metalogue

Metalogue Progress Report

PUBLISHABLE SUMMARY

Contract number :	611073
Project acronym :	Metalogue
Project title :	Multiperspective Multimodal Dialogue: dialogue system with metacognitive abilities
Workpackage, Deliverables and Tasks	Progress Report Report. Publishable summary
Duration	Months 13 - 24

Organisation Name :	DFKI
Author:	Jan Alexandersson, Andrey Girenko
Delivery date :	15.11.2015

SEVENTH FRAMEWORK PROGRAMME

Coordinator: DFKI GmbH

METALOGUE : dialogue system with metacognitive abilities

We live in the world where multimodal natural language-based dialogue is increasingly becoming the most attractive human-machine interface and it finds applications staring from information offices, cars to smart houses and smart working environments. Such interfaces offer a mode of interaction that has certain similarities with natural human communication by using a number of input and output modalities which people normally employ in communication, e.g. speech, gesture, facial expressions, pointing devices, etc. Some of these interfaces will advance to the incorporation of multimodality into virtual and augmented reality environments, e.g. using embodied conversational agents.

At the same time, existing dialogue systems, by common agreement, do not yet show interactive behaviour that is natural to its human users, and do not have the flexibility to exploit the full potential of spoken and multimodal interaction. In many instances people refuse to use available multimodal language-based interfaces because they are perceived as being too artificial and inconvenient. There are several reason why people can communicate effectively and efficiently and computer systems can not is because, e.g.

- computer dialogue systems do not have the rich experience and background knowledge that humans have
- humans are able to process and perform several actions – both task-related and communicative ones – simultaneously whereas dialogue systems typically are not. Moreover, if this happens, it mostly happens by accident rather than by design
- other than computer-based systems, humans have metacognitive abilities, that is, they are able to monitor, assess and reason about their own and their partner's performance

The goal

The overall goal of Metalogue is to develop and evaluate a multimodal interactive system that is able to implement a behaviour natural to human users. The Metalogue system should be flexible enough to exploit the full potential of multimodal interaction as well as to evaluate the developed technologies in multiperspective educational dialogue setting.

The key to achieving the Metalogue's vision of a future dialogue system is equipping the system with metacognitive capabilities. The Metalogue conversational agent provided with sufficient metacognitive skills will be able to:

1. adapt its dialogue behaviour over time according to the interlocutor's knowledge, attitude, and competence, and
2. predict other people's knowledge and intentions and show proactive dialogue behaviour,

thus, being more "human" than any know artificial systems.

In order to enable deeper understanding of metacognitive processes and the nature of the acqisition of such skills, the Metalogue system will have shared and varied responsibilities of observing, monitoring, experiencing and executing different tasks, by presenting similar materials in multiple contexts enabling self-reflection, by becoming aware of different strategies and how they work.

Education and training as applications

The main Metalogue application area will be education and tutoring with a multi-perspective support. The reasons to choose this application area include the following:

- It was proven that having better metacognitive skills can help learners learn better and self-regulate their learning across domains and contexts.
- The educational dialogue and tutoring interventions provide useful constraints and a dialogue framework. Simply saying, it is feasible to attain the ambitious goal of creating a dialogue system, which looks "human", in such setting.
- Educational dialogue systems have potentially a high economic and social impact, given the high recognition of lifelong learning as fundamental to Europe's long-term success.

It has been scientifically proven that metacognition is critically important in negotiation. This is because it significantly influences decision-making processes. Therefore, Metalogue devotes attention to negotiation tasks. Thus, Metalogue aims for two main types of negotiation tasks and for two main types of target users:

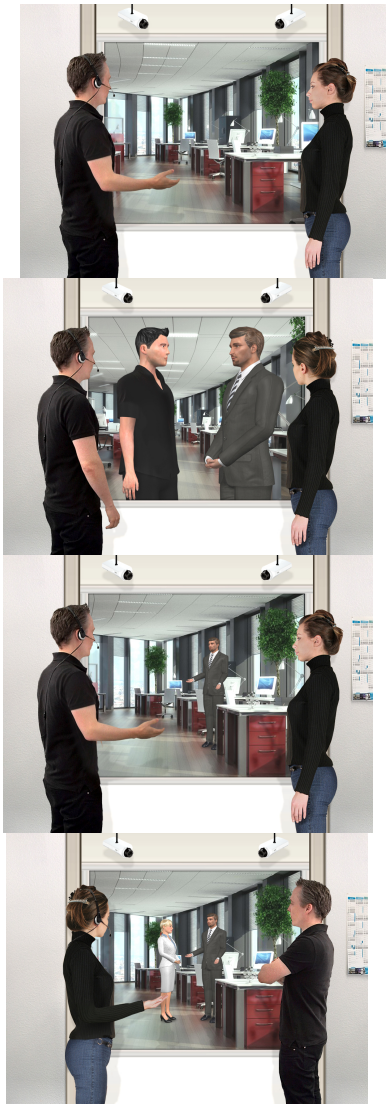
- **Young entrepreneurs and policy debating**: The Metalogue system will be used to train young entrepreneurs in debating over policy issues. Use cases will be related to educational scenarios, having as users young trainees related to the Youth Parliament & the Hellenic Parliament Foundation's educational activities along with their teachers. The project will observe and improve the metacognitive abilities of the trainees, creating societal abilities and skills of the new generation of citizens, introducing them into the modern world issues, including rules, obligations, rights, social behaviour and responsibility.
- **Call centre employees and customer negotiations**: The Metalogue dialogue system will be used to train call centre employees in customer negotiations with a focus on governmental service provision. It is designed to deliver a realistic training experience and to make it possible to give quantitative evaluations of how well a given call went. The customer service domain allows room for negotiation and is a prudent choice to test the Metalogue system.

How it will work

Metacognitive skills for where both the system and the user will be of importance in different settings, where they will play different roles in the training environment, for example:

1. The system passively observes tutoring session between a user and a tutor. We refer to this as "observing mode" since the system keeps track of human-human dialogue without interfering in it. The complete discourse is recorded for further analysis.

2. The system observing "scripted" tutoring sessions thereby has possibility to intervene, asking relevant questions and influencing the tutoring process by instruct the user to change his behaviour/performance.



3. The system re-plays a user's performance in real time so that the user can observe his/her own performance and the user has the opportunity to discuss it with the tutor. The system will additionally allow for re-entering the tutoring session at any point, thus improving the performance as it went wrong.

4. The system actively plays the role of one of the dialogue participants. Human tutor observes, evaluates and guides.

5. The system acts as a tutor thereby guiding multiple users.

The Metalogue system is designed to make use of a wide array of different modalities, such as spoken natural language, facial expressions, body posture and bio-sensory data. Where appropriate, modalities are designed to be symmetric: a modality available on the input side for user input will also have a counterpart on the output side. That means that the ability to use facial recognition will be complemented by the generation of facial expressions on the virtual characters. This allows for a more natural interaction and is especially useful in multi-party interactions. In such settings, some communication channels may be exclusive to two communication partners, e.g. speech, but other participants can still interact in parallel, e.g., using gestures.

To make the process, in which the Metalogue system is applied, transparent the project has defined so-called user journeys, e.g. **Figure 1**. There are three possible journeys, namely "tutor guided", "Learner Initial Self Assessment" and "Wholly Learner Self Directed". **Figure 1** depicts the former journey by which the learner's journey is conducted by a tutor. The other two journeys serve the purpose of less tutor involvement where the latter is a completely tutor-free journey.

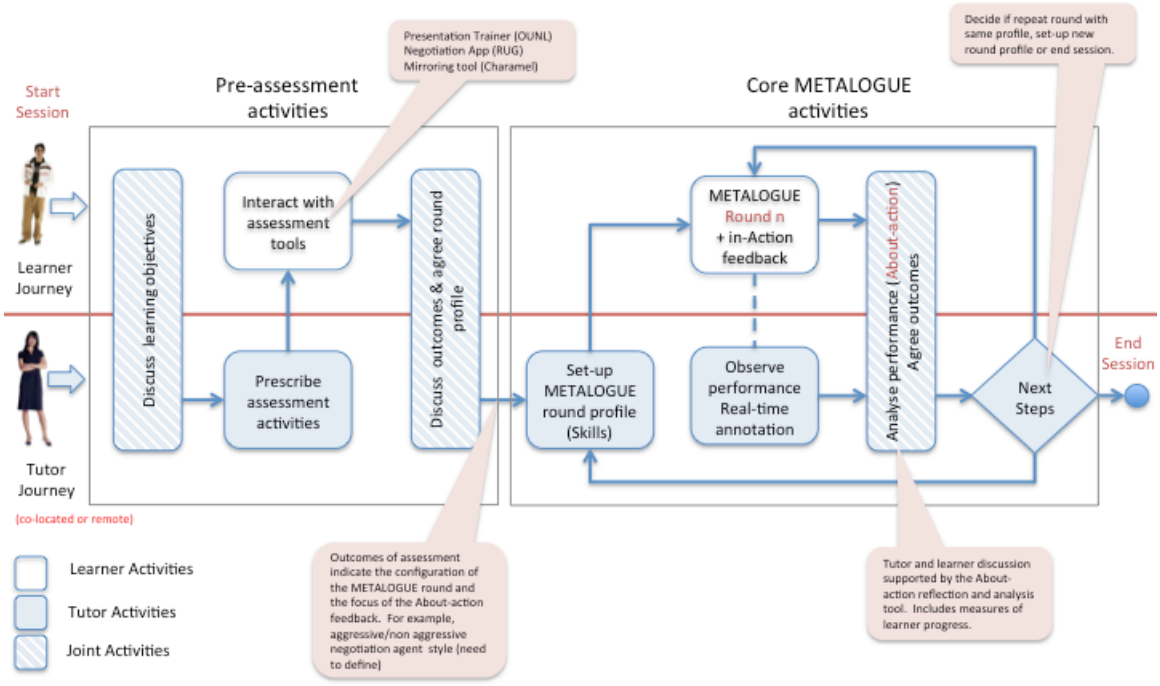


Figure 1. A tutor guided "user journey". The top half of the figure represents the learner actions, the lower that of the tutor. Initially, the learner and tutor interacts with the purpose of defining the learner's objectives. It follows prescription and interaction with pre-assessment tools after which the learner and tutor agree on the learner's tutoring objectives. It follows a sequence of sessions with the Metalogue system including set-up, training and performance analysis. At the end of each round, a joint decision on continue/exit is taken based on the system's and tutor's assessment of the learner's performance. The system deploys two novel feedback mechanisms: "in-action feedback" providing real-time feedback during the interaction and "about-action feedback", a dashboard containing the complete interaction decorated with system-detected hotspots of the learners inappropriate behaviour.

Expected impact

1. Metalogue constitutes a significant step in deeper understanding of metacognitive processes and nature of the acqisition of metacognitive skills. Having acquired this, it will be easier to design computer systems behave similar to humans. Thus, Metalogue will contribute to Europe's scientific leadership in this field.
2. Metalogue will push forward the technological foundations for designing multimodal dialogue systems employing a spectrum of communication modalities making human-computer interaction more effective and natural.
3. By developing and evaluating educational environment for debate and negotiation training of young European citizens Metalogue will contribute towards such policies as "ET 2020" – a European policy as the Strategic framework for European cooperation in education and training.
4. Advances developed by Metalogue will become available to European high-tech industry, first of all SMEs, through open-sourcing of major results, therefore inspiring innovation and commercialization of research results.

Progress achieved after two years

- Several cognitive models have been developed allowing the modelling of metacognitive processes of humans in the negotiation setting. These models are not only used to extend the Metalogue prototype system with metacognitive capabilities, they have been shown

to help in developing metacognitive capabilities and improve strategic adaption capabilities of humans. Finally, some of these models passed the "Turing Test"! Furthermore, the project has extended the initially planned interaction scenario debate to that of negotiations and adapted its technology and models accordingly.

- The project has developed, tested and piloted the methodology of multimodal corpus development including data collection, synchronisation, annotation and processing. The complete Metalogue corpus will be released to the scientific community and consists of more than 10 hours of mono- and multimodal interaction, which has been annotated with ISO 24617-2 Dialogue Acts, emerging standards on semantic relations in discourse. The corpus is utilised to design and train the Metalogue system and its models, it divides into four parts:
    - 15 smoking ban debates from the Hellenic Youth Parliament
    - 24 dialogues of single-issue negotiation (Game of Nines) and 26 multi-issue negotiations (smoking ban) from the University of Groningen
    - 3 UK Youth Parliament sessions from 35 speakers on three topics: sexual relationship education; tuition fees; job opportunities for young persons
    - 5 Vodafone UK Call Centre Agent web chats on mobile phone deal negotiation
- The first and second Metalogue system pilots have been completed. The former – pilot 1 – corresponds to the "observing mode" scenario, the latter to the user-as-experiencer mode, which will be evaluated within the scope of the second pilot. The first pilot was deployed in June 2015 and evaluated, evaluation of the second pilot is scheduled for Q1, 2016. The evaluation of pilot 1 was performed within the Hellenic Youth Parliament and involved both student learners and Youth Parliament Foundation experts as tutors.
- On the technical side, the project has developed a flexible integration platform based on ZMQ allowing the easy integration of both legacy as well as novel modules. The project plans to release this to the scientific community. During the design and implementation, the project has developed novel NLP algorithms, e.g. a segmenting Viterbi algorithm based on a conditional signal model. The project has adapted a Kaldi-based ASR with respect to the Metalogue corpus achieving WER=30%. Perhaps the most outstanding aspect of the Metalogue system is the combination of a multi-thread architecture dialogue manager and the integration of a metacognitive module with multiple strategy control capabilities.
- The projects dissemination activities include so far more than 50 project presentations, social media activities and 20 peer-reviewed publications at conferences and journals. Highlights include two ECTEL awards of project demonstrations: the jury selected best-demonstration award in 2014 and "peoples' choice award" in 2015. The project has produced a Call-Centre Demo promo video for industrial uptake activities. The project is about to extend the ISO 24617-2 standard on Dialogue Acts by incorporating negotiation moves and possibly adding semantic content drawn from sibling standards, such as ISO 24617-{1,5 and 8} on events, semantic roles and semantic relations respectively. The project developed an XSD-Scheme and a conversion tool for the ISO DiAML which have been donated to ISO.

To conclude, given its complexity and constraints, the project has demonstrated depth and breadth in its activities by meeting challenges ranging from detailed technical development and integration, to elaborating the user context, through to considering the commercial viability of the end product. The Metalogue project is about to show that multimodal dialogue system with metacognitive capabilities enables a new level of flexibility when it comes to adapting its strategy to the situation superior to previous systems.