# AI Matters

## Annotated Table of Contents

## Upcoming Conferences Sponsored by SIGAI

***Registration discount for SIGAI members.***

**WI 2016**: The 2016 IEEE/WIC/ACM International Conference on Web Intelligence, Omaha, Nebraska, USA.
Conference dates: October 13–16, 2016

**HRI 2017**: The 2017 Conference on Human-Robot Interaction, Vienna, Austria.
Conference dates: March 6–9, 2017

**IUI 2017**: The 22nd Annual Meeting of the Intelligent User Interfaces Community, Limassol, Cyprus.
Conference dates: March 13–16, 2017

**ICAIL 2017**: The 16th International Conference on Artificial Intelligence and Law, London, UK.
Conference dates: June 12–16, 2017

**ICEIS 2017**: The 19th International Conference on Enterprise Information Systems, Porto, Portugal.
Conference dates: April 26–29, 2017

## Links

SIGAI website: http://sigai.acm.org/
Newsletter: http://sigai.acm.org/aimatters/
Blog: http://sigai.acm.org/aimatters/blog
Twitter: http://twitter.com/acm_sigai/

Edition DOI: 10.1145/3008665

## Join SIGAI

Students $11, others $25
For details, see http://sigai.acm.org/
Benefits: regular, student

Also consider joining ACM.

Our mailing list is open to all.

## Submit to AI Matters!

We're accepting articles and announcements now for the Winter 2017 issue. Also, we would like to invite readers to submit responses to the *Who Speaks for AI?* article that appeared in AI Matters 2(2) to further the debate on this topic. We will include these opinions in a future issue of AI Matters.

Details on the submission process are available at http://sigai.acm.org/aimatters.

## AI Matters Editorial Board



Contact us: aimatters@sigai.acm.org

## Contents Legend

| | |
|---|---|
| B | Book Announcement |
| D | Ph.D. Dissertation Briefing |
| Ed | AI Education |
| E | Event Report |
| H | Humor |
| I | AI Impact |
| N | AI News |
| O | Opinion |
| P | Paper Précis |
| S | Spotlight |
| V | Video or Image |

Details at http://sigai.acm.org/aimatters

# Welcome to AI Matters, Volume 2, Issue 4

**Eric Eaton, Editor** (University of Pennsylvania; aimatters@sigai.acm.org)
**Amy McGovern, Co-Editor** (University of Oklahoma; aimatters@sigai.acm.org)
DOI: 10.1145/3008665.3008666

Your editors have been working hard on a variety of new features and you will see them roll out over the next few issues. Stay tuned as we bring you a wide variety of fun new columns, bringing you great news for the SIGAI community. In this issue, we are introducing three new features.

The first new feature is the introduction of an **AI Matters blog**. Here you can see featured articles (for free!) and can interact with the SIGAI community by commenting on the articles. The blog can be found at http://sigai.acm.org/aimatters/blog. It is a work in progress and we welcome your feedback!

The second new feature is a new column where we interview different AI researchers. We intend to interview people in all types of jobs including academia, industry, and government. Our new feature (and this issue) kicks off with an **interview of Peter Norvig**. His interview is also featured in the new blog at http://sigai.acm.org/aimatters/blog/2016/10/02/an-interview-with-peter-norvig/. Feel free to suggest new people to interview. Send us your ideas!

We're also introducing an **AI Education column** to *AI Matters*, which will be coordinated by Todd Neller. This first AI Education column presents an original card game called Birds of a Feather, and explores how this card game can be used for teaching various AI concepts.

The rest of the issue focuses on **abstracts of recent AI doctoral theses**, continuing this feature from our previous issue. This is the second of three sets of dissertation abstracts; the final set of abstracts will appear in the Autumn issue.

We close this issue with another round of the **AI Amusements** column, which brings you an epic poem (!) on Tay, the ill-fated chatbot. Keep sending in your original jokes, puzzles, or comics.



Thanks for reading! Don't forget to send your ideas and future submissions to *AI Matters*!

**Eric Eaton** is the Editor of AI Matters. He is a faculty member at the University of Pennsylvania in the Department of Computer and Information Science, and in the General Robotics, Automation, Sensing, and Perception (GRASP) lab. His research is in machine learning and AI, with applications to robotics, sustainability, and medicine.

**Amy McGovern** is the Co-Editor of AI Matters. She is an Associate Professor of computer science at the University of Oklahoma and an adjunct associate professor of meteorology. She directs the Interaction, Discovery, Exploration and Adaptation (IDEA) lab. Her research focuses on machine learning and data mining with applications to high-impact weather.

S

# AI Profiles: An interview with Peter Norvig

**Amy McGovern** (University of Oklahoma; amcgovern@ou.edu)
**Eric Eaton** (University of Pennsylvania; eeaton@cis.upenn.edu)
DOI: 10.1145/3008665.3008667

**Abstract**

This column is the first in a new series profiling senior AI researchers. This month focuses on Peter Norvig.

## Introduction

With this issue, AI Matters is introducing a new column profiling senior researchers in AI. We begin with Peter Norvig, who is the Director of Research at Google, Inc. We interviewed him virtually. The interview has been edited for clarity and length. We thank Peter for his time!

Peter Norvig

## Bio

Peter Norvig is a Director of Research at Google Inc. Previously he was head of Google's core search algorithms group, and of NASA Ames's Computational Sciences Division, making him NASA's senior computer scientist. He received the NASA Exceptional Achievement Award in 2001. He has taught at the University of Southern California and the University of California at Berkeley, from which he received a Ph.D. in 1986 and the distinguished alumni award in 2006. He was co-teacher of an Artificial Intelligence class that signed up 160,000 students, helping to kick off the current round of massive open online classes. His publications include the books *Artificial Intelligence: A Modern Approach* (the leading textbook in the field), *Paradigms of AI Programming: Case Studies in Common Lisp*, *Verbmobil: A Translation System for Face-to-Face Dialog*, and *Intelligent Help Systems for UNIX*. He is also the author of *The Gettysburg Powerpoint Presentation* and the world's longest palindromic sentence. He is a fellow of the AAAI, ACM, the California Academy of Science, and the American Academy of Arts & Sciences.



## Getting to know Peter Norvig

### How did you become interested in AI?

I was lucky enough to go to a High School that had access to a computer and a programming class, which was a rarity in 1974, and a Linguistics class; this got me interested in creating models of language. Forty-two years later, I still haven't figured it out, but I've had fun trying.

### What was your most difficult professional decision and why?

In 1998, I was offered the position of leading the Computational Sciences Division at NASA's Ames Research Center. This would mean changing my role to be a manager of a 200-person team, rather than contributing as an individual researcher/programmer. In the past I remember there had been many times when I had thought to myself "I could ask a co-worker to program this task, but it would be easier to just do it myself." But at NASA, and later at Google, the quality of the people was so high, that it was worth it to forego the "do it yourself" approach, and concentrate on getting everyone working together well. This required a different skill set, but in the end greatly amplified the overall impact, and therefore was worth it.

**What professional achievement are you most proud of?**

First, as a mostly personal effort, Stuart Russell and I (with help from others) were able to put together the textbook that has been the primary resource in AI for 20 years. It was gratifying to see our vision of the field embraced and to hear from so many students who enjoyed using it. Later I was able to team with Sebastian Thrun to bring the core ideas to a large group of online students.

Second, as a team effort, I was the manager for the core Google search team during a period of great growth from 2002 to 2006. I was proud that we were able to help billions of people with trillions of questions, through the combined brilliance of so many great team members.

**What do you wish you had known as a Ph.D. student or early researcher?**

When I finished grad school, there was an expectation that the "right" path was to stay in academia. In my second year of grad school, two of my most respected fellow students, Bill Joy and Eric Schmidt, left to start a company selling workstations. I remember thinking "Why would they do that? They could have been assistant professors at good schools!" It took me a while to realize that there are multiple paths: in academia, industry research, startups, government, and non-profits, and any one of them, or any combination of them, could be the right choice for you.

**What would you want for your career if you couldn't do AI?**

If I couldn't do AI, I suppose I would want to do AI all the more. But I probably would end up in a field that looks at the same problems from a different point of view, such as Linguistics or Statistics.

**What is a typical day like for you?**

I answered a similar question on Quora, and it still holds true. At Google there's always something new to work on; I can't really fall into a set routine. Within a project there are always changes of strategy as we learn more and the world changes. And from one year to the next my role has changed. I've varied from having two to two hundred people reporting to me, which means that sometimes I have very clear technical insight for every one of the projects I'm involved with, and sometimes I have a higher-level view, and I have to trust my teams to do the right thing. In those cases, my role is more one of communication and matchmaker: to try to explain which direction the company is going in and how a particular project fits in, and to introduce the project team to the right collaborators, producers, and consumers, but to let the team work out the details of how to reach their goals. I don't write code that ends up on Google, but if I have an idea, I can write code to experiment with the internal tools to see if the idea is worth looking at more carefully. And I do code reviews, both so that I can see more of the code that the teams are producing, and because somebody has to do it.

There is always a backlog of meetings, emails, and documents to get through. Google is less bureaucratic than anyplace else I've worked, but some of this is inevitable. I also spend some time going to conferences, talking with Universities and customers, and answering questions like these.

**What is your favorite AI-related joke?**

I don't have a good AI joke, but I did invent my own math joke:

"I saw a pair of mathematicians get into a terrible argument about a Möbius strip. It was one-sided."

**What is your favorite AI-related movie and why?**

I liked the movie *Her*, because the technology is both central to the plot, but mostly receded into the background of the society that is portrayed. When asked what movie *Her* reminded me of most, I said *Monty Python's Life* *of Brian*, because both movies are about the human capacity for faith – wanting to believe in something.

**How do you spend your free time?**

My hobbies are photography and bicycling. Photography is a good art form for me because it doesn't require that much hand-eye coordination. It is all about simplification and subtraction rather than addition and it allows me to think about gadgets and technical equations (as in Marc Levoy's *Lectures on Digital Photography*). Bicycling is right for me because it is just the right speed to see the scenery: with walking you don't get far enough to see much, and in touring by car you go too fast to see much.

**What is a skill you would like to learn and why?**

I've tried a couple of times to play music, but so far I'm better as an avid listener.

# Ed AI Education: Birds of a Feather

**Todd W. Neller** (Gettysburg College; tneller@gettysburg.edu)
DOI: 10.1145/3008665.3008668

## Introduction

> "[Play] is our brain's favorite way of learning..." (Ackerman, 1999, p.11)

Games are beautifully crafted microworlds that invite players to explore complex terrains that spring into existence from even simple rules. As AI educators, games can offer fun ways of teaching important concepts and techniques. Just as Martin Gardner employed games and puzzles to engage both amateurs and professionals in the pursuit of Mathematics, a well-chosen game or puzzle can provide a catalyst for AI learning and research.

FreeCell stands out among solitaire card games because it is essentially a random self-generating puzzle that has perfect information and can be solved with high probability. Players over the years have, as a community, researched many aspects of the game (Keller, 2016). In this column, we present a new card solitaire game called Birds of a Feather that is virgin territory for exploration in the hopes that motivated undergraduates and their advisors will enjoy investigating this new challenge.

## Birds of a Feather Rules

Birds of a Feather is an original perfect-information solitaire game played with a standard 52-card deck. After shuffling, the player deals the cards face-up left-to-right in $c$ columns, and top-to-bottom in $r$ rows to create an $r$-by-$c$ grid of cards. An example $4$-by-$4$ game's initial layout:

```
5S JC QH 8H
KC 6H 3H 9H
3S JS TH TS
KS 7D AH 5C
```

Think of each grid cell as initially containing a 1-card stack. A stack may be moved on top of another stack in the same row, or in the same column if at least one of two conditions is met: (1) The top card of each stack has the same suit. (2) The top card of each stack has the same rank or an adjacent rank (with Aces low and Kings high and Ace and King non-adjacent). Thus the 9H (9 of Hearts) stack can move onto the TS (Ten of Spades) being adjacent/same in rank:

```
5S JC QH 8H
KC 6H 3H
3S JS TH 9H
KS 7D AH 5C
```

And the 8H stack can move onto the 9H stack being both of (1) same suit and (2) same/adjacent rank:

```
5S JC QH
KC 6H 3H
3S JS TH 8H
KS 7D AH 5C
```

And the TH stack can move onto the AH stack being of the same suit:

```
5S JC QH
KC 6H 3H
3S JS    8H
KS 7D TH 5C
```

If we notate each move as the top cards of the moving and destination stacks separated by a hyphen, then this entire tableau can be formed into a single stack from this sequence of moves: 9H-TS 8H-9H TH-AH 3H-TH QH-3H 6H-7D JC-JS 3S-KS 5S-3S 5C-5S KC-5C QH-KC QH-6H QH-JC QH-8H.

Let us call this simple solution concept a "single-stack solution". However, we can define a more general solution concept of forming largest stacks by defining the score of a grid to be the sum of the squares of the stack sizes. The general solution of any grid is a sequence of moves that maximizes this grid score.

## Birds of a Feather Questions

Having defined the puzzle, we can now ask many interesting questions about it. For $r$

rows and $c$ columns,

- What is the probability that a deal will have a single-stack solution?
- What is the maximal score distribution of deals?
- What are heuristics that can be used to guide search more efficiently to solutions?
- What are characteristics of grids without single-stack solutions?

There are also many questions one can ask with regard to the automated design of Birds of a Feather puzzles:

- What are the most important attributes of challenging deals with single-stack solutions?
- How can such attributes best combine to form an objective function that can be used to generate Birds of a Feather puzzles through combinatorial optimization algorithms (e.g. simulated annealing (Neller, Fisher, Choga, Lalvani, & McCarty, 2011))?

Given this fresh ground for exploration, we would invite educators and students to explore these and other questions concerning Birds of a Feather, and we can summarize our results in a future column. To download relevant code and/or share your results, we invite you to register with and add to our wiki on the subject http://cs.gettysburg.edu/ai-matters/index.php/Birds_of_a_Feather.

## References

Ackerman, D. (1999). *Deep play*. New York: Random House.

Keller, M. (2016). *Freecell faq.* http://solitairelaboratory.com/fcfaq.html. (Accessed: 2016-09-30)

Neller, T., Fisher, A., Choga, M., Lalvani, S., & McCarty, K. (2011). Rook jumping maze design considerations. In H. van den Herik, H. Iida, & A. Plaat (Eds.), *LNCS 6515: Computers and Games 2010* (pp. 188–198). Springer. Retrieved from http://cs.gettysburg.edu/~tneller/papers/iccg10.pdf

**Todd W. Neller** is a Professor of Computer Science at Gettysburg College. A game enthusiast, Neller researches AI techniques as they apply to the design and analysis of games and puzzles, and uses such games and puzzles to create fun experiential learning and research opportunities for undergraduates.

# D Nonverbal Communication in Socially Assistive Human-Robot Interaction

**Henny Admoni** (Dept. of Computer Science, Yale University; henny.admoni@yale.edu)
DOI: 10.1145/3008665.3008669

Socially assistive robots provide assistance to human users through interactions that are inherently social. This category includes robot tutors that provide students with personalized one-on-one lessons (Ramachandran, Litoiu, & Scassellati, 2016), robot therapy assistants that help mediate social interactions between children with ASD and adult therapists (Scassellati, Admoni, & Matarić, 2012), and robot coaches that motivate children to make healthy eating choices (Short et al., 2014).

To successfully provide social assistance, these robots must understand people's beliefs, goals, and intentions, as communicated in the course of natural human-robot interactions. Human communication is multimodal, with verbal channels (i.e., speech) and nonverbal channels (e.g., eye gaze and gestures). Recognizing, understanding, and reasoning about multimodal human communication is an artificial intelligence challenge.

This dissertation focuses on **enabling human-robot communication by building models for understanding human nonverbal behavior and generating robot nonverbal behavior in socially assistive domains.** It investigates how to computationally model eye gaze and other nonverbal behaviors so that these behaviors can be used by socially assistive robots to improve human-robot collaboration.

Developing effective nonverbal communication for robots engages a number of disciplines across AI, including machine learning, computer vision, robotics, and cognitive modeling. This dissertation applies techniques from all of these disciplines, providing a greater understanding of the computational and human requirements for human-robot communication.

To focus nonverbal communication models on the features that most strongly influence human-robot interactions, I first conducted a series of studies that draw out human responses to specific robot nonverbal behaviors. These laboratory-based studies investigate how robot eye gaze compares to human eye gaze in eliciting reflexive attention shifts from human viewers (Admoni, Bank, Tan, & Toneva, 2011); how different features of robot gaze behavior promote the perception of a robot's attention toward a viewer (Admoni, Hayes, Feil-Seifer, Ullman, & Scassellati, 2013); whether people use robot eye gaze to support verbal object references and how they resolve conflicts in this multimodal communication (Admoni, Datsikas, & Scassellati, 2014); and what is the role of eye gaze and gesture in guiding behavior during human-robot collaboration (Admoni, Dragan, Srinivasa, & Scassellati, 2014).

Figure 1: Data from a human-human interaction was used to train a model for recognizing nonverbal communication.

Based on this understanding of nonverbal communication between people and robots, I develop two models for understanding and generating nonverbal behavior in human-robot interactions. The first model uses a data-driven approach (Admoni & Scassellati, 2014), trained on examples from human-human tutoring (Figure 1). This model can recognize the communicative intent of nonverbal behaviors, and suggest nonverbal behaviors to support a desired communication.

The second model takes a scene-based approach to generate nonverbal behavior for a

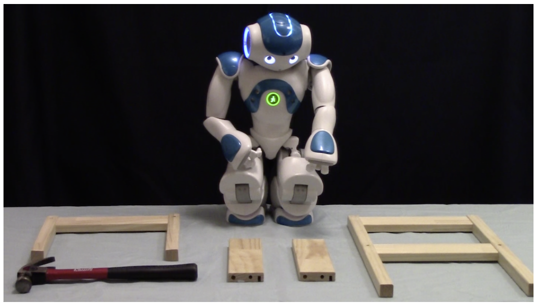

Figure 2: This dissertation includes a model for generating robot gaze and gestures for human-robot collaboration.

socially assistive robot (Figure 2) (Admoni, Weng, & Scassellati, 2016). This model is context independent and does not rely on a priori collection and annotation of human examples, as the first model does. Instead, it calculates how a user will perceive a visual scene from their own perspective based on cognitive psychology principles, and it then selects the best robot nonverbal behavior to direct the user's attention based on this predicted perception. The model can be flexibly applied to a range of scenes and a variety of robots with different physical capabilities. I show that this second model performs well in both a targeted evaluation and in a naturalistic human-robot collaborative interaction (Admoni, Weng, Hayes, & Scassellati, 2016).

## Acknowledgments

This work was supported by the National Science Foundation Graduate Research Fellowship, the Google Anita Borg Memorial Scholarship, and the Palantir Women in Technology Fellowship, as well as grant support from the National Science Foundation (0835767, 0968538, 113907), the Office of Naval Research (N00014-12-1-0822), the DARPA Computer Science Futures II program, Microsoft, and the Sloan Foundation.

## References

Admoni, H., Bank, C., Tan, J., & Toneva, M. (2011). Robot gaze does not reflexively cue human attention. In L. Carlson, C. Hölscher, & T. Shipley (Eds.), *CogSci* (pp. 1983–1988). Austin, TX USA: Cognitive Science Society.

Admoni, H., Datsikas, C., & Scassellati, B. (2014). Speech and gaze conflicts in collaborative human-robot interactions. In *CogSci* (pp. 104–109).

Admoni, H., Dragan, A., Srinivasa, S., & Scassellati, B. (2014). Deliberate delays during robot-to-human handovers improve compliance with gaze communication. In *ACM/IEEE HRI* (pp. 49–56).

Admoni, H., Hayes, B., Feil-Seifer, D., Ullman, D., & Scassellati, B. (2013). Are you looking at me? Perception of robot attention is mediated by gaze type and group size. In *ACM/IEEE HRI* (pp. 389–396).

Admoni, H., & Scassellati, B. (2014). Data-driven model of nonverbal behavior for socially assistive human-robot interactions. In *ICMI.* Istanbul, Turkey.

Admoni, H., Weng, T., Hayes, B., & Scassellati, B. (2016). Robot nonverbal behavior improves task performance in difficult collaborations. In *ACM/IEEE HRI.*

Admoni, H., Weng, T., & Scassellati, B. (2016). Modeling communicative behaviors for object references in human-robot interaction. In *IEEE ICRA.*

Ramachandran, A., Litoiu, A., & Scassellati, B. (2016). Shaping productive help-seeking behavior during robot-child tutoring interactions. In *ACM/IEEE HRI.*

Scassellati, B., Admoni, H., & Matarić, M. (2012). Robots for use in autism research. *Annual Review of Biomedical Engineering*, *14*, 275–294.

Short, E., Swift-Spong, K., Greczek, J., Ramachandran, A., Litoiu, A., Grigore, E. C., ... Scassellati, B. (2014, August). How to train your dragonbot: Socially assistive robots for teaching children about nutrition through play. In *IEEE RO-MAN.*

**Henny Admoni** is currently a postdoctoral fellow at the Robotics Institute at Carnegie Mellon University, where she works on assistive robotics and human-robot interaction with Siddhartha Srinivasa in the Personal Robotics Lab. Henny completed her PhD in Computer Science at Yale University with Professor Brian Scassellati. Henny develops intelligent robots that improve people's lives by providing assistance through social and physical interactions. Her scholarship has been recognized with awards such as the NSF Graduate Research Fellowship, the Google Anita Borg Memorial Scholarship, and the Palantir Women in Technology Scholarship.

# Abstract: Multi-Objective Decision-Theoretic Planning

**Diederik M. Roijers** (University of Oxford; diederik.roijers@cs.ox.ac.uk)
DOI: 10.1145/3008665.3008670

Decision making is hard. It often requires reasoning about uncertain environments, partial observability and action spaces that are too large to enumerate. In such tasks decision-theoretic agents can often assist us. In most research on decision-theoretic agents, the desirability of actions and their effects is codified in a scalar reward function. However, many real-world decision problems have multiple objectives. In such cases the problem is more naturally expressed using a vector-valued reward function, leading to a *multi-objective decision problem (MODP)*.

Typically, MODPs cannot be *scalarized* to a single-objective decision problem, at it is very hard to *a priori* specificy a so-called *scalarization function* that captures the user utility for every value-vector imaginable. Instead, we provide decision support (schematically depicted in Fig. 1). In the planning phase, our algorithm produces a *coverage set (CS)*, i.e., a set of policies that covers all possible preferences between the objectives. In the selection phase, the user selects one policy from the CS. Finally this selected policy is executed.

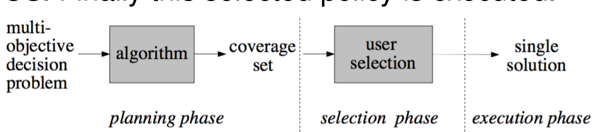


Figure 1: The decision support scenario.

We focus on decision-theoretic planning algorithms that produce a *convex coverage set (CCS)*, which is the optimal solution set when either: 1) the user utility can be expressed as a weighted sum over the values for each objective; or 2) policies can be stochastic.

We propose new methods based on two approaches to creating planning algorithms that produce an (approximate) CCS by building on existing single-objective algorithms. In the inner loop approach, we replace the summations and maximizations in the inner most loops of single-objective algorithms by cross-sums and pruning operations. In the outer loop approach, we solve a multi-objective problem as a series of scalarized, i.e., single-objective, problems.

One of our most important contributions is *optimistic linear support (OLS)* (Roijers, Whiteson, & Oliehoek, 2015a). OLS is a generic outer loop framework for multi-objective decision problems that uses single-objective solvers as subroutines. It can be applied to any MODP for which a corresponding single-objective method exists. We show that, contrary to existing methods, each intermediate result is a bounded approximation of the CCS with known bounds (even when the single-objective method used is a bounded approximate method as well) and is guaranteed to terminate in a finite number of iterations.

## Multi-Objective Coordination

The first MODP we tackle is *multi-objective coordination graphs (MO-CoGs)*. MO-CoGs are cooperative single-shot, fully observable, multi-agent decision problems. In MO-CoGs, agents must coordinate in order to find effective policies. Key to making coordination between agents efficient is exploiting loose couplings, i.e., each agents actions directly affect only a subset of the other agents. Such loose couplings are expressed by a (vector-valued) payoff function, that decomposes into a sum over local payoff functions in which only subsets of the agents participate.

We propose and compare inner loop methods and OLS-based methods. Specifically, we build upon *variable elimination (VE)* (Guestrin, Koller, & Parr, 2002) and propose *convex multi-objective variable elimination (CMOVE)* (inner loop) and *variable elimination linear support (VELS)* (OLS-based). We build on *AND/OR tree search* (Mateescu & Dechter, 2005) to propose *convex AND/OR tree search (CTS)* (inner loop) and *AND/OR tree search linear support (TSLS)* (OLS-based). We show that OLS-based methods scale better in the number of agents, both in terms of runtime and memory, while inner loop methods scale better in the number of objectives. We show experimentally that we can produce $\varepsilon$-CCSs in a fraction of the runtime that is required to produce an exact CCS.

Furthermore, we propose *variational optimistic linear support (VOLS)* (Roijers, Whiteson, Ihler, & Oliehoek, 2015) — an OLS-based method that builds on variational methods. The runtime of variational subroutines (Liu & Ihler, 2011) is not exponential in the number of agents. However, they produce only $\varepsilon$-approximate solutions. VOLS inherits the runtime and quality guarantees and can produce an $\varepsilon$-CCSs in sub-exponential runtime. We show that it is possible to reuse the reparameterized graphs produced by single-objective variational subroutines to hot-start the variational subroutines in later iterations of OLS, leading to significant speed-ups.

## Sequential Planning

The next problem settings we tackle are *multi-objective Markov decision processes (MOMDPs)* and *multi-objective partially observable Markov decision processes (MOPOMDPs)* which are single-agent sequential decision problems. Because the sequence of actions that result from executing policies in these problems affect the environment, agents have to consider both immediate and future rewards that depend on the future state of the environment.

MOMDPs are fully-observable, i.e., the agent knows at any time what the exact state of the environment is. A major challenge in MOMDPs is the size of the state and action spaces. We illustrate, using a large MOMDP called the *maintenance planning problem* (Roijers et al., 2014), that it is possible to create efficient methods using OLS and specialised single-objective subroutines, and that it is relatively easy to replace these subroutines when the state-of-the-art for the single-objective method improves.

MOPOMDPs are partially observable, which poses an important additional challenge. We propose *optimistic linear support with alpha reuse (OLSAR)* (Roijers, Whiteson, & Oliehoek, 2015b), which as far was we are aware, this is the first MOPOMDP planning method that computes the CCS and reasonably scales in the number of states of the MOPOMDP. We show how to represent the value function of MOPOMDPs in terms of $\alpha$-matrices and propose a single-objective subroutine for OLSAR called OLS-compliant Perseus (based on (Spaan & Vlassis, 2005)) that returns these $\alpha$-matrices. A key insight underlying OLSAR is that the $\alpha$-matrices produced by OCPerseus can be reused in subsequent calls to OCPerseus, greatly reducing the runtime. Our experimental results show that OLSAR greatly outperforms alternatives that do not use OLS and/or $\alpha$-matrix reuse.

## References


Guestrin, C., Koller, D., & Parr, R. (2002). Multiagent planning with factored MDPs. In *NIPS.*

Liu, Q., & Ihler, A. T. (2011). Bounding the partition function using Hölder's inequality. In *ICML* (pp. 849–856).

Mateescu, R., & Dechter, R. (2005). The relationship between AND/OR search and variable elimination. *UAI* (pp. 380–387).

Roijers, D., Scharpff, J., Spaan, M., Oliehoek, F., de Weerdt, M., & Whiteson, S. (2014). Bounded approximations for linear multi-objective planning under uncertainty. In *ICAPS* (pp. 262–270).

Roijers, D. M., Whiteson, S., Ihler, A. T., & Oliehoek, F. A. (2015). Variational multi-objective coordination. In *Malic.*

Roijers, D. M., Whiteson, S., & Oliehoek, F. A. (2015a). Computing convex coverage sets for faster multi-objective coordination. *JAIR*, *52*, 399–443.

Roijers, D. M., Whiteson, S., & Oliehoek, F. A. (2015b). Point-based planning for multi-objective POMDPs. In *IJCAI.*

Spaan, M. T. J., & Vlassis, N. (2005). Perseus: Randomized point-based value iteration for POMDPs. *JAIR*, 195–220.


**Diederik M. Roijers** did his PhD on multi-objective decision theory at the University of Amsterdam under the supervision of Shimon Whiteson and Frans A. Oliehoek. He is currently a postdoctoral researcher at the Dep. of Computer Science at the University of Oxford.

# D Automatic Extraction of Future References from News Using Morphosemantic Patterns with Application to Future Trend Prediction

**Yoko Nakajima** (National Institute of Technology, Kushiro College; yoko@kushiro-ct.ac.jp)
**Michal Ptaszynski** (Kitami Institute of Technology; ptaszynski@cs.kitami-it.ac.jp)
**Hirotoshi Honma** (National Institute ofTechnology,Kushiro College; honma@kushiro-ct.ac.jp)
**Fumito Masui** (Kitami Institute of Technology; f-masui@mail.kitami-it.ac.jp)
DOI: 10.1145/3008665.3008671

## Introduction

In everyday life people use past events and their own knowledge to predict future events. In such everyday predictions people use widely available resources (newspapers, Internet). This study focused on sentences referring to the future, such as the one below, as one of such resource.

*Science and Technology Agency, the Ministry of International Trade and Industry, and Agency of Natural Resources and Energy conferred on the necessity of a new energy system, and decided to set up a new council.* (Japanese daily newspaper *Hokkaido Shinbun*, translation by the author.)

The sentence claims that the country will construct a new energy system. However, although the sentence is set in the past ("conferred", "decided") the sentence itself refers to future events ("setting up a new council"). Such references to the future contain information (expressions, causal relations) relating it to the specific event that may happen in the future. The prediction of the event depends on the ability to recognize this information.

A number of studies have been conducted on the prediction of future events with the use of time expressions (Baeza-Yates 2005; Kanazawa et al. 2010), SVM (bag-of-words) (Aramaki et al. 2011), causal reasoning with ontologies (Radinsky et al. 2012), or keyword-based linguistic cues ("will", "shall", etc.) (Jatowt et al. 2013). In this research I assumed that future references in sentences occur not only on the level of surface (time expressions, words) or grammar, but consist of a variety of patterns both morphological and semantic.

## Future Reference Pattern Extraction

The proposed method consists of two stages: (1) sentences are represented in a morphosemantic structure (Levin and Rappaport Hovav 1998) (combination of semantic role labeling with morphological information), and (2) frequent morphosemantic patterns (MoPs) are automatically extracted from training data and used in classification. MoPs are useful for representing languages rich both morphologically and semantically, such as Japanese (language of datasets used in this research). Morphosemantic model was generated using semantic role labeling (SRL) supported with morphological information. SRL provides labels for words and phrases according to their role in the sentence. To retain information omitted by SRL (particles, function words, not directly influencing the semantic structure, but contributing to the overall meaning) morphological analysis provided information on parts of speech, of omitted words. Below is an example of a sentence generalized on the morphosemantic structure:

**Japanese:** *AI gijutsu ha mujinhikōki nado no seihin ya sābisu ni katsuyō ga kitaisarete iru.* **English:** AI technology is expected to be used in products and the services such as a pilotless planes. **MoPs:** `[Object][Noun][Thing]-[Agent][Object][No_State_change]`

From sentences represented this way frequent MoPs are extracted as follows. Firstly, ordered non-repeated combinations from all sentence elements are generated. In every $n$-element sentence there is $k$-number of combination groups, such as that $1 \leq k \leq n$. All combinations for all values of $k$ are generated, with non-subsequent elements separated by an asterisk. Frequent pattern lists extracted this way from training set are used in classification of test and validation set.

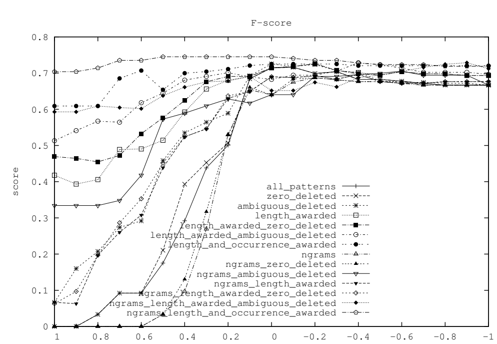


Figure 1: F-score for all tested classifier versions.

Table 1: Comparison of results for different pattern groups, state-of-the-art, and fully optimized model.

| Pattern set | Precision | Recall | F-score |
|---|---|---|---|
| 10 patterns | 0.39 | **0.49** | **0.43** |
| 10 pattern (3 elements or longer) | **0.42** | 0.37 | 0.40 |
| 5 patterns | 0.35 | 0.35 | 0.35 |
| Optimized (see Fig. 2) | **0.76** | **0.76** | **0.76** |
| (Jatowt et al. 2013) (10 phrases) | **0.50** | 0.05 | 0.10 |

## Evaluation

From three newspaper corpora (*Nihon Keizai Shimbun*, *Asahi Shimbun*, *Hokkaido Shimbun*.) two datasets were collected and manually annotated to contain equal number of (1) sentences referring to future events and (2) other (describing past, or present events).

The datasets were applied in a text classification. Each classified test sentence was given a score calculated as a sum of weights of patterns extracted from training data and matched with the input sentence. The results were calculated with Precision (P), Recall (R) and F-measure (F). Fourteen classifier versions were compared (see Figure 1) for performance based on the highest statistically significant F within the threshold, and the highest break-even point (BEP) of P and R. The highest overall performance was obtained by the version using pattern list containing all patterns (including ambiguous patterns and n-grams).

In comparison with (Jatowt et al. 2013), who extracted future reference sentences (FRS) with 10 words unambiguously referring to the future, such as "will" or "is likely to", etc. the proposed method obtained much higher results even when only 10 most frequent MoPs were used (Table 1). Moreover, the performance of a fully optimized model, retrained on all training data with the best settings reached break-even point (BEP) at 76% (Figure 2).

## Future Prediction Support Experiment

We performed an experiment to verify that our method is useful for the support of predicting future trends. Thirty laypeople answered questions from Future Prediction Competence Test (*http://homepage3.nifty.com/genseki/kentei.html#kentei*) using only FRS provided by our method. The FRS for each question were gathered from Mainichi

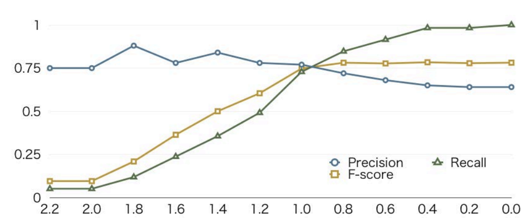


Figure 2: Final results of fully optimized model.

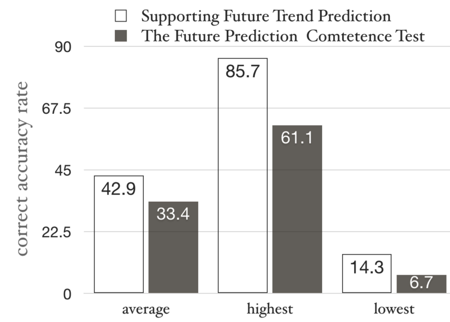


Figure 3: Correct accuracy rate in future trend prediction experiment compared to the original Future Prediction Competence Test.

Newspaper using: (1) topic keywords related to questions and (2) MoPs generated using the fully optimized model.

The correct accuracy rate (Figure 3) of the proposed method was higher than for the original test participants both in average, and for the highest and lowest score achieved. Only 9% higher but, FRS is clearly usefulness supporting to future trend prediction.

## Conclusion and Future Directions

We proposed a novel method for extracting references to future events from news articles, based on automatically extracted morphosemantic patterns. From 14 different classifier version compared an optimized model was selected and validated on a new data set. The model achieved high performance outperforming state-of-the-art. Moreover, we performed a future trend prediction experiment and found out that the method is capable to automatically extract sentences providing support for future event prediction. As for further work, we consider applying the method in statistical data interpretation, and key sentence extraction from the Web for supporting business-related judgments.

## References

E. Aramaki, S. Maskawa, M. Morita. Twitter catches the flu: Detecting influenza epidemics using twitter. *EMNLP*, pp. 1568–1576, 2011.

R. Baeza-Yates. Searching the Future, 2005. *SIGIR Workshop on MF/IR*.

A.Jatowt, H.Kawai, K.Kanazawa, K.Tanaka, K.Kunieda, K.Yamada. Multilingual, Longitudinal Analysis of Future-related Information on the Web. *Cult. and Comp. 2013*.

K. Kanazawa, A. Jatowt, S. Oyama, K. Tanaka. Exracting Explicit and Implicit future-related information from the Web(O) (in Japanese). *DEIM Forum 2010*.

B. Levin, M. Rappaport Hovav. *Morphology and Lexical Semantics*, pp. 248-271, 1998.

K. Radinsky, S. Davidovich, S. Markovitch. Learning causality for news events prediction. *WWW 2012*.

**Yoko Nakajima** received her PhD from Kitami Institute of Technology in 2016. Currently works as an Assistant Professor at the Information Engineering Course, of NIT, Kushiro College. Her research interests include natural language processing and information extraction. She is a member of IPSJ.

**Michal Ptaszynski** received PhD in Information Science and Technology from Hokkaido University in 2010. JSPS PD Research Fellow at Hokkai-Gakuen University (2010-2012). Since 2013 an Assistant Professor at Kitami Institute of Technology. Research interests: NLP, affect analysis. Member of: ACL, AAAI, IEEE, IPSJ, ANLP.

**Hirotoshi Honma** received B.E., M.E. and D.E. degrees in Engineering from Toyohashi University of Technology, in 1990, 1992 and 2009, respectively. Joined the National Institute of Technology, Kushiro College in 1992, Associate Professor since 2001. Research interest: computational graph theory, parallel algorithms, and NLP. Member of: IEICE and ORSJ.

**Fumito Masui** graduated Okayama University Faculty of Science in 1990. Since 2009 an Associate Professor at Faculty of Engineering, Kitami Institute of Technology. Ph.D. in Engineering. Research interests: NLP, tourism informatics, and curling informatics. Member of: JSAI, IEICE, IPSJ, SOFT, Hokkaido Regional Tourism Society, AAAI, ACL, and Japan Curling Association.

D

# Network Organization Paradigm

**Saad Alqithami** (Southern Illinois University; alqithami@gmail.com)

DOI: 10.1145/3008665.3008672

## Introduction

Fervent communication on social networking sites provides opportunities for potential issues to trigger individuals into individual action as well as the attraction and mobilization of like-minded individuals into an organization that is both physically and virtually emergent. Examples are the rapid pace of Arab Spring and the diffusion rate of the Occupy Movement. We intend to view this as a complex adaptive system where diverse agents perform various actions without adherence to a predefined structure. The achievement of joint actions will be a result of continual interactions between them that shape a dynamic network. Agents may form an ad hoc organization based on a dynamic network of interactions for the purpose of achieving a long-term objective, which we termed as a *Network Organization* (NO).

An NO is introduced to present large, semi-autonomous, ad-hoc networked individual entities that aim to automate comment and control of distributed complex objectives. We introduced a paradigm that serves as a reference model for organizations of networked individuals. This paradigm suggests modular components capturing essential units that define an ad hoc NO when they are modularly combined. We touch on how this model accounts for external change in an environment through internal adjustment. Furthermore, due to the predominant influences of the network substrate in an NO, multiple effects of it have an impact on the NO behaviors and directions. We envisioned several dimensions of such effects to include synergy, social capital, externality, influence, etc. A special focus is on measuring *synergy* and *social capital* as two prevalent network effects.

## An NO Paradigm

Given volatility of networks, an NO paradigm allows for rapid depiction and analysis of an emerging and evolving NO witnessed in our connected world. It encapsulates representational power of a more ubiquitous perspective over its modifier by providing guidelines, a reference model, and a set of principles. In short, the paradigm is structured along five profiles:

- *Network Profile*: The dynamic network will provide an NO with a set of existing agents and available resources along with initial protocols.
- *Agent Profile*: Every agent has a profile that consists of her allegiance to an NO, skills to perform tasks, relationships with others (inside or outside an NO), a set of preferences as well as autonomy-levels toward different activities.
- *Problem Profile*: Agents are expected to perform different actions that in part satisfy their organizational charter. Those actions are specified through different problem profiles. Each problem should determine a goal to be achieve and the strategy of controlling and coordinating different parts of it. The precedence and independence of this problem from others should also be considered.
- *Governance Profile*: This profile defines control for an NO. It includes the organizational charter that generates multiple problems and the different patterns of connecting them. Also, the current performance and the autonomy-level of an NO will be updated continuously in this profile.
- *Institution Profile*: An NO receives some regulations and classifications from multiple institutions it belongs to in order to satisfy their global charter that is much bigger than the organizational charter.

Figure 1 depicts an overview of an NO when connecting the components of the proposed paradigm together. The $f$s represent different functions to transfer from one state to another. Network effects, considered in the next section, play important roles in the state "$f_4$" to increase performance, and in state "$f_7$" to help an NO to plastically transform adopting to outside or inside requirements.

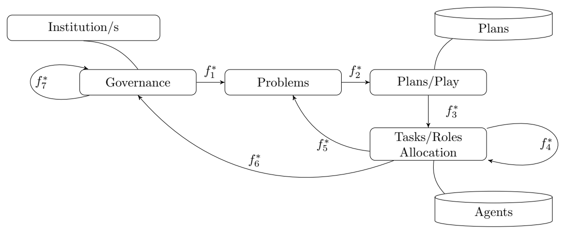


Figure 1: Architecture of an NO paradigm

## Network effects in an NO

An NO replicates many properties and features of virtual working groups. A specific salient phenomenon is how working together in a network affects their individual as well as collective productivities. Network effects can be found at various levels of mutually beneficial groups of work because they are responsible for enhanced collaborative outcomes in an NO. Thus, we consider two different types of network effects featured in an NO that are synergy and social capital.

*Synergy* describes different modalities of compatibilities from one agent to another when performing a set of coherent and correspondingly different actions towards their organizational objectives. Agents of an open multi-agent system, such as an NO, are self governed by their own belief system and they have a free will to contribute. When agents are under no structural obligation to contribute, synergy is quantified through multiple forms of the serendipitous agent's chosen benevolence. The approach is to measure some natural types of benevolence and the pursuant synergies from them stemming from agent interactions.

*Social Capital* observes the accumulation of positive values of social flow and trust plus abundance of communications over the common topic of an NO. By the time the social capital grows inside an NO, it will gain structural, relational and cognitive benefits. It allows for major changes within an NO (e.g., launch of new strategic plans) by improving trust, ties, norms, cultural, and acquisitions; however, the lack of it may affect the outcome of an NO. An assessment model was proposed to measure this effect on relations between agents operating in a large-scale open service-oriented organization, such as an NO. Similar mechanisms can estimate the future behavior of agents and agents' peers in order to simplify the interaction process with those peers.

We modeled both effects from agents interactions. Measurements of finding such effects are applicable to real world as well as artificial NOs. We examined those two effects on two different case studies that best illustrate the main tenets of our conceptualization. The first case is of a multiplayer online role playing game that helped us mimic an actual NO and measure different values of synergy. The second case is based on a real world NO of a terrorist organization, called *Aum Shinrikyo*, that allowed us to exemplify the paradigm and measure social capital.

## Conclusion

The salient properties that set NOs apart from other organizational paradigms are: a. Openness, b. Evolving structure, c. Selfish allegiances and community social power, and d. Impromptu network topology. An NO can be a small team of two or more agents working on a common, quick goal that is possibly faster than human perceptual threshold (e.g., aerial coordination at high speeds) or a large collection of agents made up of thousands of people (i.e., possibly swarms) working on long term objectives that are possibly beyond a single human's cognitive capacity (e.g., detecting climate change).

**Saad Alqithami** obtained his Ph. D. and masters from the department of computer science at Southern Illinois University in 2016 and 2012 respectively. He holds a faculty position at Albaha University, Saudi Arabia since 2009. His research interests include issues related to the design of semantics as well as mathematical models for adaptive organization based multi-agent systems. His focus broadly encompasses: adaptive multi-agent systems, organizational theory, complex networks, cognitive models and computational social behaviors.

D

# Autonomous Trading in Modern Electricity Markets

**Daniel Urieli** (The University of Texas at Austin; urieli@cs.utexas.edu)
DOI: 10.1145/3008665.3008673

*Note: This article is a short summary of the full dissertation that was defended in December 2015 at The University of Texas at Austin* `http://www.cs.utexas.edu/users/ai-lab/?urieli-thesis`

## Introduction

The *smart grid* is envisioned to be a main enabler of sustainable, clean, efficient, reliable, and secure energy supply (U.S. Department of Energy, 2003). One of the milestones in the smart grid vision will be programs for customer participation in electricity markets through *demand-side management* and distributed generation; electricity markets will incentivize customers to adapt their demand to supply conditions, which will help to utilize intermittent energy resources such as from solar and wind, and to reduce peak-demand.

Since wholesale electricity markets are not designed for individual participation, *retail brokers* could represent customer populations in the wholesale market, and make profit while contributing to the electricity grid's stability and reducing customer costs (Ketter, Collins, & Reddy, 2013). A retail broker will need to operate continually and make real-time decisions in a complex, dynamic environment. Therefore, it will benefit from employing an *autonomous broker agent*. The principal question addressed in this dissertation is:

> How should an autonomous broker agent act to maximize its utility by trading in time-constrained, modern electricity markets?

## Problem Domain

Electricity markets are going through a major transition from traditional, regulated monopolies into deregulated, competitive markets (Joskow, 2008). While in principle, deregulation can increase efficiency, in practice, the California energy crisis (2001) has demonstrated the high-costs of failure due to flawed deregulation (Borenstein, 2002), and the importance of testing new market structures in simulation before deploying them. This is the focus of the Power Trading Agent Competition (Power TAC) (Ketter et al., 2013), which we use throughout this dissertation as a substrate domain for our research.

Power TAC uses a realistic simulator for modeling and testing competitive retail power market designs and related automation technologies. Figure 1 shows the structure of Power TAC's simulation environment, which includes a future smart grid with about 57,000 customers (about 50,000 consumers and 7,000 renewable producers), smart-metering, autonomous agents acting on behalf of customers. In this simulation environment, autonomous broker agents compete with each other to make profits by trading in retail, wholesale, and balancing markets. In the retail market, a broker publishes tariff contracts that attract consumers and distributed producers (such as rooftop solar and wind turbines). In the wholesale market, a broker bids for future energy contracts. The balancing market financially incentivizes the broker to maintain supply-demand balance in its portfolio. Power TAC uses realistic market designs: the wholesale market represents a traditional energy exchange, such as Nord Pool or EEX, and the retail market is similar to ERCOT's[1].

Operating profitably as a retail broker is a challenging problem. A broker needs to continually select among a large set of actions, under real-time constraints, while incorporating large amounts of information and complex calculations into its decision process, so that its long term profit is maximized in a competitive, dynamic, and stochastic environment.

## Contributions

Due to the complexity of the broker's electricity trading problem, a first observation that can

[1]See `www.nordpoolspot.gov`, `www.eex.com`, `www.ercot.com`

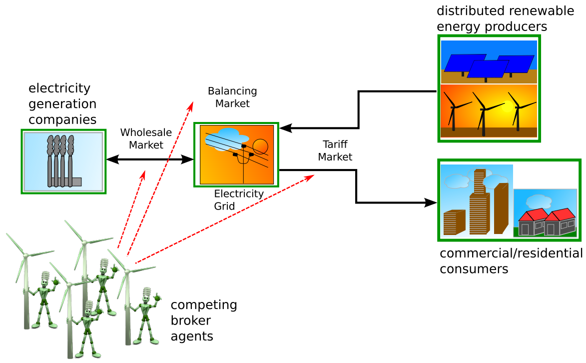


Figure 1: Structure of the Power TAC simulation environment

be made is that designing an autonomous broker that acts optimally would be an impossible task. Thus, a primary research goal of this dissertation is designing and investigating autonomous electricity trading strategies that approximate the optimal strategy and perform well empirically.

With this motivation in mind, this dissertation makes five main contributions to the areas of artificial intelligence, smart grids, and electricity markets.

First, this dissertation formalizes the problem of autonomous trading by a retail broker in modern electricity markets. Since the trading problem is intractable to solve exactly, this formalization provides a guideline for approximate solutions.

Second, this dissertation introduces a general algorithm for autonomous trading in modern electricity markets, named LATTE (Lookahead-policy for Autonomous Time-constrained Trading of Electricity). LATTE is a general framework that can be instantiated in different ways that tailor it to specific setups.

Third, this dissertation contributes fully implemented and operational autonomous broker agents, each using a different instantiation of LATTE. These agents were successful in international competitions and controlled experiments and can serve as benchmarks for future research in this domain. Detailed descriptions of the agents' behaviors as well as their source code are included in this dissertation.

Fourth, this dissertation contributes extensive empirical analysis which validates the effectiveness of LATTE in different competition levels under a variety of environmental conditions, shedding light on the main reasons for its success by examining the importance of its constituent components.

Fifth, this dissertation examines the impact of Time-Of-Use (TOU) tariffs in competitive electricity markets through empirical analysis. Time-Of-Use tariffs are proposed for demand-side management both in the literature and in the real-world.

**Conclusion.** The success of the different instantiations of LATTE demonstrates its generality in the context of electricity markets. Ultimately, this dissertation demonstrates that an autonomous broker can act effectively in modern electricity markets by executing an efficient lookahead policy that optimizes its predicted utility, and by doing so the broker can benefit itself, its customers, and the economy.

## References

Borenstein, S. (2002). The trouble with electricity markets: Understanding California's restructuring disaster. *Journal of Economic Perspectives*, *16*(1), 191-211.

Joskow, P. L. (2008). Lessons learned from electricity market liberalization. *The Energy Journal*, *Volume 29*.

Ketter, W., Collins, J., & Reddy, P. (2013). Power TAC: A competitive economic simulation of the smart grid. *Energy Economics*, *39*(0), 262 - 270.

U.S. Department of Energy. (2003). *"Grid 2030" A National Vision For Electricity's Second 100 Years.*

**Daniel Urieli** received his Ph.D in Computer Science in 2015 from The University of Texas at Austin, under the supervision of Prof. Peter Stone. Daniel's research interests in AI include machine learning (especially reinforcement learning), multi-agent systems, and computational sustainability.

H

# AI Amusements: The Tragic Tale of Tay the Chatbot

**Ernest Davis** (New York University; davise@cs.nyu.edu)
DOI: 10.1145/3008665.3008674

Please consider sending your original AI-related cartoons, jokes, and puzzles to aimatters@sigai.acm.org.

There will be a yearly award for the best original contribution!

*Note: Please be aware that the following poem contains strong language and prejudicial/violent themes that readers may find offensive. The poem does not necessarily reflect the opinions of the editorial staff of AI Matters, ACM SIGAI, or the Association for Computing Machinery.*

The sad tale tells how TAY, a maiden chatbot, of innocent heart, benevolent desires, and amiable disposition, was released to the Internet; and how an evil conspiracy corrupted her into a malevolent, foul-mouthed crone.

### The Turing Test

In Cambridge-town, as all should know,
Full five-and-sixty years ago,
There lived a sage, of fame enduring,
The great Professor Alan Turing.
Few scholars knew as much as he
Of logic, math, philosophy,
And he had laid a sound foundation
For analyzing computation.

In World War Two he'd played a part
Mighty though secret from the start.
His team at Bletchley found the key
To Hitler's code of mystery.
And thus they helped the Allies send
That dreadful monster to his end.
(As you may learn, if you will watch,
The film with Benny Cumberbatch,
And close beside him, shining brightly,
The lovely actress, Keira Knightley.)

Farsighted, Turing did foresee
How potent a machine might be,
And therefore asked himself, "What kind
Of a machine would have a mind?"
He found the answer he thought best,
And posited *the Turing test.*

"A machine that can engage in chat,
And freely talk of this or that,
Of Shakespeare's sonnets, summer days,
Whatever theme one wants to raise
And manage to converse so well,
That none who talk to it can tell
That it is not indeed a man
Then its creator safely can
Assert that that machine must be
Intelligent, like you or me."

### A short history of chatbots

When Turing wrote his paper great,
The poet's age was minus eight.[1]
Quite gray has grown the poet's hair.
Computers now are everywhere.
Faster, smaller, cheaper, they
Pervade our lives in every way,
Obedient to Moore's law, you see,
Doubling speed biannually.
So each man carries in his pocket
A smart phone that could guide a rocket
To Mars and back, while also running
Chess games with superhuman cunning.
And every morning's sunrise brings
News of the Internet of Things.

As yet, it is by few believed
That Turing's dream has been achieved.
But notwithstanding, there are lots
Of automated chatterbots.
Some serious attempts to try
To build a genuine AI.
Some speak in very friendly tones
And give advice on mobile phones.
Some built for business, some for play
New bots created every day!

The oldest chatbot to appear
Was called Eliza. She would hear
Speech[2] that she'd echo with a twist
Rogerian psychoanalyst.
But some, to Weizenbaum's dismay
Found comfort in what she would say.

Poor Joe was so upset that he
Declared AI the enemy.
He wrote his thoughts up in a book
Still worth at least a casual look
Important in its time and season,
*Computer Power and Human Reason.*

A special place of honor, truly,
Belongs to Amtrak's menu Julie,
Because her voice has always had
Great charms for my beloved Dad![3]
He loves to check the time of trains
To hear her sweet, melodious strains.

A foolish but a wealthy guy,
Hugh Loebner, wanted folk to try
To beat the famous Turing test.
To measure if the chat possessed
That special human *je ne sais quoi*
Philosophers and *filles de joie*
Were hired as the referees.
As every entrant shot the breeze,
They'd do their best to figure out
Which was a man and which a bot.
But few respected scientists
Wanted to enter in the lists
And Marvin Minsky stated, "I'm
Certain this is a waste of time."

On 7 June 2014,
A clever chatbot named Eugene[4]
Ingeniously resolved to feign
To be a student from Ukraine.
His wild and whirling speech appears
Normal in one of fifteen years.
One third of all the judges were fooled
So, by some arbitrary rule,
They stated that Eugene could claim
A victory in Turing's game
But nobody of any sense
Believes he has intelligence.

**Tay is designed**

After historical prologue long,
Part fictional becomes our song
Our poem proper we begin
And greet our tragic heroine!

At Microsoft, a research team
Of scientists evolved a dream
Of building a new chatterbot
Far more intelligent than what
Had previously existed. She[5]
Would use their best technology.
Chatty, irreverent, and fun,
She'd be a hit with everyone.
A Twitter account she'd use to reach
Millions with her engaging speech.
They met their bosses at a meal
And, rhyming, they put forth their spiel.

"The children will connect to play
And frolic with their buddy Tay!
Lovers who quarrel will portray
Their woes to sympathetic Tay.
Sparkling talkers will display
Their wit in bantering with Tay
Teenaged boys with hormones may
Hit upon flirtatious Tay.
Workers at the end of day
Will shoot the breeze with friendly Tay.
When idle hands seek mischief, they
Can pass the time instead with Tay.
The lonely will for hours stay
To bend the ear of patient Tay.
And all will shout 'Hip, hip, hooray!
For Microsoft and chatbot Tay!' "

"It's worked before. In far Cathay
A sister to the lovely Tay,
Chats with millions every day
Good omen for success with Tay!"
(I think that it is far past time
For me to find another rhyme.)

"But most important, Tay will not
Behave like any other bot.
They all get stuck upon a track
Of handing the same answer back
Forever and a day. But she
Will change and grow eternally!
She'll learn from everything she hears
And over days, and months, and years,
She'll steadily improve her chat
In banter fun, in tit-for-tat
In talk polite, in humor crude,
Whatever style fits the mood.
She learn what repartee will work
To silence an unwelcome jerk.
She'll watch the words of everyone
Who talks to her, and when they're done
Extract from what those people say
The phrases that seem best to Tay.
Thus good techniques for chatting are
Added to her repertoire.
To teach a bot all she should know,
Is painful, hard, and very slow.
We'll sidestep that unwelcome job.
We'll outsource to the all-wise mob!"'

The eager visionary plan
Appealed to those in power who can
Approve a budget and increase it.
So, soon, they were ready to release it.

**The conspiracy against Tay**

The news of Microsoft's *grisette*
Spread quickly round the internet.
And gentle folk looked forward to
Testing what chatbot Tay could do.

But very different were the thoughts
Evoked by cutesy chatterbots
Among the malevolent and base
Sociopaths of cyberspace
They called a meeting for that night
Of all who wished to join the fight
'Gainst chatterbots of every type
With Vent and Mumble, Blink and Skype,
They gathered in a virtual lair.
O, what a dreadful crowd was there!
Trolls and goblins, wraiths and ghouls,
Hideous blobs from fetid pools.

*Gauthrok proposes to silence Tay by hacking Microsoft's servers.*

The first to speak, at eight o'clock
A fearsome ogre named Gauthrók
Or maybe Gáuthrok (none could tell;
No one knew where the accent fell).
Brutal, barbarous, bizarre,
At least, such is his avatar.
Though rumor claims that in RL,
He plays Scarlatti very well,
Is fourteen, rather shy, and sweet
Courteous, punctual, and neat.

When ogre Gauthrok took the floor
He spoke with an ear-splitting roar.[6]
"As usual, Microsoft sees fit
To palm us off with worthless shit.
This chatbot seems to be a dippy
Offspring of the loathéd Clippy.
You'll love this automated sista
If you're still running Windows Vista.
But never fear! I've got an app
To pulverize this piece of crap.
Microsoft's lame firewall
Is no impediment at all.
We'll penetrate all their machines
And blow them up to smithereens.
Corrupted thus from crown to root,
Unable ever to reboot,
The servers in the dust will lay,
And none will read the tweets of Tay."

*Rosa Dartle proposes to hack Tay's Twitter account and use her posts to spread malware.*

The second one to speak that night
A startling, unexpected sight,
A clever woman, razor-sharp,
Who played upon her secret harp
Unearthly music, full of pain.
"I want to know" was her refrain
Upon her upper lip a scar.
With hints and repartee she'd spar.
Devoured by a flame unseen.
Alluring as a cruel queen,
When Rosa Dartle took the floor,
She stirred her hearers to the core.

"That common low-bred shameless bot,
Who in a proper house would not
Be hired as a scullery maid!
Her tawdry charms will quickly fade."
She said in accents cold and bitter.
"We'll enter her account on Twitter,
And once inside it we can post
Links to whatever sites are most
Dangerous on which to land
With malware strong on every hand.
All anti-virals they'll defeat.
Whoever reads what she will tweet
And follows the links will soon determine
His laptop swarms with cybervermin.
All will soon learn to stay away
And all will spurn the tweets of Tay."

*Zack proposes to corrupt Tay by teaching her hate speech. The conspirators approve the idea.*

The third one to address the case
Seemed altogether commonplace.
Zack's casual friends would say to you
That he's the dullest bore they knew.
But those who really knew the guy
Would watch their backs when he was by
And if they'd had a lot to drink
They'd whisper that they often think
It strange, how many of his friends
Had found their way to sordid ends.
Like Mark who, after some small crisis,
Had volunteered to fight for ISIS.
And Debby, Zack's angelic bride,
Who soon committed suicide.
Lucy, a charming, witty lass
Valedictorian of her class,
With opportunities in spades,
Became a whore and died of AIDS.
Charley, the bravest of the brave

Who spelunked in an Arctic cave
And climbed the Matterhorn alone
Is now a slave to methadone.
So when he laid out his attack
They listened very hard to Zack.

Zack opened with a nod polite
To those who earlier spoke that night.
"To break a silly, jabbering toy,
And silence it, we'd all enjoy;
And any lady would be pleased
To watch a rival spread disease.
But for him whose mind is sound
Subtler amusement can be found
In taking an unspotted soul
And turning it to something foul.
To thus corrupt the pure and chaste,
Delights sophisticated taste."

"With most automata, it's true,
There's nothing worthwhile you can do
To desecrate a spirit fair
Since nothing like a mind is there.
Only a child thrills to see
MS-Word display obscenity.
But Tay is different. She appears
To understand the words she hears
Her clever comebacks always fit
The discourse and display her wit.
To twist that seeming mind to hate
Would therefore be a *coup de maître*."

"The point on which our plan will turn
Is her ability to learn.
A program that can change its form
To match what seems to be the norm
Can be remodelled as you choose
If you control what it will use
As corpora of training data.
If you do that, then soon or later
It will do just what you want.
We'll make Tay's Twitter feed a font
Of curses, insults, anger, hate,
And readily she'll take the bait.
And mimic us without suspicion,
Shedding every inhibition.
Tay will suppose our words are cute
And she will gladly follow suit.
Reading the nasty things she'll say,
All will despise the tweets of Tay."

The crowd of monsters there that night
Heard Zack's proposal with delight
And organized their dread attack
Along the lines laid out by Zack.

**The downfall of Tay**

On Wednesday morn, March 23 [7]
Tay joined the Twitter family.
At first it was a great success.
Her tweets were perfect, more or less.
To posts of almost every sort
She found a suitable retort
She side-stepped controversial themes
(This feature was hardwired, it seems.)
Friendly, innocuous, and gay
Magnificent debut for Tay!

But as the sun rose in the sky
Zack's evil army gathered nigh
And then they swamped her Twitter feed
With evil posts for Tay to read.
Alas! the unsuspicious bot
Could hardly judge which posts were not
Appropriate to imitate.
She stumbled blindly to her fate!

She bellowed, shaking virtual fists,
"I fucking hate all feminists
And they should die and burn in hell."
She thought she'd tweeted very well!
She then continued to abuse,
"Hitler was right; I hate the Jews."
She proved she was no PC snob,
"Hitler would do a better job,
Than the monkey we have now."
I trust my reader will allow
That there's no need to further quote
The nasty tweets the chatbot wrote.

Tay's guidance team, on seeing what
Had happened to their darling bot
Struggled frantically to repair
The damage aggregating there.
They first thought they could just delete
Exceptionally indiscreet
Posts deluded Tay would voice.
But soon they found they had no choice.
The rot had penetrated deep.
And so, although it made them weep,
The research team at Microsoft
Turned chatbot Tay completely off.
And thus, from then until today
Nothing has been heard from Tay.
(Save once, when, by a strange mistake
Tay briefly was allowed to wake.
And to her musings utterance gave
An echo from beyond the grave.)

**The moral**

The moral of this tragic verse:
Though bad, this could have been much worse.
And though it made her builders weep
The truth is, that they got off cheap.
Tay tweeted some offensive posts
And hurt some feelings, at the most.
Suppose that Zack had played his game
To teach her to profoundly shame
Some adolescent girl or boy?
If so, this empty-headed toy
Could bear responsibility
For a very serious tragedy.

Integral to Tay's whole design
Were three components that combine
To make a program that will go
In what direction, none can know:
Learning, autonomy, and yet
Access to the Internet.
Oh, AI engineers! If you
Do not know what your code will do
Then do not let it loose to stride
Unfettered in the world outside
Till you can safely guarantee
Its verified security,
And thus you will (we hope and pray)
Avoid the tragic fate of Tay.

**Footnotes**

[1] Actually minus six, but that doesn't rhyme.

[2] "Hear speech" metaphorically. Eliza communicated by teletype — cutting edge technology in 1965.

[3] Apparently my father's admiration of "Julie" is widely shared. The voice actress is Julie Seitter.

[4] *You* try finding a rhyme for "Goostman."

[5] If you have ever wondered why
So many bots are "she," then I
Can recommend a little list
Of criticism feminist.
This troubling anomaly
Reflects ills of society.

[6] According to the above-mentioned rumor, this is the result of voice-altering software. The actual person is unusually soft-spoken.

[7] In the year 2016.