HOKKAIDO UNIVERSITY

Title	Bacteria Lingualis in The Knowledge Soup : A Webcrawler with Affect Recognizer for Acquiring Artificial Empathy
Author(s)	Rzepka, Rafal; Komuda, Radoslaw; Araki, Kenji
Citation	Biologically Inspired Cognitive Architectures, Papers from the 2009 AAAI Fall Symposium, Arlington, Virginia, USA, November 5-7, 2009
Issue Date	2009-10-31
Doc URL	https://hdl.handle.net/2115/63627
Type	conference paper
File Information	1306.pdf

Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP

Biologically Inspired Cognitive Architectures II:
Papers from the AAAI Fall Symposium (FS-09-01)

Bacteria Lingualis in The Knowledge Soup – A Webcrawler with Affect Recognizer for Acquiring Artificial Empathy

Rafal Rzepka
Graduate School of
Information Science and Technology
Hokkaido University, Sapporo, Japan
kabura@media.eng.hokudai.ac.jp

Radoslaw Komuda
Faculty of Theology
Nicolaus Copernicus University
Torun, Poland
komuda@stud.umk.pl

Kenji Araki
Graduate School of
Information Science and Technology
Hokkaido University, Sapporo, Japan
araki@media.eng.hokudai.ac.jp

Abstract

In this paper we describe our vision on how to create a system that acquires moral rules by learning affective meaning of human experiences written on the web.

Introduction

Our research on developing what Yudkowsky (2007) named *Friendly AI* is mostly based on the vast textual WWW resources and on a presumption that most people follow the Ethic of Reciprocity (Golden Rule in Ethics). Our webcrawler for collecting human experiences is equipped with an affect recognizer that allows automatic emotion categorization of text. This can lead to creation of an "internal advisor" for various systems where human feedback is needed to make a decision but there is no supervisor around. Children learn by drawing morals from a parents' or other people's stories, however we need machines to search for them and interpret them automatically to make such a process efficient. As machines are still not able to fully perceive world with its five senses, third person's experiences become crucial source of knowledge. We assume that to intepret them correctly the *morals* must be drawn by understanding affective meaning of human acts.

Toward Moral Rules Acquisition

Inductive learning approach is quite efficient for the knowledge acquisition task, however ambitious "artificial infant" projects disappear without striking success after showing a promising start. We assume it is because with limited (or none) sensory input and only real supervisors - the learning process soon meets its limits. For this reason we decided to automatize the supervising process by using Bacteria Lingualis self-learning abilities briefly described in the next section. For testing this idea (and acquiring an empathic machine) we chose moral rules acquisition as ethics is a field when *common sense knowledge* collides with basic *emotions* - two objects of Bacteria's processing.

Bacteria Lingualis Webcrawler

Our idea of an imaginary beneficial bacteria living in the world of language was introduced in (Rzepka & Araki 2003). Its main purpose is to acquire bits of knowledge by using Japanese language particles as enzymes and to evaluate these bits in terms of 10 kinds of emotions. There are four parts of our "emotional webclawler": **Flagellum**, which is a motor for traversing the "Soup of Knowledge" - WWW in our case, **Affect Receptors** which calculate emotional categories of every written experience[1], **Concrete and Abstract Knowledge Memory** which allows to store not only the specific examples but also the abstracted conclusions and **Good Feeling Pursue Cell** which plays a role of a *utilitarian instinct* while applying the system and its knowledge to other applications.

Conclusions and Future Work

In this abstract we introduced our idea to combine our previous methods to facilitate the task of developing an empathic machine. We follow a presumption that WWW provides enough commonsense facts which (after affective categorization) will lead to correct ethical reasoning. Our presentation will show results of preliminary experiments which we wish to discuss with researchers developing human-level learners.

References

Yudkowsky, E. 2007. Artificial Intelligence as a Positive and Negative Factor in Global Risk In Nick Bostrom and Milan Cirkovic (eds.), *Global Catastrophic Risks* Oxford University Press

Rzepka, R., and Araki, K. 2003. Bacterium lingualis - the web-based commonsensical knowledge discovery method. In *6th International Conference on Discovery Science (DS 2003)*, volume 2843 of *Lecture Notes in Artifical Intelligence*, 453–460. Berlin, Heidelberg, New York: Springer-Verlag.

Ptaszynski, M., Dybala, P., Shi, W., Rzepka, R.; and Araki, K. 2009. Towards Context Aware Emotional Intelligence in Machines: Computing Contextual Appropriateness of Affective States In *Proceedings of IJCAI 2009 Conference, Pasadena, USA*, 1469-1474.

[1]We use 10 categories: joy / delight, anger, sorrow / sadness, fear, shame / shyness / bashfulness, liking / fondness, dislike / detestation, excitement, relief and surprise / amazement; details in (Ptaszynski et al. 2009).

Copyright © 2009, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.