ARL-RP-0560 • Dec 2015

US Army Research Laboratory

SHERLOCK: Simple Human Experiments Regarding Locally Observed Collective Knowledge

by A Preece, W Webberley, D Braines, N Hu, T La Porta, E Zaroukian, and J Z Bakdash

Reprinted from the International Technology Alliance Collaboration System [accessed 2015 Dec 2]. https://www.usukitacs.com/node/3069. Paper presented at the 2015 US/UK International Technology Alliance Fall Meeting; 2015 Sep 15–17; College Park, MD.

NOTICES

Disclaimers

ARL-RP-0560 ● DEC 2015

US Army Research Laboratory

SHERLOCK: Simple Human Experiments Regarding Locally Observed Collective Knowledge

by A Preece and W Webberley
School of Computer Science and Informatics, Cardiff University, Cardiff, UK

D Braines
Emerging Technology Services, IBM United Kingdom, Ltd, Hursley Park, Winchester, UK

N Hu and T La Porta
Department of Computer Sciences and Engineering, The Pennsylvania State University, University Park, PA

E Zaroukian and J Z Bakdash
Human Research and Engineering Directorate, ARL

Reprinted from the International Technology Alliance Collaboration System [accessed 2015 Dec 2]. https://www.usukitacs.com/node/3069. Paper presented at the 2015 US/UK International Technology Alliance Fall Meeting; 2015 Sep 15–17; College Park, MD.

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED (From - To)**
December 2015	Reprint Report	October 2014–September 2015

4. TITLE AND SUBTITLE

SHERLOCK: Simple Human Experiments Regarding Locally Observed Collective Knowledge

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

A Preece, W Webberley, D Braines, N Hu, T La Porta, E Zaroukian, and J Z Bakdash

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

US Army Research Laboratory
ATTN: RDRL-HRS-E
Aberdeen Proving Ground, MD 21005-5425

8. PERFORMING ORGANIZATION REPORT NUMBER

ARL-RP-0560

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

Reprinted from the International Technology Alliance Collaboration System [accessed 2015 Dec 2]. https://www.usukitacs.com/node/3069. Paper presented at the 2015 US/UK International Technology Alliance Fall Meeting; 2015 Sep 15–17; College Park, MD.

14. ABSTRACT

We describe the design of human-machine conversation experiments that support the evaluation of our contextaware approach in coalition decision making at or near the network edge. The studies, named SHERLOCK (for Simple Human Experiments Regarding Locally Observed Collective Knowledge), involve humans participating in simple intelligence, surveillance, and reconnaissance (ISR) tasks in physical environments, using mobile devices in situ, or online. Experimental tasks are undertaken by multiple participants operating as a coalition, with collaboration mediated by machine agents. We illustrate the SHERLOCK using two specific experiment designs. In each one, the participants' task is to locate a number of target individuals and identify their specific features. Experiment 1 is a crowdsourcing "whodunit" scenario involving intelligence in synthetic, natural, and hidden situations. All human input is via a conversational agent which mediates information sharing between participants. Experiment 2 is an ISR asset assignment scenario, using simulated sensing assets selected either by a human, via an algorithm, or a combination of the human and machine. Participants' success in both experiments depends on their ability to make effective use of the conversational system, including being able to provide interpretable information to the conversational agent, and to obtain information via the conversational interface to assist them.

15. SUBJECT TERMS

controlled natural language; conversational interface; human machine interaction; human-computer collaboration; intelligence, surveillance, and reconnaissance

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
					E Zaroukian
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			**19b. TELEPHONE NUMBER (Include area code)**
Unclassified	Unclassified	Unclassified	UU	8	410-278-9050

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18

SHERLOCK: Simple Human Experiments Regarding Locally Observed Collective Knowledge

A. Preece*, W. Webberley*, D. Braines†, N. Hu‡, T. La Porta‡, E. Zaroukian§ and J. Z. Bakdash§

*School of Computer Science and Informatics, Cardiff University, Cardiff, UK; contact email: PreeceAD@cardiff.ac.uk

†Emerging Technology Services, IBM United Kingdom Ltd, Hursley Park, Winchester, UK

‡Department of Computer Science and Engineering, The Pennsylvania State University, USA

§Human Research and Engineering Directorate, US Army Research Laboratory, Aberdeen Proving Ground, USA

***Abstract*—We describe the design of human-machine conversation experiments that support the evaluation of our context-aware approach in coalition decision making at or near the network edge. The studies, named SHERLOCK (for Simple Human Experiments Regarding Locally Observed Collective Knowledge), involve humans participating in simple intelligence, surveillance, and reconnaissance (ISR) tasks in physical environments, using mobile devices in situ, or online. Experimental tasks are undertaken by multiple participants operating as a coalition, with collaboration mediated by machine agents. We illustrate the SHERLOCK using two specific experiment designs. In each one, the participants' task is to locate a number of target individuals and identify their specific features. Experiment 1 is a crowdsourcing "whodunit" scenario involving intelligence in synthetic, natural, and hidden situations. All human input is via a conversational agent which mediates information sharing between participants. Experiment 2 is an ISR asset assignment scenario, using simulated sensing assets selected either by a human, via an algorithm, or a combination of the human and machine. Participants' success in both experiments depends on their ability to make effective use of the conversational system, including being able to provide interpretable information to the conversational agent, and to obtain information via the conversational interface to assist them.**

I. Introduction

Over the past four years we have been researching technologies to support human-machine collaboration in the context of coalition intelligence, surveillance, and reconnaissance (ISR) tasks [1]. We have focussed on approaches using controlled natural language (CNL) to provide representations of information and knowledge that are human readable and writable, as well as machine-processable. Such approaches allow the machine to perform computational reasoning over a knowledge base while expressing rationale that is human-understandable. The overall goal of this research is to use human-machine collaboration to enhance human cognition.

Our recent focus has been on behavioural research to test the effectiveness of the technology for people using simulated ISR tasks. To this end, we have designed a platform for running a series of experiments in which human participants work alone or collectively on gathering synthetic and natural information either in situ or online. Participants address the tasks by interacting with a CNL agent through dialogues in which they can use both natural language (NL) and CNL. The tasks involve the collection of locally observed information, such as simulating activities humans would perform on patrol or while operating a remote sensing system. However, tasks are simplified to allow participation without any specific ISR training as well as easing some aspects of the natural language processing (NLP) performed by the agent (since NLP is not the focus of our research).

In the experiments, the CNL agent and supporting infrastructure are designed to assist the participants in their tasks. The precise details of what assistance is provided in each case vary from experiment to experiment; generally, tasks involve information collection, sharing, or adding context to local observations to make them more meaningful (i.e., turning information into knowledge). Dependent measures include the time to enter information or make decisions, and their accuracy and/or quality. Our platform is called SHERLOCK (for Simple Human Experiments Regarding Locally Observed Collective Knowledge).[1]

This paper is structured as follows: Section II describes the main elements of the experimental world and participants' tasks. Section III summarises the features of the CNL agent that participants interact with and the assistive capabilities developed in our research. Section IV illustrates the use of the SHERLOCK platform with two example experiments. Finally, Section V concludes and points to future work.

II. Experimental World

The world of SHERLOCK experiments consists of sets of: persons of interest (POIs), locations, objects, sensors (human and machine), and attributes of all of these entities. Participants are provided with a subset of information about the world, for example, a map of locations, "mugshots" of POIs, and a list of attributes of interest. Their task is to collect specific information, for example, to discover the location of POIs and the values for their particular attributes, or just to collect as much information as possible. The world is modelled in CNL, making it relatively straightforward to define sets of instances for a particular experiment setup.

Worlds have information that is synthetic, natural, or some combination of both. For example, human actors may portray fictional POIs inhabiting real-world locations, or cartoon characters may inhabit synthetic virtual locations. Participants

[1]The name was chosen to give participants a sense that their tasks involve elements of detection and deduction.

Fig. 1. The CNL agent used in a SHERLOCK game

may be tasked to report on natural real-world objects or distinctive objects placed in locations for the purposes of creating artificial scenes. Given that our research focus is not on NLP, synthetic elements are often easier for participants to describe to the agent (with lower complexity and potential for ambiguity). For example, it is easier to describe the location of an object in cases where that location has a label, such as a room number, than it is to describe an "unlabelled" location, such as a corridor or public space. Highly distinctive objects, such as pieces of fruit or sporting equipment, are also relatively easy to describe.

III. CNL Agent Capabilities

A full description of the capabilities that are the focus of our research is given in [2]. The core capability is a conversational agent that supports a set of interactions defined in a protocol, including interactions to verify machine-generated mappings of NL to CNL (*confirm*), question-answer interactions (*ask/tell*), imparting of information (*tell*), and obtaining rationale for some piece of machine-generated information (*why*). Figure 1 shows a screenshot from one of the mobile apps created for a SHERLOCK experiment. The conversation between the user ("Alun") and the agent ("Sherlock") is displayed as a conventional text "chat" thread. The user's input is in NL, which the agent maps to CNL via NLP. The user can either confirm that the CNL is an acceptable interpretation of his or her input (shown as the "Yes" message here) or reject the agent's interpretation and try again. The agent can also ask the user questions, as shown in the bottommost message.

IV. Example Experiment Designs

Design 1: "Whodunit". This design features six POIs played by human actors. Participants are assigned a set of 54 questions in total, mostly concerning (fixed) attributes of the six characters. Values for the attributes — including a character's location, shirt colour, preferred fruit, and hobby — are discoverable by visiting a set of locations around a university building. In addition, a subset of the 54 questions requires participants to locate real-world objects in the vicinity of the given locations (e.g., works of art and historical artefacts). Participants gain points for submitting answers to questions and for reporting other real-world objects in the environment (including both "mundane" items, such as furniture and office equipment, and a number of "anomalous" objects, such as balloons and toys that would not normally be found in a university building). The CNL agent mediates the collection of information as a crowdsourcing exercise. Each participant has access to a shared dashboard where questions are considered "settled" once a sufficient number of consistent answers has been received. Different groups of participants can be equipped with different variants of the CNL agent (e.g., a version supporting only confirmatory and tell interactions, and a version that also supports ask/tell querying, allowing participants to access other players' submitted information). Gradually, as information is acquired, participants can uncover a "crime" committed by one of the POIs against another.

Design 2: ISR asset assignment. In this design, POIs can move between locations. Participants can only view locations via simulated sensing assets (e.g., colour or B&W cameras). The participants' goal is to locate the POIs (and maintain current location information) and report on visible attributes of the POIs (e.g., carried objects, shirt colour). Participants can work alone or in a coalition team with sharing of assets and information. The agent capabilities vary in terms of the degree of intelligent and predictive assistance given to participants in terms of asset selection, using state-of-the-art algorithms [3].

V. Conclusion

Initial results from SHERLOCK experiments, currently being written for publication, indicate that untrained users can quickly become productive with the CNL agent, allowing them to cooperate effectively with other "coalition members". So far, approximately 45 people (undergraduate students and industry researchers) have participated in SHERLOCK experiments. Further experiments will be run through April 2016.

Acknowledgements

This research was sponsored by the US Army Research Laboratory and the UK Ministry of Defence and was accomplished under Agreement Number W911NF-06-3-0001. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the US Army Research Laboratory, the US Government, the UK Ministry of Defence or the UK Government. The US and UK Governments are authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon.

References

[1] A. Preece, D. Pizzocaro, D. Braines, D. Mott, G. de Mel, and T. Pham, "Integrating hard and soft information sources for D2D using controlled natural language," in *Proc 15th International Conference on Information Fusion*, 2012.

[2] A. Preece, C. Gwilliams, C. Parizas, D. Pizzocaro, J. Z. Bakdash, and D. Braines, "Conversational sensing," in *Proc Next-Generation Analyst II (SPIE Vol 9122)*. SPIE, 2014.

[3] N. Hu, T. L. Porta, D. Pizzocaro, and A. Preece, "A system architecture for decision-making support on ISR missions with stochastic needs and profit," in *Proc Ground/Air Multisensor Interoperability, Integration, and Networking for Persistent ISR IV (SPIE Vol 8742)*. SPIE, 2013.

1 DEFENSE TECHNICAL
(PDF) INFORMATION CTR
DTIC OCA

2 DIRECTOR
(PDF) US ARMY RESEARCH LAB
RDRL CIO LL
IMAL HRA MAIL & RECORDS
MGMT

1 GOVT PRINTG OFC
(PDF) A MALHOTRA

1 ARMY RSCH LAB – HRED
(PDF) RDRL HRM D
T DAVIS
BLDG 5400 RM C242
REDSTONE ARSENAL AL
35898-7290

1 ARMY RSCH LAB – HRED
(PDF) RDRL HRS EA DR V J RICE
BLDG 4011 RM 217
1750 GREELEY RD
FORT SAM HOUSTON TX
78234-5002

1 ARMY RSCH LAB – HRED
(PDF) RDRL HRM DG K GUNN
BLDG 333
PICATINNY ARSENAL NJ
07806-5000

1 ARMY RSCH LAB – HRED
(PDF) ARMC FIELD ELEMENT
RDRL HRM CH C BURNS
THIRD AVE
BLDG 1467B RM 336
FORT KNOX KY 40121

1 ARMY RSCH LAB – HRED
(PDF) AWC FIELD ELEMENT
RDRL HRM DJ D DURBIN
BLDG 4506 (DCD) RM 107
FORT RUCKER AL 36362-5000

1 ARMY RSCH LAB – HRED
(PDF) RDRL HRM CK J REINHART
10125 KINGMAN RD BLDG 317
FORT BELVOIR VA 22060-5828

1 ARMY RSCH LAB – HRED
(PDF) RDRL HRM AY M BARNES
2520 HEALY AVE
STE 1172 BLDG 51005
FORT HUACHUCA AZ
85613-7069

1 ARMY RSCH LAB – HRED
(PDF) RDRL HRM AP D UNGVARSKY
POPE HALL BLDG 470
BCBL 806 HARRISON DR
FORT LEAVENWORTH KS
66027-2302

1 ARMY RSCH LAB – HRED
(PDF) RDRL HRM AR J CHEN
12423 RESEARCH PKWY
ORLANDO FL 32826-3276

1 ARMY RSCH LAB – HRED
(PDF) HUMAN SYSTEMS
INTEGRATION ENGR
TACOM FIELD ELEMENT
RDRL HRM CU P MUNYA
6501 E 11 MILE RD
MS 284 BLDG 200A
WARREN MI 48397-5000

1 ARMY RSCH LAB – HRED
(PDF) FIRES CTR OF EXCELLENCE
FIELD ELEMENT
RDRL HRM AF C HERNANDEZ
3040 NW AUSTIN RD RM 221
FORT SILL OK 73503-9043

1 ARMY RSCH LAB – HRED
(PDF) RDRL HRM AV W CULBERTSON
91012 STATION AVE
FORT HOOD TX 76544-5073

1 ARMY RSCH LAB – HRED
(PDF) RDRL HRM DE A MARES
1733 PLEASONTON RD BOX 3
FORT BLISS TX 79916-6816

8 (PDF) ARMY RSCH LAB – HRED
SIMULATION & TRAINING
TECHNOLOGY CENTER
RDRL HRT COL G LAASE
RDRL HRT I MARTINEZ
RDRL HRT T R SOTTILARE
RDRL HRT B N FINKELSTEIN
RDRL HRT G A RODRIGUEZ
RDRL HRT I J HART
RDRL HRT M C METEVIER
RDRL HRT S B PETTIT
12423 RESEARCH PARKWAY
ORLANDO FL 32826

1 (PDF) ARMY RSCH LAB – HRED
HQ USASOC
RDRL HRM CN R SPENCER
BLDG E2929 DESERT STORM DR
FORT BRAGG NC 28310

1 (PDF) ARMY G1
DAPE MR B KNAPP
300 ARMY PENTAGON
RM 2C489
WASHINGTON DC 20310-0300

ABERDEEN PROVING GROUND

12 (PDF) DIR USARL
RDRL HR
L ALLENDER
P FRANASZCZUK
K MCDOWELL
RDRL HRM
P SAVAGE-KNEPSHIELD
RDRL HRM AL
C PAULILLO
RDRL HRM B
J GRYNOVICKI
RDRL HRM C
L GARRETT
RDRL HRS
J LOCKETT
RDRL HRS B
M LAFIANDRA
RDRL HRS D
A SCHARINE
RDRL HRS E
D HEADLEY
E ZAROUKIAN