# Consciousness of Crowds – The Internet as Knowledge Source of Human Conscious Behavior and Machine Self-Understanding

**Rafal Rzepka** and **Kenji Araki**
Language Media Laboratory
Graduate School of Information Science and Technology
Hokkaido University
Kita-ku Kita 14 Nishi 9, 060-0814 Sapporo, Japan
kabura,araki@media.eng.hokudai.ac.jp

### Abstract

This position paper introduces our approach to the artificial consciousness problem, approach which we developed during our experiments with automatic common sense retrieval and affective computing for open-domain talking systems. The data about the common behaviors and usual feelings is taken from the Internet, not from a particular programmer or user. Behavior patterns are retrieved based on Schankian *scripts*, usual emotional consequences of particular actions are calculated for every possible outcome, and the most positive action is chosen to be performed after a confirmation. The system is not only learning the conscious behavior, it can also "know" what it is or not capable of by using simple text-mining.

## Introduction

If our assumption is correct and the self can be simplified to an existence being a referee during an eternal match between inborn instincts and socially-learned common sense, then at least affective computing must be combined with common sense processing to understand or simulate (which we do not think is needed) the human consciousness.
From the very beginning of our trials with open-domain conversational systems we wanted to make programs we cannot fully control. First, driven by willingness to avoid manual creation of a database for the system's artificial personality, we decided to develop an algorithm calculating average personality features of Japanese Internet users. If most of users usually wore jeans and t-shirts, our "Mr. Internet" also "wore jeans" and "he" also liked beer because most of Japanese did. It was only a trick but it helped us noticing one thing - users were more interested in "simple-minded" (but "minded") systems than tricky ones which "pretend" to be intelligent or these just being passive (Rzepka 2005). We believe that the idea of "average thinking" could bring hints or even whole methods for creating a new level of Artificial Intelligence. Next we will introduce our methods and state our suggestions on possible applications that deal with conscious actions.



## Methods

Here we introduce very briefly our research and the ideas behind them together with some technical solutions although more specific descriptions are available in particular cited papers.

### GENTA Project

Basic functions of GENTA are described below.

**Positiveness and Artificial Experience** A machine needs to be able to tell good from bad. We divided an emotional reaction into 5 levels - negative, slightly negative, neutral, slightly positive and positive which are calculated by searching for words, phrases or clauses with their neighbors containing "emotional indicators". This is how we mimic the effects of Human emotional experiences gathering process as we think it is important for ones development of the "self".

**Usualness As Basic Common Sense Measure** A machine also needs to be able to tell correct from wrong. We assumed that if some combination of words or phrases does not exist on the Internet it can be treated as abnormality (n-grams used). We assume that if something is usual, normally the system does not react or reacts the same way it usually does. Hypothesis saying that the majority is always correct is controversial but output is usually very natural. Knowing what is uncommon gives us an idea of own originality which also seems to be a part of our conscious selves.

**Felific Calculus** This is an extension of Positiveness calculation that we currently work on. By calculating intensity, duration, certainty or uncertainty, propinquity or remoteness, fecundity, purity and extent, we can achieve quite sophisticated algorithm for measuring the emotional load of actions and their consequences. As we mentioned above, we assumed that every human's deed or word has a hedonistic backgrounds - that we always pursue a Good Feeling most often without realizing it. We based our system on classic hypothesis (Bentham 1789) saying that whatever we do, we do for feeling better. But we do not implement this "egoistic ego" into a computer, we set it as the fixed knowledge about a human user. If Positiveness becomes low or anything threaten the high Positiveness value - adequate action should be performed. If a user says he/she is cold, the agent has to calculate if it is negative state for its "master", and if

so, check what kind of actions are usually performed. Then after confronting the action possibilities with own functions and environment - the best action is proposed. By "own functions" we mean build-in functions - they can be inputted beforehand or learned by physical interaction.

### Bacterium Lingualis

For retrieving more complicated language structures and dependencies we use our multi-functional web crawler called Bacterium Lingualis (Rzepka & Araki 2003). Its three most important functions are as follows.

**Exception Processing** As the context is the clue for most failures of universal rules we have to find as many eventual exceptions of behaviors, opinions and rules, as possible. Even very simple Positiveness measures may be not true depending on who is talking about the measured matter. For most people sweets are pleasant but for a mother of a little baby it can be nightmare ruining the offspring's teeth.

**Schankian Scripts Retrieval** We use classic ideas (Schank & Abelson 1997) to retrieve and calculate common behavior patterns which combined with Positiveness calculation give the system information about what consequences may a given action bring. For example neutral word "escaping" becomes negative while inside of "robbery script" or positive in "picnic script".

**Causal Rules Retrieval** This works on the same basis as Scripts Retrieval but uses several Japanese "if" forms which have abilities to categorize causal dependencies. In this case Usualness of single happenings becomes more important - if a Script cannot be created, it can be made from single causalities generalized semantically by using a thesaurus. If there is not enough data for "quarreling with professor" maybe "quarreling with teacher" could bring possible consequence patterns.

## "I Am Only a Computer"

In our case computer's "self" can be set as "average man" (the idea of *consciousness of crowds*) or "average computer". The latter uses "a computer" as a keyword to find what is possible for computers and what is not. The method is very naive but based on, in our opinion, very fair idea - especially for a conversational agent. When a user invites an agent for a coffee, a computer can easily guess that "its kind" is not capable of such task. There is a big amount of noise because of on-line science-fiction stories but we think it can be to some extent eliminated by statistical methods.

## Possible Implementations

We have already suggested that a machine should be "conscious" enough to forecast danger for us (Rzepka & Araki 2005). There are many ways to use systems that can analyze a given situation, guess the previous actions, possible plans and goals, then estimate the consequences and how a user will feel about it. From very simple house devices (adjective "cold" in "home" context brings "possible flu" alert and "close the window" as the best action proposal) to sophisticated conversational agents which we are interested in fascinated by the need of universality in this broad topic.

## Questions Left To Be Answered

Is such "limited motivation" a part of "consciousness"? We do not hope for many positive answers for this question. Is there a chance for any freedom in such a "somebody else's" will? Majority would not agree. But is our will completely free? Aren't we slaves of the never-ending pursue for own satisfaction and common sense which prevents us taking shortcuts on our way to this satisfaction? The main message we want to pass to the artificial consciousness community is that most people behave very similar way even if every single one has his/her own consciousness. It is probably safer to have a machine which keeps to behavior standards and the manufacturers would not be afraid to implement such algorithms. Are we suggesting that consciousness is not needed? No, we are suggesting that "common consciousness" and "average self" could be an option which becomes easier and easier to implement along with development of large scale text-mining and affection analysis. We also suggest that Turing test for consciousness might work exactly the same way as the one for intelligence - isn't it enough if a human believes that its partner is has consciousness? Because we have never had an opportunity to confront our ideas with consciousness specialists, by our paper we would also like to trigger the discussion on new possibilities brought by massive data sets helpful for analyzing *what it is like* or *what it means* to have a consciousness. Does knowledge of action's meaning creates a conscious being? Purely functionalistic approach is tempting but "automatic pursue of (user's) good feeling" hypothesis as main engine for all behaviors brings possibilities for new interpretations of intentionality, rationality or phenomenology. Can one be conscious for others, not for the self, without freedom?

## References


Rzepka, R.; Ge, Y. and Araki, K. 2005. Naturalness of an Utterance Based on the Automatically Retrieved Commonsense. In *Proceedings of IJCAI 2005 - Nineteenth International Joint Conference on Artificial Intelligence, Edinburgh, Scotland*.

Bentham, J. 1789. *An Introduction to the Principles and Morals of Legislation*. London: T. Payne.

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.

Schank, R., and Abelson, R. 1977. *Scripts, Plans, Goals and Understanding: an Inquiry into Human Knowledge Structures*. Hillsdale, NJ: L. Erlbaum.

Rzepka, R.; and Araki, K. 2005. What Statistics Could Do for Ethics? - The Idea of Common Sense Processing Based Safety Valve. In *Machine Ethics, Papers from AAAI Fall Symposium, Technical Report FS-05-06*, 85–87.