Initial Plan

Bilingual Natural Language Knowledge Generation from Simple Scene Images

Robert Harris

C1334989

Supervisor: Professor Alun Preece

Module Code: CM3203

Module Title: One Semester Individual Project

Credits: 40

January 30th 2017

Project Description

Human Computer Collaboration in recent times has had a heavy interdisciplinary focus, on ways a human agent can best collaborate with a machine. The two traditional channels of HCC involve giving computers human-like abilities, generally focusing on language; with the second being a machines ability to work alongside humans. Research suggests that a combination of both approaches is the most effective way of encouraging human-computer collaboration [1].

With HCC in mind, this project will aim to create a software agent that can play the SHERLOCK (Simple Human Experiments Regarding Locally Observed Collective Knowledge) tactical intelligence game, which involves humans participating in intelligence, surveillance, and reconnaissance (ISR) tasks in physical environments [4]. The motivation behind this project is too develop the machine agent so it's much closer to a "full" player rather than what it currently is. It is hoped the machine agent will be able to "perceive" scenes "first hand", compared its current implementation which currently "perceives" the world in a "second handed" manner. This agent would need two key abilities: Perceiving the world (Recognising simple poster scenes) and describing the world (In multiple natural languages which can be communicated to human players).

One example of HCC that is both human readable and machine compatible is through Controlled Natural Language (CNL). This compatibility is achieved through sub setting a traditional natural language by introducing restrictions on vocabulary. One of the most prominent CNLs was developed by IBM, and is defined by the International Technology Alliance as a form of Controlled English (CE) [2]. Although inspired by CLCE (Common Logic Controlled English), ITA CE is less strict in terms of precision: It has an "Informal meaning and a semi-formal mapping to predicate logic". The two most common forms of language rules are "logical rules" and "rationale" statements. A "logical rule" would look similar to the following: "if (the person X has the person Y as brother) and (the person Z has the person X as father) then (the person Z has the person Y as uncle) ". A "rationale" rule may look like: "("the plan has failed" because "there was a misunderstanding".)" [3].

One approach of enabling human machine collaboration through Controlled English is by using CENode (Controlled English Node). CENode enables Human-Machine conversations at the network edge and allows humans to update and query a shared knowledge base directly using ITA CE. CENode also supports versatility in deployment, functioning in web applications or as a standalone app on handheld device or servers. Developed in JavaScript, CENode is relatively lightweight controlled natural language knowledge based system, which is ideal for handheld and ubiquitous devices [5].

The primary deliverable will be based on the SHERLOCK game which is an application which implements the CENode library, along with the "ask and tell" conversational protocol. The "ask and tell" protocol allows the user to "query" the

knowledge base for previously crowd sourced knowledge or “tell” the knowledge base about previously unknown and potentially useful knowledge. The project will aim to assist the human in the “telling” action of the protocol, by delegating object classification to an intelligent system.

With the primary aim to classify objects within images in the SHERLOCK game, its hoped classifications can be returned in a bag like data structure. E.g {**Elephant, Eliffant**} or {**Soccer, Pel-Droed**}. This will satisfy the description ability of the agent. This bag data structure will allow for a bilingual classification in both English and Welsh, with the desired string language specified by the human. Depending on the language specified, the CNL string would have to be in either ITA Controlled English or Controlled Welsh. Controlled Welsh can be implemented following the specification outlined by Elliot Howell in his Welsh Natural Language Knowledge-Base [6].

Classifying images should aide the human in the surveillance and reconnaissance of game objectives. Although humans are better at interpreting natural language and images than computers, using a computer to classify images is one of many approaches to achieve a bilingual response. The end goal is to eventually aide Collective Situational Understanding (CSU) between humans who do not share a common language.

Using a human agents smart phone device, the human will have the option to take a snapshot of a scene with their camera and submit it for classification. This will satisfy the perception ability of the agent. Classified objects can then be used to construct a CNL string in Welsh or English. Experimentation will be needed to decide on levels of automation while using this functionality. Decisions will be made on whether to classify image objects and automatically submit them to the knowledge base, or whether a human should oversee the classification process and provide conformation on the constructed CNL string.

Image processing will be achieved using the open source vision library OpenCV [7]. Released under a BSD license, OpenCV is free to use for both academic and commercial use. OpenCV would be ideal for this project as it has support for classification techniques such as SIFT, SERF and MESR, along with compatibility with all major operating systems. Although OpenCV is implemented in C/C++, is has support for Python wrappers which will enable me to implement a Python image classification system. An advantage of using wrappers, is that there’ll be no performance decrease compared to using the native C/C++ implementation. This is because the wrappers will be executing the native C/C++ code from the Python calls.

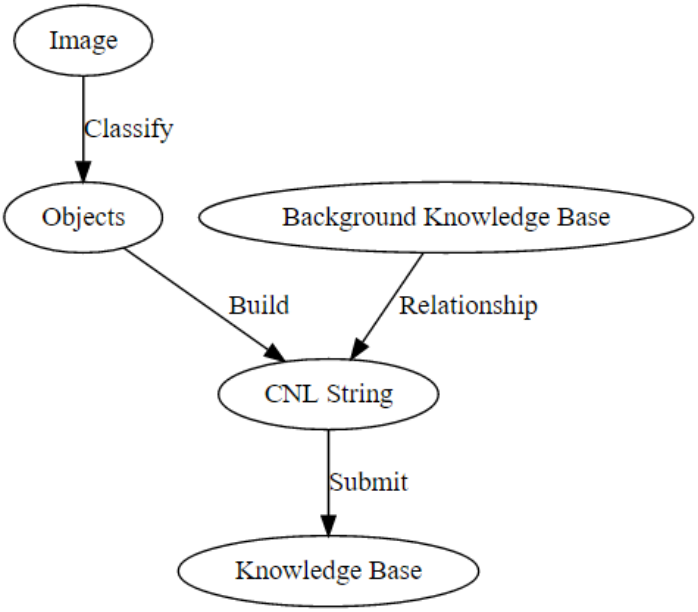
As we are classifying non-real-time scenes with white, non-congested backgrounds, image classification should not be too difficult given enough training data.

A challenge faced in using OpenCV will be linking my Python image classification system to the JavaScript source code of SHERLOCK. Image classification would only work in the always online version of the game, as images will need to contact a vision library located on a server. The consequence of this is that image classification will be unavailable to users running the offline version of the game.

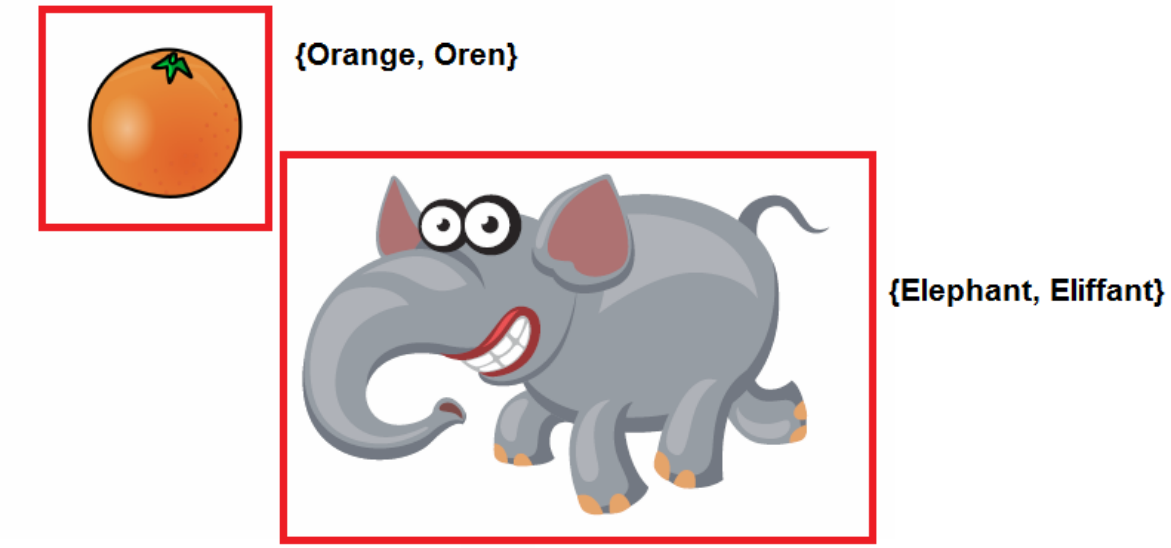
This problem could potentially be addressed in future developments beyond the scope of this project.

The primary target will see an instance of the Welsh and English versions of the SHERLOCK game receive bilingual strings describing objects, returned from a still scene image using a smartphones camera device.

The below diagram aims to highlight a rough flow of the classification process;



An image may be classified bilingually as such;



It is hoped a CNL string in both languages may be constructed in each version of the game.

English CNL: *The **Elephant** eats the **Orange***

Welsh CNL: Mae **Eliffant** yn bwyta **Oren**

Project Aims & Objectives

The aims and objectives described below outline the learning outcomes and deliverables this project hopes to achieve. They are listed in no explicit order, with no consideration for weight or time of the objective. Deliverables will be clearly noted, with all aims and objectives evaluated at the end of the project.

- An understanding of appropriate Levels of Automation (LOAs).
    - Measure automation level from 1 – 10. Level 1 represents the computer offering no assistance; human taking all action and decisions. Level 10 represents the computer acting autonomously, ignoring the human.
- An understanding of relevant image classification techniques and how to deploy them.
    - Conduct research into open source computer vision libraries such as OpenCV. Can expand my knowledge in these areas by studying past final year projects, completing online tutorials and consulting university lecturers. (Can study Allan Lainchbury's Final year project titled "Object Recognition on Mobile Devices" [8])
    - Will need to download OpenCV, setup the environment, setup paths, directories and make them version controlled. Analyse how different techniques (eg: SIFT, MSER, SURF) classify images in Python and analyse their outputs.
    - Deploy my preferred technique in a Telegram [11] prototype.
    - **Deliverable 1: Working prototype of bi-lingual image classification using telegram (Return English and Welsh string of objects found)**
- An understanding of areas in the SHERLOCK code base where modifications will need to be made to support user image input.
    - Pull the code base from git.
    - Learn the JavaScript programming language
    - Achieve an understanding of SHERLOCKS's backend semantics.
    - Add camera support to SHERLOCK
    - Link image inputs to the Python vision library backend.
- A demonstration of bilingual classification in SHERLOCK
    - Demonstrate the modified SHERLOCK code base to a select group of users.
    - Gather feedback.
    - Make modifications to the implementation based on feedback.
    - **Deliverable 2: Demonstration of image classification being used to aide CNL input by returning bilingual objects strings.**

- Implementation of automatically generated relationships between objects
    - Construction of strings formed from CNL for both Welsh and English classifications. (e.g: If the image classification recognises an **elephant** object and a **pineapple** object then the relationship should be **eats**)
    - Will need to implement in both Welsh and English version of the game
    - **Deliverable 3: Demonstration of automated relationships between classified objects**
- Testing the modifications to the backend and UI
    - Perform documented testing of new features
- A detailed report of the projects outcomes and findings.
    - A report containing an abstract of the project, research, approaches on deliverables and an outline of the overall findings.
    - **Deliverable 4: Report**

Ethical Considerations

Reviewing the research ethics and guidelines outlined by the school [9], I have concluded to the best of my ability that no planned work will need to be reviewed by the Research Ethics Committee (REC). During development and testing, no personally identifiable or compromising information will be collected or processed by myself or the system. If any personal data is accidentally collected, then appropriate parties will be informed and action can be taken. If this stance on ethics is changed at any point, then the REC will be informed and appropriate documentation will be completed.

Work Plan

The following work plan has been developed to ensure suitable time frames have been allocated to research, development and implementation of each task described. University contact hours, volunteering commitments, coursework deadlines and holidays have been considered in the outline of this plan. With week 1 starting at the **23/01/2017**, the work plan outlines the next 15 weeks of the projects life span including research time, notable events and deliverables.

Pre-Week 1
24/10/2016 – 22/01/2017

- Background research about the project's origin, recent developments and future objectives of CENode/ SHERLOCK experiments.
- Meetings with Professor Alun Preece on how to build on the Welsh CNL outlined by Elliot Howells.
- Supervisor meetings discussing scope and deliverables and implementation of the project.

Weeks 1 – 2
23/01/2017 – 05/02/2017

- Write and submit initial plan. 30/01/16.

- Research image classification techniques in the OpenCV library to classify SHERLOCK scene objects
- Prototype using the Telegram messenger API to send images and receive classifications.
- Start reading the SHERLOCK code base and understand its semantics.
- Build an image repository of the different scenes in the game.

Weeks 3 – 4
06/02/2017 – 19/02/2017

- Scope the SHERLOCK code base and identify areas that need to be modified to allow for image input.
- Explore whether it would be possible to merge both Welsh and English versions of SHERLOCK into a single environment.
- Continue research into classification techniques.
- Decide on levels of automation available while playing SHERLOCK.
- **Deliverable 1: Telegram prototype to Alun Preece.**
- Receive feedback from prototype.

Weeks 5 – 6
20/02/2017 – 05/03/2017

- Add functionality so the user can decide to receive English / Welsh classifications.
- Add functionality to access user's camera within SHERLOCK.
- Update the apps UI to accommodate new functionality. (Produce sketches)
- Link the image input to Python classification software.
- **Deliverable 2: SHERLOCK application which can classify objects based off image input.**
- Evaluate Demo.

Weeks 7 – 8
06/03/2017 – 19/03/2017

- Explore how to add relationships between objects using the background knowledge base.
- Implement relationships between classified objects for both English and Welsh CNL (e.g: Elephant **eats** pineapple/ Eliffant **bwyta** Pinafal).

Weeks 9 – 10
20/03/2017 – 02/04/2017

- Write a short report/ detailed notes on progress, findings, and outcomes so far. Evaluate the progress made so far and ways one could do things differently in the future.
- Reflect on the bilingual functionality, by reviewing its implementation, effectiveness, and accuracy.

Weeks 11 – 12
03/04/2017 – 16/04/2017

- Test new functionality with use case documentation. A nominated tester will complete use cases.

- **Deliverable 3: A bilingual natural language knowledge generator to assist a human playing the SHERLOCK game.**

Weeks 13 – 14
17/04/2017 – 30/04/2017

- Produce draft of final report.
- Amend final draft.

Week 15 Onwards
01/05/2017 – 09/06/2017

- Finalise report and submit by 05/05/2017.
- Prepare for Viva. Date TBA.
- **Deliverable 4: Report.**

Supervisor Meetings

Meetings will be held with my supervisor Professor Alun Preece in a fortnightly fashion to supply face to face updates on the project. Regular interaction will also take place on the Slack [10] instant messaging application to discuss more urgent queries and issues.

Currently scheduled meetings include:

- Thursday, January 19th 2017
- Thursday, February 2nd 2017
- Thursday, February 16th 2017
- Thursday, March 2nd 2017
- Thursday, March 16th 2017
- Thursday, April 13th 2017
- Thursday, April 27th 2017
- Thursday, May 11th 2017

Gantt Chart

Week	1	2	3	4	5	6	7	8	9	10	11	12	13	14	>15
Start of Week	23/01/2017	30/01/2017	06/02/2017	13/02/2017	20/02/2017	27/02/2017	06/03/2017	13/03/2017	20/03/2017	27/03/2017	03/04/2017	10/04/2017	17/04/2017	24/04/2017	01/05/2017
Task															
Initial Plan															
OpenCV Image classification research															
Prototype with Telegram															
Scope the SHERLOCK code base															
Explore levels of automation															
Deliver a functioning proof of concept with Telegram															
Make SHERLOCK compatible to image input															
Link Python backend with image input															
Deliver bilingual classification within SHERLOCK															
Research how to add bilingual relationships															
Implement bilingual relationships															
Produce a lessons learnt/ evaluation report															
Testing															
Deliver a bilingual natural language knowledge generator															
Write final report															
Viva															

References

[1] L. Terveen. Overview of human-computer collaboration. Knowledge-Based Systems, 67–69, 1995.

[2] IBM, IBM Controlled Natural Language Processing Environment. [Online]. Available: https://www.ibm.com/developerworks/community/groups/service/html/communityview?communityUuid=558d55b6-78b6-43e6-9c14-0792481e4532. Accessed: Jan. 29, 2017

[3] T. Kuhn, A Survey and Classification of Controlled Natural Languages. [Online]. Available: http://www.aclweb.org/anthology/J14-1005. Accessed: Jan. 29, 2017

[4] A. Preece, C. Gwilliams, C. Parizas, D. Pizzocaro, J. Z. Bakdash, D. Braines. Conversational Sensing, 5-7, 2014. [Online]. Available: http://arxiv.org/pdf/1406.1907.pdf. Accessed: Jan. 29, 2017

[5] Cardiff University, IBM UK, CENode.js. [Online]. Available: http://cenode.io. Accessed: Jan. 29, 2017

[6] E. Howells, "Final Report: Welsh Natural Language Knowledge-Base," in https://www.cs.cf.ac.uk/PATS2/. [Online]. Available: https://www.cs.cf.ac.uk/PATS2/@archive_file?c=&p=file&p=538&n=final&f=1-1110117_-_Final_Report.pdf. Accessed: Jan. 29, 2017.

[7] "OpenCV," 2016. [Online]. Available: http://opencv.org/. Accessed: Jan. 29, 2017.

[8] A. Lainchbury, "Object Recognition on Mobile Devices,". [Online]. Available: https://www.cs.cf.ac.uk/PATS2/@archive_file?c=&p=file&p=338&n=final&f=1-C1265364_Allan_Lainchbury_Final_Report.pdf. Accessed: Jan. 30, 2017.

[9] "Research Ethics,". [Online]. Available: http://users.cs.cf.ac.uk/I.Spasic/ethics/. Accessed: Jan. 30, 2017.

[10] Slack, "Slack: Where work happens," Slack. [Online]. Available: https://slack.com/. Accessed: Jan. 30, 2017.

[11] "Telegram – a new era of messaging," Telegram. [Online]. Available: https://telegram.org/. Accessed: Jan. 30, 2017.