Proceedings of the Twenty-Ninth International
Florida Artificial Intelligence Research Society Conference

On Modeling the Interplay Between Opinion Change and Formation

Fei Yu and Eugene Santos, Jr.
Thayer School of Engineering, Dartmouth College, Hanover, NH 03755
{Fei.Yu.TH,Eugene.Santos.Jr}@dartmouth.edu

Abstract

We have a limited understanding of how an opinion is originated, how an opinion gets conveyed, and how the communicated opinion is perceived and processed by others. Extant research concerning "opinions" primarily focuses either on the conditions and determinants for opinion formation or on opinion change induced by external influence. Furthermore, existing cognitive/computational models typically address a single opinion formation or change process and rarely consider the interplay between these processes. We propose a new computational model for opinions that recognizes the learning nature of opinion change and the decision-making nature of opinion formation: what has been learned through internalizing an external influence guides how decisions are made to externalizing cognitive processes. The Double Transition Model (DTM) represents a networked space of reasoning processes providing a computational framework of opinions encompassing both formation, change, and their continuous interplay. We apply DTMs to a simulated dyadic, multiple episode opinion change and formation problem to determine how to best train an advocate to convince others regarding an institution's ranking.

Introduction

What are opinions? A workable definition from Wikipedia states: an opinion is a "[… subjective] statement […], i.e. based on that which is less than absolutely certain, and is the result of emotion or interpretation of facts. […] opinion may be the result of a person's perspective, understanding, particular feelings, beliefs, and desires." We further consider opinions to be personal beliefs (Krueger, 1996) – an opinion is belief as it is derived rather than recalled; an opinion is personal as it can be derived in diverse ways by different agents. The continuing interest in opinion mining research (e.g., sentiment analysis, etc.) has produced technologies across domains to elicit individual and group opinions. Yet, we are still scratching the surface to understand (cognitively and computationally) how an opinion is originated, how an opinion and the information supporting and explaining it gets conveyed, and how the communicated opinion is perceived and processed by others.

In existing computational models (Weisbuch et al. 2003; Yildiz et al. 2011), the opinion formation process and opinion change process are studied separately. The area of opinion formation focuses on the conditions and determinants for opinion formation (Watts & Dodds 2007) while the area of opinion change focuses on identifying internal and external (social) influence (Friedkin & Johnsen 2011). There is nothing learned from an opinion change process besides the opinion value itself and there are no decisions being modeled for an opinion formation process besides the opinion value itself (Martin et al. 2005).

Our goal is to develop a computational framework to learn one's cognitive state space by considering one's perceptions and behaviors. The use of this framework is twofold: 1) Infer an individual's opinion and its change, and be able to explain the process; and, 2) Infer group dynamics on opinions. The key idea in designing the framework is based on the following observation: that of the *learning nature of opinion change* (Kelman 1961) and the *decision-making nature of opinion formation* (Clayton 1997). These refer to changes in one's own beliefs or at least supporting information when opinion change occurs. In turn, such ever-changing beliefs also yield different opinions later for the same or similar subject matters. The decision-making nature of opinion formation refers to the motivation driving a specific opinion. To put it more formally, the interplay of opinion change and formation is described as: 1) the internalization of an external influence (opinion change) guides the externalization of internal cognitive processes (opinion formation); and, 2) the externalization of an internal cognitive process (opinion formation) becomes external influents that can in turn influence others' opinions.

The key insight to our model is to exploit the interplay between opinion change and the opinion formation process. The core of our framework is an inter-connected knowledge-base of cognitive states. The opinion change process would affect the current cognitive state, while the opinion formation process can be considered as a "walk" through the cognitive states. Such a walk as we shall see can be computed from a Markov Decision Process (MDP).

We present a new probabilistic cognitive model called the Double Transition Model (DTM) which is a networked space of reasoning processes where each node represents a cognitive state with different degrees of query and knowledge organization and incompleteness. The output of a reasoning process is an emitting probability representing a possible opinion. The edges within a DTM denote how a query and knowledge differ between connecting states reflecting the influence that leads to opinion change. Edges can capture influence both exogenous and endogenous.

Furthermore, we applied the DTM to simulating a dyadic, multiple episode (multi-pair) opinion change and formation problem where the goal is to train an individual to convince others regarding a target opinion. The process of training introduces learning episodes through exposure to different trainers that affects the trainee's opinion formation process and their own opinions. This is modeled as an MDP using Q-learning (Sutton and Barto 1998). After training, the trainee is then set to convincing a population of testers on the target opinion. We examine the successful consensus rates and number of interaction turns used by the trainee. Our results indicate that variations in trainee personal beliefs and goals are directly reflected in the effectiveness of the training process.

Approach

We define an opinion change process as the process of internalizing one's external influence, and we define an opinion formation process to be the process of externalizing one's internal cognitive processes. We define an opinion formation task as undergoing a series of opinion formation and opinion change processes on one issue in a single context. Intuitively, an opinion formation task can include situations such as a librarian wanting to suggest a good textbook on Artificial Intelligence, congress members trying to reach consensus in a debate, or two or more family members collaborating on the Christmas gifts to buy.

A framework that processes a sequence of opinion formation tasks allows us to achieve a better understanding of how an external influence gets internalized (opinion change), how an internal opinion gets externalized (opinion formation), and the interaction between these two processes in both short-term (one task) and long-term (sequence of tasks) resolutions. To begin with, we can exploit the close interplay between formation and change processes within a task in order to learn the mechanisms of internalization.

The challenges of modeling an opinion formation task lies in modeling both the learning and decision-making aspects within an interactive environment. We recognize that the individual performing an opinion formation task is essentially engaged in a sequential decision-making problem with a specific goal in mind.

It is fairly easy to hypothesize what the external influents are by considering the behavior of an individual engaged in an opinion formation task. These external influents can come in a wide variety of forms such as receiving/observing an opinion from someone, perceiving/sensing the environment for evidence, actively seeking information, the question itself that asks for opinions, an action performed, or behavior (e.g. threats (Carver 1977)), punishment and external surveillance (Hoekstra 1995)), that can be observed from others. To simplify the problem, we concentrate on influents in the form of messages since what has been internalized is much less clear as it all happens within a human brain which, of course, lacks "visibility". A variety of social theories have identified that the internalization may include the knowledge basis (McGuire 1968) from which an opinion is formed, a value system (Kelman 1961), and sentiment towards others (Robinson et al. 2006). To be consistent with our simplification, we concentrate on the underlying knowledge base and the reasoning process from which an opinion can be derived.

Our goal here is to develop a knowledge-centric computational framework that processes sequential opinion formation tasks. Consider an opinion sequence for some individual A: $\{[o_1], [o_2, o_3, o_4], [o_5], [o_6, o_7]\}$ where each o_i is a degree of belief of an opinion over time i. The sequence in brackets is a sequence of opinions one has formed within the same context on one issue (e.g., one has been reading a collection of news articles on a presidential candidate, or one is having a heated back and forth discussion on who should be the next president). Each bracketed sequence is an opinion formation task. Lastly, well-established social theories have concluded that assumptions of $o_4 \equiv o_5$ rarely holds. In other words, the final opinion formed within an interactive environment may not carry over into the next task – non-persistence (Pierro et al. 2012).

Double Transition Model (DTM)

An individual faces a sequence of different opinion formation tasks $\{t_1, t_2, t_3, t_4, t_5\}$, each with different subsequences of external and internal changes caused by the (task-interleaved) actions taken by the individual:

$$\{\langle a_1^1, s_1\rangle, \langle a_1^2, s_2\rangle, \langle a_2^1, s_3\rangle, \langle a_1^3, s_4\rangle, \langle a_2^2, s_5\rangle, , \dots, \langle a_5^{m_5}, s_n\rangle\}$$

where a_j^k is the k[th] action for task j and s_l is the individual's state (knowledge, beliefs, experiences, etc.) after l actions taken. The decision process behind task t_1 is $\{\langle a_1^1, s_1\rangle, \langle a_1^2, s_2\rangle, \langle a_1^3, s_4\rangle, \langle a_1^4, s_4\rangle\}$.

A *Double Transition Model (DTM)* (Figure 1) consists of two sub-models: a *query transition graph (QTG)* and a *memory transition graph (MTG)*. Each node in a QTG represents a single query at a time. Each query is captured when an opinion is being requested by someone else. Therefore, the creation of a new query is triggered by a task while the new creation of a memory is triggered by perception. The QTG reflects changes to the query over time. As a complement, the MTG embodies the changes in the memory of the individual as actions are taken over time. At any given time, the states of the QTG + MTG determine the answer (opinion) of the individual. Actions (internal and external) derived from the decision result in transformations in one or both graphs. Thus, the decision processes for the tasks is simply the simultaneous (cross-product) transition "walk" through the QTG and MTG.

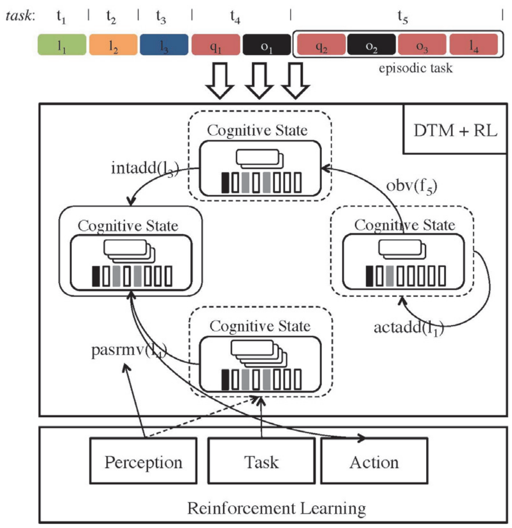

Figure 1. DTM + RL

DTMs are defined as follows: A *query* is a propositional logic statement where the terms are different assignments to random variables (rvs) in U, the universe of rvs. The space of all possible queries is denoted ∇. For example, "Do you support abortion (r.v. A)?" is represented as query $?(A = support \vee A = against)$. A *query transformation function (qtf)* F_q is a function from ∇ to ∇, e.g., substituting assignment $A = a$ with $A = b$ in the query, simplifying by adding or removing terms (assignments), and even negating a clause are possible transformations. Each task t_i has an associated *query space* $\varphi(t_i) = \langle q(t_i), T_q(t_i)\rangle$ where $q(t_i)$ is a finite set of queries and $T_q(t_i)$ is a finite set of qtfs. Basically, each transformation is a change in the question being considered by the individual driven by both internal and external factors of the individual and his/her decisions/actions. For example, transformations include changing hypothesis or simplifying a query.

Def 1. A *QTG* $\boldsymbol{Q}$ is an undirected graph $(\boldsymbol{V^Q}, \boldsymbol{E^Q})$, where $\boldsymbol{V^Q}$ is a finite subset of $\bigcup_i \boldsymbol{q}(\boldsymbol{t_i})$ and $(\boldsymbol{v_1^Q}, \boldsymbol{v_2^Q}) \in \boldsymbol{E^Q}$ only if $\boldsymbol{F_q}(\boldsymbol{v_1^Q}) = \boldsymbol{v_2^Q}$ for some $\boldsymbol{F_q} \in \boldsymbol{T_q}(\boldsymbol{t_i})$ and $\boldsymbol{i}$.

The representation of a query can simply be a vector such as $[X, ?, ?, 3, \ldots, 2]$ where a numeric value represents either a category or a discretized number for a feature of an entity (instead of a proposition), a question mark represents unknown value for a feature, and X represents the target feature of an entity. For example, one may ask another for opinions about a university: "This school has tuition around \$42,000, yearly enrollment around 11,000 and an acceptance rate around 7.7%, do you think this is a tier 1 school?" The other features that are not included in an inquiry are treated as unknown, and the targeted feature is the rank of the school. In Figure 1, unknown features are represented by white rectangles, known features are grey rectangles, and targeted features are black.

Let Ξ be the space of probabilistic networks over U which form the underlying knowledge/memory of the individual in each cognitive space that can be defined on U. A *memory transformation function (mtf)* F_m is a function from Ξ to Ξ which represents how memory changes. It could be a simple factual change to more sophisticated changes introducing new linkages and re-arrangement of existing correlations. For example, a person may need to make decisions under time pressure so must rely on heuristics such as recent experience instead of time-consuming, fully detailed analysis (see (Yu 2013) for mechanisms to represent changes in reasoning heuristics in terms of memory transitions). This is a transformation where recent memory and knowledge is emphasized by biasing the underlying probability distribution as reflected in the probabilistic network. Let $\Gamma = (K, T_m)$ be the *memory space* where K is a finite set of probabilistic networks and T_m is a finite set of mtfs operating on K.

Def 2. A *MTG* $\boldsymbol{M}$ is an undirected graph $(\boldsymbol{V^M}, \boldsymbol{E^M})$ where $\boldsymbol{V^M}$ is a finite set from $\boldsymbol{K}$ and $(\boldsymbol{v_1^M}, \boldsymbol{v_2^M}) \in \boldsymbol{E^M}$ only if $\boldsymbol{F_m}(\boldsymbol{v_1^M}) = \boldsymbol{v_2^M}$ for some $\boldsymbol{F_m} \in \boldsymbol{T_m}$.

With the formulations of both query transition graphs and memory transition graphs as the basis for transformations, we combine them to form DTMs as follows:

Def 3. A *DTM* $\boldsymbol{D}$ induced by QTG $\boldsymbol{Q}$ and MTG $\boldsymbol{M}$ is the undirected graph $(\boldsymbol{V^D}, \boldsymbol{E^D})$ where $\boldsymbol{V^D} = \boldsymbol{V^Q} \times \boldsymbol{V^M}$ and there is an edge between $\boldsymbol{v_1^D} = (\boldsymbol{v_1^Q}, \boldsymbol{v_1^M})$ and $\boldsymbol{v_2^D} = (\boldsymbol{v_2^Q}, \boldsymbol{v_2^M})$ if and only if (1) $\boldsymbol{v_1^Q} = \boldsymbol{v_2^Q}$ or $(\boldsymbol{v_1^Q}, \boldsymbol{v_2^Q}) \in \boldsymbol{E^Q}$ and (2) $\boldsymbol{v_1^M} = \boldsymbol{v_2^M}$ or $(\boldsymbol{v_1^M}, \boldsymbol{v_2^M}) \in \boldsymbol{E^M}$.

DTMs allow for probabilistic reasoning and can interact with other DTMs. Thus, the nodes in the MTG and QTG represent such networks and inferences to be computed, respectively. As we shall see, certain DTMs for opinion change can be mapped to Markov Decision Processes (MDPs) which allows the ordering and selection of decisions and actions to be learned through Reinforcement Learning (RL) (Sutton and Barto 1998). This ordering is critical because changes to the environment and individuals can have different effects with different orderings.

Representing Opinion Change and Formation

A DTM can be used to compute an opinion given the target/current cognitive state, but why does a state change happen and why are new states constructed? Consider the space of all possible cognitive states arising from any/all mappings from perceptions (incl. biases, acceptance, experience, etc.) and tasks (incl. query variants, miscommunication, interactions, etc.) into a cognitive state. The cognitive state can also reflect any number of opinion inferences using different reasoning style such as heuristics. As such, a DTM is constructed from this space and represents a particular instance/individual. From this point of view, to answer our question about why (and how) DTM's are dynamically formed goes directly to our recognition of the learning nature of opinion change and the decision-making nature of opinion formation. The former occurs as we are dynamically building each DTM while the latter as we are choosing the transitions to take. Together, the decision processes is simply the simultaneous (cross-product) transition "walk" through the QTG and MTG.

Figure 1 also demonstrates how a DTM functions in a dynamic environment. As shown in the sequence on the top of the figure, l_1 to l_3 are perceived learning episodes in three consecutive steps corresponding to tasks t_1 to t_3. Tasks t_4 and t_5 are *opinion formation tasks*: For example, in t_4, an external individual asks for an opinion on query q_1. The DTM then forms an opinion o_1 and sends it out. In the task sequence shown at the top of the figure, the non-black rectangles are received information including perceptions (represented by learning episodes) and task information (represented as queries and external opinions, e.g. o_3). The black rectangles are outgoing messages. We consider t_4 to be a *non-episodic opinion formation task* as it does not maintain an interactive session with the individual who asks for opinions. In real-world problems, asking a librarian's opinion can be considered a non-episodic opinion formation task where the individual tends to be the domain expert. Non-episodic opinion formation tasks are common also when the problem itself is not controversial.

Lastly, t_5 is an *episodic opinion formation task*. Unfortunately, determining both the cognitive state currently in use and the likely transition (and potential construction) to another cognitive state cannot be derived solely from a DTM as we mentioned above. To model episodic opinion formation tasks, we further classify such a task with regards to whom an individual interacts with: A *non-dynamic one-to-one episodic opinion formation task* focuses on how to choose an action to fulfill his goal whereas his interlocutor also has a particular goal to fulfill. A goal can simply be reaching opinion consensus, changing the other's opinion, or do not care. A *dynamic one-to-one episodic opinion formation task* is that an individual engages in an opinion formation task with different people one at a time.

Given that we can define goals and actions based on a space of possible episodic opinion formation tasks, the DTM state graph, opinion emission, actions, and goals form the components underlying a sequential decision-making process, thus accounting for both the learning nature of opinion change and the decision-making nature of opinion formation. We can then solve an episodic opinion formation task by defining it as a Markov decision problem (MDP): At each time step, determine what is the best action to take to accomplish one's goal – e.g., a goal function that considers two aspects: 1) minimizing the gap between two players' opinions for next time step (both captured as DTMs), and 2) minimizing the change between his own opinions between two steps. It is very important to emphasize here that an episodic opinion formation task subsumes a non-episodic opinion formation task by defining a finite-horizon MDP with the horizon equal to one. This is a simple yet powerful generalization so that the episodic and non-episodic tasks can be modeled as MDPs. Given that each individual is unlikely to have perfect knowledge of another's DTM, if training experience is available, a Q-learning method can improve an agent's strategy by repetitively participating in episodic opinion formation tasks. Thus, our framework handles dynamic opinion formation tasks that consist of a sequence of opinion formation and change with external influences from multiple sources.

A Dyadic, Multiple Episodes Model

We now describe a decision problem in an episodic opinion formation task between two people as follows:

> Two agents, e_1 and e_2, are exchanging influents with each other guided by their respective goals. At each time step, agent e_1 needs to decide an action to take.

The goal for agent e_i can be specified as

$$\lim_{t\to+\infty}\left\{\gamma_i\left|o_i^{t+1}-o_i^t\right|+\varsigma_i\left|o_i^{t+1}-o_j^{t+1}\right|\right\}=0$$

where o_i^t is the opinion for agent e_i at time t, $i \neq j$, and $\gamma_i, \varsigma_i \in [0,1]$ are control parameters. The first term represents the degree of opinion change from time t to $t+1$ for e_i while the second term represents the gap in the two agents' opinions at time $t+1$. Replacing o_i^t by o_i^w where $w = 0$ or $w < t$ directly reflects the desire to revert to their original or an earlier opinion. The two goals cover the two ways to reduce the gap between two agents: one by moving e_1's opinion towards e_2's and vice versa. Parameters γ_1 and γ_2 denote values on a *malleability-idealism scale* from 0 to 1 representing an agent's willingness to change its own opinion; while parameters ς_1 and ς_2 are on a *passivity-activism scale* from 0 to 1 representing an agent's eagerness for reaching a consensus. The higher the malleability-idealism score is, the more idealistic an agent is (i.e., more unwilling to change its opinion). The higher the passivity-activism score is, the more active an agent is (i.e., more eager to reach a consensus). Each goal assesses how desirable each transition between states is for an agent.

Now consider the following hypothetical problem:

Problem Setting: We want to train advocates at Dartmouth College to be proficient at convincing others to believe it is a great university. We have materials about different universities, but unfortunately cannot recruit too many people to practice with that have a wide variety of beliefs and behaviors.

Target Questions: What type of advocate is best? How critical is representative training?

For this problem, "convincing" is when consensus is reached between the advocate and the current target individual. This is focused on situations where consensus is actively sought for but individuals may differ in the way to reach consensus. We collected and pre-processed the U.S. News 2013 College Ranking Data which contains nine attributes we mapped to 0 or 1 (Table 1 and Table 2).

Each university feature vector represents one piece of knowledge to be included or removed from memory – a single learning episode. Each learning episode (Table 2) is directly encoded as a single Bayesian Knowledge-Base (BKB) fragment (see

Figure 2). For this problem, each node in the MTG is simply a fusion of the BKB fragments (Santos et al., 2011) for each learning episode currently included, and an opinion probability is computed using the target query as evidence. Thus, we can determine the target opinion 'X' by reasoning over the fused BKB at each node in the MTG with evidence set based on the target query vector (in Table 2, attributes 2 and 9).

Table 1. U.S. News 2013 College Ranking Data Attributes

#	Feature	Criteria for Value of 1
1	Ranking	≥ 100
2	Tuition	≥ \$25,000
3	Enrollment	≥ 20,000
4	Accept ratio	≥ 35%
5	Avg freshman retention	≥ 86%
6	6-yr graduation rate	≥ 71%
7	Classes < 20	≥ 47%
8+9	SAT/ACT 25th-75th % (low+high)	≥ 1,010

Table 2. University and query feature vector samples. 'X' is the target opinion and '?' is an unknown feature value

Dartmouth	0	1	0	0	1	1	0	1	1
UCLA	0	0	1	0	1	1	1	1	1
Columbia	0	1	1	0	1	1	1	1	1
GaTech	0	0	1	1	1	1	0	1	1
Stevens IT	1	1	0	1	1	1	0	1	1
WPI	1	1	0	1	1	1	1	1	1
Query	*X*	*1*	*?*	*?*	*?*	*?*	*?*	*?*	*1*

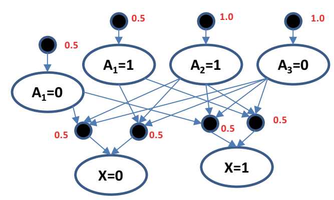

Figure 2. BKB fragment for learning episode {?,?,1,0}

We examine two goal profiles for trainees: Idealistic-Active (IA) profile ($\gamma_i = 1$ and $\varsigma_i = 1$) and Malleable-Active (MA) profile ($\gamma_i = 0$ and $\varsigma_i = 1$). As both agents' opinions are used in evaluating the goal, each state at a time step needs to include both agents' opinions. In order to transition from one state to a desired state according to the goal, every agent needs to decide on the actions to take. We simplify the problem in that each external influent contains one message (corresponding to a learning episode). Therefore, we construct a DTM with two states connected only if the difference between them is one learning episode. For our problem, the 5 schools (excl. Dartmouth) define 32 possible cognitive states (nodes) for the MTG. The QTG for this problem is relatively simple as the query is fixed throughout the simulation.

Based on our ideas above, we consider 7 possible actions for each agent: (1+2) intadd(l) & intremove(l) –internal action to add/remove learning episode l from consideration; (3+4) pasadd(l,e) & pasremove(l,e) – add/remove l from own memory as suggested by agent e; (5+6) actadd(l) & actremove(l,e) – suggest to agent e to add/remove l from their consideration; and, (7) donothing. Each action defines the transition in the MTG appropriately.

We consider a range of trainers and testers to cover individuals with different backgrounds the trainee might encounter. Trainers practice with the trainee in order for the trainee to learn (Q-learning with ϵ-greedy(0.05) behavior style) their decision-making process whereas testers are those the trainees must aim to convince that Dartmouth College is a Tier 1 ranked university. We generated 18 different types who differ in both goal profiles (γ, ς) from {(1,1), (0,1), (0,0), (1,0), (.75,.25), (25,.75)} and their behavioral styles {pureGreedy, ϵ-greedy(0.05), and τ-softmax(1.0)} (Sutton & Barto 1998) which determines their action choices with regards to their goal. Each trainee starts with some knowledge basis for their initial opinion – i.e., one out of the 32 possible university sets. This is also the same for each trainer/tester.

We consider two forms of dynamic opinion formation tasks – tasks where a trainee interacts with the same type of trainer (or a tester) and tasks where a trainee interacts with trainers (or testers) of different types. Every dynamic opinion formation task (also called one behavioral run) has 1000 episodic opinion formation tasks. Within a dynamic opinion formation task, the trainee may keep changing its opinion by interacting with trainers (or testers) in a sequential manner. If a pair of individuals do not reach consensus, we terminate their discussions after ten rounds of interactions and the trainee then continues to the next one in the queue. The initial opinion is randomly generated for the trainer (or a tester) in each opinion formation task. Within one behavioral run, a trainee may talk to agents of different types. For example, a trainee talks to pureIA-style agents for 500 times and talks to pureIP-style agents for 500 times. For our testbed, we randomly generated nine composite behavioral runs from the 18 types above. We conducted simulations for all pairs of trainee-trainer and pairs of trainee-tester. In total, we have ~140k behavioral runs.

The trainee is given the opportunity to initially determine what schools the trainer/tester considers to be top schools before beginning their discussion. This corresponds to the trainee knowing the DTM of the trainer/tester, and choice of actions is constrained when communication actions are employed, e.g., an agent cannot pasremove a learning episode l unless the other agent actremove l.

At any given time, the state of the environment in a trainee-trainer or trainee-tester interaction consists of the current nodes for the trainee and trainer DTMs plus the specific learning episodes communicated (if any). However, the trainee is not likely to have transition probabilities between states since this is in essence an on-line learning situation. Still, we have sufficient elements to apply Q-learning methods to learn the best policy (i.e., what action to choose in a given state) for our trainee. We assume trainers/testers always make the greedy choice of actions with respect to the trainee's action to achieve their goal.

Simulation Results

For the question, "What type of advocate is best?", we use consensus rate and number of turns to measure success at reaching consensus in each behavioral run. For the second question regarding representativeness, we ran simulations where the trainers were representative of the testers and cases where they were not. We compute consensus rate as the % of opinion formation tasks that result in consensus out of 1000 tasks in total. We compute an averaged number of turns to measure the speed of consensus within each behavioral run. As an opinion formation task terminates after 10 rounds of interactions, the upperbound of an averaged number of turns is 10. The remaining parameters for Q-learning are {Discount=0.5, Step size=0.1, ϵ termination=0.005, ∇ illegal action=-100}. Table 3 compares these two performance metrics for trainees with different goal profiles. Trainees with MA-style significantly outperforms IA-style trainees for both performance metrics.

Table 3. Results of representative training

	Representative Training			
	Consensus Rate		# of Turns	
	Mean	Stddev	Mean	Stddev
MA-style	0.775	0.058	1.428	1.058
IA-style	0.503	0.098	3.441	7.742
	p=	*2.75E-162*	*p=*	*1.84E-157*

Even though trainees with MA-style perform better than IA-style trainees when representative training was received, the results from non-representative (improper) training tell a slightly different story. IA-style trainees on average ($\mu_{3t} = 0.997$, $\sigma_{3t} = 0.019$) significantly take fewer turns to reach consensus compared to MA-style trainees ($\mu_{4t} = 1.098$, $\sigma_{4t} = 0.022$) with $p = 6.4E - 92$. Despite the fact that MA-style trainees perform well when the type of training they receive matches well with the testing situation, IA-style trainees perform better at convincing unexpected types of testers. The results suggest that the performance observed in training is not representative of later performance. Thus, if sufficient training can be provided, MA-style trainees are preferred as they achieve better performance compared to IA-style trainees. On the other hand, IA-style trainees are preferred if sufficient training cannot be provided. This indicates the importance of representative training.

Conclusion

We have developed a probabilistic cognitive framework that can bridge the artificial gap between opinion formation and opinion change processes in current computational models. Our simulation results highlight how learning during opinion change may impact the upcoming decisions during opinion formation. Numerous areas of future work should be explored including more detailed studies on opinion stability, alternative opinion inferencing (heuristics), efficient algorithmic development, and larger scale simulations and human subject studies, such as in agent-human negotiations (Lin et al., 2014).

Acknowledgments

This work was supported in part by AFOSR Grant Nos. FA9550-15-1-0383, FA9550-09-1-0716, ONR Grant No. N00014-15-1-2154, ONR/NPS and DHS.

References

Carver, C. S. 1977. Self-awareness, perception of threat, and the expression of reactance through attitude change. *Journal of Personality* 45(4):501-512.

Clayton, M. J. 1997. Delphi: a technique to harness expert opinion for critical decision-making tasks in education. *Educational Psychology* 17(4):373-386.

Friedkin, N. E., and Johnsen, E. C. 2011. *Social influence network theory: a sociological examination of small group dynamics*: Cambridge University Press.

Hoekstra, V. J. 1995. The Supreme Court and Opinion Change An Experimental Study of the Court's Ability to Change Opinion. *American Politics Res* 23(1):109-129.

Kelman, H. C. 1961. Processes of opinion change. *Public Opinion Quarterly* 25(1):57-78.

Krueger, J. 1996. Personal beliefs and cultural stereotypes about racial characteristics. *J of Pers and Soc Psych* 71(3):536-548.

Lin, R.; Kraus, S.; and Mazliah, Y. 2014. Training with automated agents improves people's behavior in negotiation and coordination tasks. *Decision Support Sys* 60:1-9.

Martin, T. G.; Kuhnert, P. M.; Mengersen, K.; and Possingham, H. P. 2005. The power of expert opinion in ecological models using Bayesian methods: impact of grazing on birds. *Ecological Applications* 15(1):266-280.

McGuire, W. J. 1968. Personality and attitude change: An information processing theory. In *Psychological foundations of attitudes*, 171–196. Academic Press.

Pierro, A.; Mannetti, L.; Kruglanski, A. W.; Klein, K.; and Orehek, E. 2012. Persistence of attitude change and attitude–behavior correspondence based on extensive processing of source information. *Eur J of Social Psychology* 42(1):103-111.

Robinson, D. T.; Smith-Lovin, L.; and Wisecup, A. K. 2006. Affect control theory. In *Handbook of the Sociology of Emotions*.

Santos, E.; Wilkinson, J. T.; and Santos, E. E. 2011. Fusing multiple Bayesian knowledge sources. *International Journal of Approximate Reasoning* 52(7):935-947.

Sutton, R. S., and Barto, A. G. 1998. *Introduction to reinforcement learning*. MIT Press.

Watts, D. J., and Dodds, P. S. 2007. Influentials, networks, and public opinion formation. *J of Consumer Res* 34(4):441-458.

Weisbuch, G.; Deffuant, G.; Amblard, F.; and Nadal, J. P. 2003. Interacting agents and continuous opinions dynamics. In *Heterogenous Agents, Interactions and Economic Performance*.

Yildiz, E.; Acemoglu, D.; Ozdaglar, A.; Saberi, A.; and Scaglione, A. 2011. *Discrete opinion dynamics with stubborn agents*. SSRN eLibrary.

Yu, F. 2013. A framework for computational opinions. Ph.D. diss., Thayer Sch. of Eng., Dartmouth College, Hanover, NH.