6], which regarded speech acts as conjunctions of various *Verbs* and *Nouns* as a pair (*v-n*). By including further parameters in our speech act model, we believe that our ontology is much more powerful than any of its predecessors. In particular, we directly addressed our main concerns: the intents and expectations accompanying speech acts – by including specific parameters in the model. 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 are then 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 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.
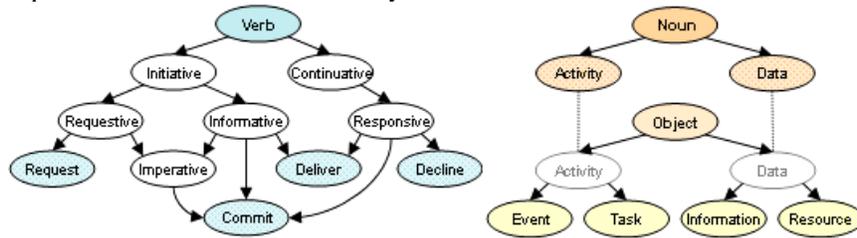

Fig. 1. Speech Act Verb, Noun and Object

In our ontology, we categorize the nouns in two major concepts: *Data*, 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 *Information* and *Resource* are Data 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 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. We therefore 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 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.

3.2 Predicting Reactions on a Speech Act

Table 1. Realized Speech Acts Combinations and Expected Reactions

Verb-Noun	Object	Subject	Description	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	Resp/Informative Resp/Imperative Resp/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-Data	Info/Reso		Request data from recipient	Requestive	Expect	Reply
Deliver-Data	Info/Reso		Deliver data	Resp/Informative	None	None
Decline-Data	Info/Reso		Decline delivering data	Responsive	None	None

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 *Expect*, where the sender expects a response on sending a speech act; and *None*, where the sender or recipient is expected to do nothing on issuing or receiving the speech act. Active reactions are *Reply*, when the recipient is expected to reply on getting a speech act; and *Perform*, for speech acts which demand an Activity, e.g. Task, from the sender or recipient on sending or getting a speech act. We apply this predictive model to our realized speech act definition as (v-(o)[s]) {ERs}$\rightarrow${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. Table 1 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.

4 Future Directions and Conclusion

We are currently evaluating how well our speech act model fits a real corpus of threaded email messages, and how it compares to previous work in [6], by using the Kappa statistic to measure human annotator agreement for both models. After considering the results, we plan to extend popular Email Clients to enable semantic email by providing: semi-automatic content metadata extraction through text analytics; context metadata retention through threaded-based email handling; and invisible semantic annotation of email (based on our ontology) with such metadata through a MIME extension that allows for an RDF content-type in the email headers.

We believe that the presented models can be a sound basis for achieving our goal: improving the data workflow efficiency for the user. Although they are generic enough to be applied to other communication media, the problem addressed here mostly concerns disconnected workflows – where implicit expectations have a larger impact on workflow efficiency. We want to achieve a scenario where email users are aided by smarter email clients that predict their actions on the basis of the semantics accompanying speech acts in email. The semantic email-aware email client will aid the user by autonomously aiding the workflow of personal information generated by email. The client will suggest the most appropriate action for speech acts the user is creating or acting upon. Rather than going through unread mails, a user will be able to periodically check or even be reminded of speech acts they were expected to act upon and never did. If supported by personal information management tools, the email client might suggest saving a task in a task list once the user commits to it. This scenario would improve the overall efficiency of email conversations.

References

1. Khosravi, H., Wilks, Y.: Routing email automatically by purpose not topic. Natural Language Engineering, Vol. 5 (1999) 237–250.
2. Goldstein, J., Sabin, R.E.: Using Speech Acts to Categorize Email and Identify Email Genres. System Sciences, HICSS2006 (2006).
3. Erikson, T.: Making Sense of Computer-Mediated Communication (CMC): Conversations as Genres. Hawaiian International Conference on System Services, HICSS2000 (2000).
4. Whittaker, S., Bellotti, V., Gwizdka, J.: Email and PIM: Problems and Possibilities. CACM.7 (2007).
5. Searle, J.: Speech Acts. Cambridge University Press (1969).
6. Carvalho, V., Cohen, W.: Improving Email Speech Act Analysis via N-gram Selection. HLT/NAACL, ACTS Workshop, New York City, NY (2006).
7. 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.
8. 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.
9. Khoussainov, R., Kushmerick, N.: Email task management: An iterative relational learning approach. Conference on Email and Anti-Spam, CEAS'2005 (2005).
10. McDowell, L., Etzioni, O., Halevey, A., Levy, H.:Semantic email. In Proceedings of the 13th World Wide Web Conference, ACM Press, New York (2004) 244-254
11. Armstrong, E.: What's wrong with email? Bootstrap Alliance (2000)