CoCoMaps Project

CMLabs | IIIM

CoCoMaps Final Report

Project Start Date: 1 September 2016
Project End Date: 31 March 2018

INTRODUCTION

The CoCoMaps Project is a joint effort by CMLabs (UK) and the Icelandic Institute for Intelligent Machines (Iceland) to integrate an architecture for robot control and interaction with humans using natural communication.

This report provides an overview of the CoCoMaps project as a whole including a summary of the tasks, deliverables, milestones and KPIs achieved.

The goal of CoCoMaps project is to demonstrate the advances made throughout this project in the development of the Collaborative Cognitive Map Architecture, including communication with multiple humans, information extraction, and more.

This report is listed as Milestone 8 in the CoCoMaps and includes the following deliverable:

- **T15.D1** Demos, results and literature publicly available

This report is split up to provide final reports on the following elements of the project, as presented in the original proposal and subsequently updated in collaboration with the project reviewers:

- Project Tasks
- Project Milestones
- Project Deliverables
- Project Impact KPIs
- Project Technical KPIs

and concludes with a summary of the results and a section on future work.

Finally, the project has published all reports, videos and results publicly at

https://cmlabs.com/cocomaps/publications

PROJECT TASKS

The following section describes the tasks for the CoCoMaps project.

T1: Specification of experimental platform

Partner Effort	Budget	Actual
CMLabs	1	1
IIIM	1	0.5
Total	2	1.5

The project started with the task of writing a specification for the experimental platform. This was completed by Month 3 as Deliverable T1.D1: Specification of Experimental Platform, which provides more technical details on this task.

T2: Integrate Psyclone support for Turtlebot

Partner Effort	Budget	Actual
CMLabs	4	4.5
IIIM	1	1
Total	5	5.5

This task was changed from the original task of integrating Psyclone for the Qbo robot (no longer available) to the revised task of integrating Psyclone to run on the new robot, the Turtlebot2. This made some aspects of the project easier, as the processing power of the new platform was higher and had better access to the individual, but the visual capabilities of the new platform were significantly reduced and the system as a whole provided less integration than originally anticipated. This resulted in more work having to be put into the integration and data analysis, but we believe that we ended up with a more powerful and capable platform for demonstrations of the project.

T3: Modules for visual detection and localisation

Partner Effort	Budget	Actual
CMLabs	4	5
IIIM	2	2
Total	6	7

Due to the change of robot platform and the lack of visual capabilities of the new robot this task was changed to focus more on vision for the purpose of natural dialogue, rather than searching for objects in the scene. The new task included integration with a more powerful face recognition engine and more advanced temporal analysis of facial data to provide clues for estimating the emotional state of the humans in the scene, the gaze direction (for conversational use) and the estimation of which human in the scene is the current speaker, if any.

T4: Modules for robot speech recognition and synthesis

Partner Effort	Budget	Actual
CMLabs	2	1.5
IIIM	0.5	1
Total	2.5	2.5

The new robot platform did not provide us with any speech tools, which required us to find our own solutions. After evaluating a number of third-party SDKs we settled on real-time speech SDKs from Nuance for both speech recognition and synthesis. We were able to successfully integrate these to run inside the Psyclone platform on each of the robots and get them to work collaboratively with the Turn-Taking and pitch tracking systems, as well as the visual speaker estimation.

T5: Port current state-of-the-art Turn-Taking system

Partner Effort	Budget	Actual
CMLabs	1	1.5
IIIM	0.5	0.5
Total	2	1.5

The original Turn-Taking system was ported to run on the new Psyclone platform and to run successfully and natively on each of the robots. It is now running as a combination of low-level modules created in C++ providing the audio data analysis and higher-level modules written in Python providing the intelligence and Turn-Taking algorithms themselves.

T6: Port prior work on Cognitive Map

Partner Effort	Budget	Actual
CMLabs	3	3
IIIM	1	1
Total	4	4

The original Cognitive Map Architecture was created to run on Honda's ASIMO robot and provided this robot with the ability to observe and learn from its interaction with its environment and one human. This work was ported to run on the robots within the new Psyclone framework to provide both the original functionality as well as foundations for the new work to be done to make the architecture collaborative. This was completed by Month 8 as Deliverable T6.D1: Current state-of-the-art implementation, which provides more technical details on this task.

T7: Task Planner

Partner Effort	Budget	Actual
CMLabs	2	1.5
IIIM	2	4
Total	4	5

The task planner started out as a monolithic structure and was then developed to run independently on each robot while providing the collaborating, negotiating and synchronising roles and tasks via the CCMCatalog. It eventually became part of the TDM – the Task Dialog Manager which now runs on each robot and orchestrates all tasks such as dialogue, search patterns, motor control, sensory feedback and task execution.

T8: Collaborative Cognitive Map

Partner Effort	Budget	Actual
CMLabs	6	6.5
IIIM	3	3
Total	9	9.5

The new architecture for the Collaborative Cognitive Map was split into two phases. The first phase provided a draft implementation which was used in Demo-1 and was completed by Month 14 as Deliverable T8.D2: Draft Collaborative Cognitive Map. This provided the robots with the ability to coordinate and collaborate on the task of searching for humans in the scene, as well as the initial parts of allowing the robots to discuss and share observations they have made via the CCMCatalog.
The second phase then added the ability to negotiate disagreements about these observations and the concepts of tasks and roles, which can also be assigned and negotiated via the CCMCatalog. This was completed by Month 17 as Deliverable T8.D3: Final Collaborative Cognitive Map, which provides more technical details on this task.

T8a: Specification and implementation of Hannover Messe Demo

Partner Effort	Budget	Actual
CMLabs	0	0.5
IIIM	0	0
Total	0	0.5

Due to the initial delays in funding confirmation and the adverse consequences to the project it was decided that working on a live demonstration for the ECHORD stand at the Hannover Messe 2017 would provide the project with an alternate momentum to get more of the vision and speech technology ready earlier, now that hiring of the required resources had fallen through. So we specified and implemented a live demo to run on a laptop which included the full Turn-Taking system, facial recognition and the first parts of the Task Dialog Manager.

T8b: Demo 0: Hannover Messe 2017

Partner Effort	Budget	Actual
CMLabs	0	0
IIIM	0	0
Total	0	0

This system was then running as a live demonstration in April 2017 at the Hannover Messe 2017 as Deliverable T8.D1: Demo 0: Hannover Messe 2017, which provides more technical details on this task.

T9: Demo 1: Collaborative Visual Detection

Partner Effort	Budget	Actual
CMLabs	0	0
IIIM	1	1
Total	1	1

This task saw the system implemented and running on two robots, collaborating via the CCMCatalog to search for human presence in the scene. This demonstration took place in Month 14 as Deliverable T9.D1: Demo 1: Collaborative Visual Detection, which provides more technical details on this task.

T10: Demo 2: Collaborative Visual Search

Partner Effort	Budget	Actual
CMLabs	0	0
IIIM	1	1
Total	1	1

By Demo 2 the full CCM Architecture has been completed which provided the system with roles, tasks and dialogue capabilities. The demonstration saw two robots working collaboratively with one human on completing a task involving interaction with a control panel, using natural dialogue. This demonstration took place in Month 19 as Deliverable T10.D1: Demo 2: Collaborative Visual Search, which provides more technical details on this task.

T11: Goal Seeking Dialogue Steering

Partner Effort	Budget	Actual
CMLabs	2	1.5
IIIM	3	3.5
Total	5	5

A big upgrade towards the capabilities needed for Demo 3 was the ability for the human to specify the task to be completed and for this we needed the dialogue system to take more of an active task in controlling the dialogue. This was implemented partly in the Task Dialog Manager and partly using shared tasks and roles via the CCMCatalog.

T12: Four-way Turn-Taking system

Partner Effort	Budget	Actual
CMLabs	2	2
IIIM	3	3.5
Total	5	5.5

The other significant change in Demo 3 was the ability to communicate with more than one human at the same time. To achieve this the Turn-Taking system was rewritten to become a Four-way Turn-Taking system. This was completed by Month 19 as Deliverable T12.D1: Four-way Turn-Taking, which provides more technical details on this task.

T13: Demo 3: Collaborative Information Extraction through Dialogue

Partner Effort	Budget	Actual
CMLabs	1	1
IIIM	1	1
Total	2	2

The final demonstration showed off the full capabilities of all the technologies implemented and integrated as part of this project and included multiple humans collaborating with multiple robots on first identifying the task to be completed and then carrying it out using natural dialogue to retrieve the information needed. This demonstration took place in Month 19 as Deliverable T13.D1: Demo 3: Collaborative Information Extraction, which provides more technical details on this task.

T14: Project Dissemination

Partner Effort	Budget	Actual
CMLabs	2	2
IIIM	2	2
Total	4	4

Throughout the project the project website and the project's social media accounts provided blog entries on the technical progress of the project. The demonstration videos have been made available on the project website as have the accompanying reports and technical papers produced. Finally, the source code for both the Psyclone platform and for the project components have been made available for download on the project website.

T15: Project Management

Partner Effort	Budget	Actual
CMLabs	1	1
IIIM	0	0
Total	1	1

The project management task ran throughout the project and dealt with practical issues such as the consequences and subsequent negotiation of the delayed funding guarantee and managing the communication with the project reviewers.

Task effort overview

Budget	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12	T13	T14	T15	Total PM
CMLabs	1	4	4	2	1	3	2	6	0	0	2	2	1	2	1	**31**
IIIM	1	1	2	0.5	0.5	1	2	3	1	1	3	3	1	2	0	**22**
Total	**2**	**5**	**6**	**2.5**	**1.5**	**4**	**4**	**9**	**1**	**1**	**5**	**5**	**2**	**4**	**1**	**53**
Actual																
CMLabs	1	4.5	5	1.5	1.5	3	1.5	6.5	0	0	1.5	2	1	2	1	**32**
IIIM	0.5	1	2	1	0.5	1	4	3	1	1	3.5	3.5	1	2	0	**25**
Total	**1.5**	**5.5**	**7**	**2.5**	**2**	**4**	**5.5**	**9.5**	**1**	**1**	**5**	**5.5**	**2**	**4**	**1**	**57**

PROJECT MILESTONES

The following section describes the milestones for the CoCoMaps project.

M1: Kick-off Meeting

The project kick-off meeting took place early in Month 1 and started the collaboration between the project partners.

M2: Support for TurtleBot platform

The new robot platform was supported by the end of Month 7, a bit later than expected due to the delay in funding and the inability to hire the robot expert staff we expected.

M3: Demonstration 0

Demo 0 was an additional demonstration which was not planned in the original proposal, but which made sense for the project to schedule in and which benefitted the ECHORD project as our demonstration was added to the group stand. This took place in Month 8 at the Hannover Messe 2017 and showed off a system which worked with dialogue and turn-taking in human conversation.

M4: Demonstration 1

Demo 1 was the first demonstration using the new robots. It showed off collaborative search for humans and took place at the IIIM offices in Month 14.

M5: Collaborative Cognitive Map complete

The final version of the CCM architecture was complete by Month 16. This saw support for sharing and negotiating about observations, tasks and roles and provided the foundation for Demo 2 and Demo 3.

M6: Demonstration 2

Demo 2 took place at the IIIM offices in Month 19. It added speech dialogue and turn-taking, first seen in Demo 0, to the robots who were now able to communicate with a human to extract information about a known task and subsequently use this to complete the task.

M7: Demonstration 3

Demo 3 concluded the demonstrations by adding support for dialogue with multiple humans and obtaining new tasks from the humans, using more multi-modal input to support the turn-taking and dialogue system. It took place at the IIIM offices in Month 19.

M8: Project Completed

The project completed after one month’s agreed extension by the end of Month 19.

PROJECT DELIVERABLES

The following section describes the deliverables for the CoCoMaps project. These are also available online at **https://cmlabs.com/cocomaps/publications**.

T1.D1: Specification of Experimental Platform

This report contains the technical specification of the experimental platform which was implemented during the project and demonstrated fully in Demo 3. You can download the report from **this link**.

T6.D1: Current state-of-the-art implementation

This report details the initial implementation of the Cognitive Maps Architecture after it had been ported to the new platform. You can download the report from **this link**.

T8.D1: Demo 0: Hannover Messe 2017

This demonstration video and accompanying report show the first project technology demonstration at the Hannover Messe in April 2017. You can download the report from **this link** and you can see the video here:

https://youtu.be/HqtN0499Vq8

T8.D2: Draft Collaborative Cognitive Map

This report describes the initial draft implementation of the Collaborative Cognitive Map architecture, subsequently used in Demo 1. You can download the report from **this link**.

T9.D1: Demo 1: Collaborative Visual Detection

This demonstration video and accompanying report show the second project technology demonstration carried out in November 2017. You can download the report from **this link** and you can see the video here:

https://youtu.be/maAh8N3nRFg

T8.D3: Final Collaborative Cognitive Map

This report describes the final version and implementation of the Collaborative Cognitive Map architecture, subsequently used in Demo 2 and Demo 3. You can download the report from **this link**.

T10.D1: Demo 2: Collaborative Visual Search

This demonstration video and accompanying report show the third project technology demonstration carried out in March 2018. You can download the report from **this link** and you can see the video here:

https://youtu.be/FsFgwil0or4

T12.D1: Four-way Turn-Taking

This report deals with the technology and implementation of the upgrade to four-way turn-taking within the Collaborative Cognitive Map architecture, subsequently used in Demo 3. You can download the report from **this link**.

T13.D1: Demo 3: Collaborative Information Extraction

This demonstration video and accompanying report show the fourth and final project technology demonstration carried out in March 2018. You can download the report from **this link** and you can see the video here:

https://youtu.be/Xb492moHr20

T15.D1: Demos, results and literature publicly available

This is this report the final report of what was achieved in the CoCoMaps project by going through the tasks, milestones, deliverables and KPIs. You can download the report from **this link**.

PROJECT IMPACT KPIS

The following section describes the Impact KPIs for the CoCoMaps project.

1: Industrial collaborations

As part of the project we have discussed our technology with a number of commercial partners. Several of these have shown significant interest and two were able to provide us with formal letter of interest. As these contain personal identifiable information copies of them will be provided separately and not as part of the publicly available information for the project.

2: Psyclone framework

The original Psyclone platform was a closed-source commercial product. Version 2 of Psyclone was rewritten from scratch with the intention to make it available under dual licence; a commercial licence allowing companies to enhance the product without having to make their changes public and LGPL to allow collaboration in the continued development of the product as a whole. As part of the CoCoMaps project we released Psyclone 2 as open source under the LGPL licence and version 2.0.1 is now available for public download from

https://cmlabs.com/psyclone/download

3: Academic collaborations

Similarly we have discussed our technology with a number of academic partners. Quite a few of these have shown significant interest and three were able to provide us with formal letter of interest. As these contain personal identifiable information copies of them will be provided separately and not as part of the publicly available information for the project.

4: Psyclone + project bundle, ready for commercially funded integration projects

To ensure continued interest and usability of the components developed as part of this project and of the full system functionality as a whole we have additionally made all the source code for our components available for download free to use by anyone under a BSD-style licence. We were not able to provide download access to the third-party commercial libraries used in the system, but we are of course happy to help anyone interested in obtaining these or finding alternative technologies as replacements. Version 1.0.90 of the CoCoMaps source code is now available for public download from

https://cmlabs.com/cocomaps/downloads

PROJECT TECHNICAL KPIS

The following section describes the Technical KPIs for the CoCoMaps project.

KPI 1: Ability of current state of the art running on one computer

In Demo-3 CoCoMaps runs on three computers - on each of the robot's computers and on a third offboard computer. Each robot computer runs the full system which includes the state of the art for turn-taking and the cognitive map. This KPI has been met.

KPI 2: Ability of real-world robot-robot interaction using new collaborative CMArch

In Demo-3 and previous demonstrations the robots demonstrate the ability to share and query data, as well as negotiate roles and tasks in near real-time, with (near) 100% success rate for several runs, and within very acceptable time frames (see Tables 7-10). This KPI has been met.

KPI 3: Ability of real-world multi-robot-human interaction using collaborative CMA and speech

Demo-3 has shown that our system is able to allow multiple robots to collaborate successfully both between themselves using the CCM architecture and with multiple humans via speech and natural dialogue (see Tables 8-10). This KPI has been met.

KPI 4: Efficiency of collaborative detection of humans

Collaborative detections of humans were proven in Demo-1 and further refined and significantly improved in Demo-2 and Demo-3. The efficiency measurements are provided in Table 7 and show that two robots can more effectively detect humans when collaborating on the task. This KPI has been met.

KPI 5: Efficiency of collaborative tracking of humans

Collaborative tracking of humans was proven in Demo-1 and further refined and significantly improved in Demo-2 and Demo-3. The efficiency measurements were provided in Table 7 and show that two robots can more effectively track humans when collaborating on the task. This KPI has been met.

KPI 6: Efficiency of collaborative information extraction through dialogue

Collaborative information extraction from humans was proven in Demo-2 and further refined and improved in Demo-3. The efficiency measurements were provided in Table 10 and show that the robots can both extract information and perform a remote task at the same time. This KPI has been met.

KPI 7: Efficiency of collaborative task extraction through dialogue

Collaborative task extraction from humans was proven in Demo-3. The efficiency measurements were provided in Table 10 and show that the robots can both extract the task and the required information as well as perform a remote task at the same time. This KPI has been met.

KPI 8: Real-time algorithms for the estimation of the emotional state of the humans and speaker estimation from facial expressions and head movement

The algorithms for estimating emotional state of humans and speaker estimation were used and shown in Demo-3. The efficiency measurements for both were provided in Table 10 and show that the system is able to use visual analysis of the humans' faces to estimate emotions, head movement and speaking activity. This KPI has been met.

KPI 9: Human-leg and torso tracker using 3D information from the navigation camera

The algorithms for detecting legs and torso were used in the HumanDetector module, used in Demo-3 to detect and track the 3D position of the human. The efficiency measurements for this were included in the Human Detection entry in Table 7 and show that the robots are able to estimate the position in the room by using 3D depth information to find either the torso or the legs of the person. This KPI has been met.

KPI 10: Participant Negotiation Module, distributed reasoning/data fusion system for estimation of observations of the participants.

The module for negotiating observations of humans were demonstrated in Demo-3 and the functionality was used every time the two robots shared information about their observations. Specifically, this was measured in the **Person Identified: Collab.** entry in Table 7. This shows that the robots are able to discuss and negotiate both observations, roles and tasks with other robots via the CCMCatalog. This KPI has been met.

PROJECT DISSEMINATION

1 Website of experiment
The project team set up the main project website at the start of the project.
http://cocomaps.org

2 Press release 1
In the first months of the project the project team issued the first project press release to alert the media and public to the fact that the project had started.
http://download.cmlabs.com/cocomaps/CoCoMaps_Press_release.pdf

3 Press release 2
In connection with the project's attendance at the Hannover Messa 2017 the project team issued the second press release of the project.
http://download.cmlabs.com/cocomaps/CoCoMaps_Press_Release_Hannover_2017.pdf

4 Final demo Press release
After the project had completed and all the reports and videos were available on the project website the project team issued the final press release of the project.
http://download.cmlabs.com/cocomaps/CoCoMaps_Final_Press_Release.pdf

5 Multi media report
Throughout the project we have added blog posts to our website and social media and as video recordings became available we added these to the Media part of the project website.
https://cmlabs.com/cocomaps/media

6 Networking w customers (Marel)
We have had two meetings with Marel who now have their own research department. Discussions are ongoing, but they are very interested in the real-time CoCoMaps architecture which allows multiple robots to negotiate observations, roles and tasks through the CCMCatalog.

7 Networking w customers (Magic Leap)
We engaged with early stage discussions with Magic Leap which are still ongoing. We hope that the next step will happen towards the end of 2018.

8 Networking w customers (Honda)
We engaged with early stage discussions with Honda which are still ongoing. We hope that the next step will happen over the summer 2018.

9 Attendance to trade fairs (Consumer Technology Association / CES)
The project team attended CES 2017 and met with a number of potential technology providers and future customers. Amongst the technology partners were audio and visual processing companies and a number of elderly care providers showed interest in the project's technology.

10 Attendance to trade fairs (Hanover Messe 2017)
The project team attended and demonstrated the early stage project technology at Hannover Messe 2017. This involved the two-way turn-taking system and speech input and output.

11 Attendance to trade fairs (Hanover Messe 2018)
Attendance for Hannover Messe 2018 was not offered by the Echord Group so no attendance was possible. We hope to have the opportunity again in 2019.

12 Attendance to scientific conferences (CES in the US booked and scheduled)
This task appears to be a duplicate of task 9 above – we believe that it was added to the monitoring system by mistake.

13 Attendance to scientific conferences (Hanover Messe 2017)
This task appears to be a duplicate of task 10 above – we believe that it was added to the monitoring system by mistake.

14 Attendance to scientific conferences (Hanover Messe 2018)
This task appears to be a duplicate of task 11 above – we believe that it was added to the monitoring system by mistake.

15 Create posters/leaflets/roll-ups
Flyers were created for handout at the CES 2017 in January 2017 and both flyers and posters were created for the Hannover Messe April 2017.
http://download.cmlabs.com/cocomaps/Natural_Human-Robot_Interaction.pdf
http://download.cmlabs.com/cocomaps/Hannover_Poster.pdf

16 Social media
Throughout the project we have added blog posts to our website and social media.
https://twitter.com/cocomapsorg
https://www.facebook.com/cocomaps/

PROJECT RESULTS

The four demonstrations – Demo-3 in particular – have successfully demonstrated multi-party collaboration capabilities using dialog, integrated with navigation and appropriate visual competencies and virtual control panel interaction, where two robots work in an environment extracting directions from humans. Specifically, the robots collaborate and communicate with each other and with two humans to perform a task initiated by the humans. The collaboration involves extracting the task to perform including sub-tasks as well as individual actions dynamically from two humans during the session via 4-way dialogue. Task and role assignment is done by the robots dynamically during the interaction. The communication and dialogue acts are both time- and context-dependent.

We have provided detailed reports, demonstration videos and measurements to document that each milestone has been met and each KPI has been delivered as part of the project and within the project running period and budget.

FUTURE WORK

The CoCoMaps project has demonstrated the commercial potential of the project technologies and that they are ready to be integrated into real-world projects. Further tuning and task specific optimisation is of course necessary for each specific application and use, however we have demonstrated both the general purpose nature of the architecture, and its ability to be readily integrated for new applications.

A large amount of work has gone into using multi-modal sensor data and near real-time processing to enhance social Turn-Taking in a human centric environment and more work lies ahead to improve the performance and accuracy of this. This includes investigating what additional data can be usefully added to further augment the pool of information, such as using a directional microphone array to assist the speaker identification and noise handling.