Order of Magnitude Reasoning in Modeling Moral Decision-Making

Morteza Dehghani Emmett Tomai Ken Forbus Matthew Klenk
Qualitative Reasoning Group, Northwestern University
2133 Sheridan Road, Evanston, IL 60201 USA
{morteza, etomai, forbus, m-klenk}@northwestern.edu

Abstract

We present a cognitive model of moral decision-making, MoralDM, which models psychological findings about utilitarian and deontological modes of reasoning. Current theories of moral decision-making extend beyond pure utilitarian models by relying strongly on contextual factors that vary with culture. In MoralDM, the impacts of secular versus sacred values are modeled via qualitative reasoning, using an order of magnitude representation. We present a simplified version of ROM(R) (Dague, 1993) and discuss how it can be used to capture people's degree of quantity sensitivity. MoralDM uses a combination of first-principles reasoning and analogical reasoning to determine consequences and utilities of moral judgments. A natural language system is used to produce formal representations for the system from psychological stimuli, to reduce tailorability. We compare MoralDM against psychological results in moral decision-making tasks and show that its performance improves with experience.

Introduction

While traditional models of decision-making in AI have focused on utilitarian theories, there is considerable psychological evidence that these theories fail to capture the full spectrum of human decision-making. In particular, research on moral reasoning has uncovered a conflict between normative outcomes and intuitive judgments. This has led some researchers to propose the existence of deontological moral rules, which could block utilitarian motives. Consider the starvation scenario (from Ritov & Baron 1999) below:

> A convoy of food trucks is on its way to a refugee camp during a famine in Africa. (Airplanes cannot be used.) You find that a second camp has even more refugees. If you tell the convoy to go to the second camp instead of the first, you will save 1000 people from death, but 100 people in the first camp will die as a result.
>
> Would you send the convoy to the second camp?

While the utilitarian decision would send the convoy to the second camp, participants were more likely to choose to send the convoy to the first camp. Baron and Spranca (1997) suggested the existence of *sacred values*, which are not allowed to be traded off, no matter what the consequences. Further, they suggest that these sacred values "arise out of deontological rules about actions rather than outcomes". In our example, given that life is a sacred value, people often refuse to take an action which would result in taking lives.

This paper describes a cognitively motivated model of moral decision-making, called MoralDM, which operates in two modes of decision-making: utilitarian and deontological. MoralDM models the different impacts of secular versus sacred values via qualitative reasoning, using an order of magnitude representation. To reduce tailorability, a natural language understanding system is used to semi-automatically produce formal representations from psychological stimuli re-rendered in simplified English. MoralDM combines first-principles reasoning and analogical reasoning to implement rules of moral decision-making and utilize previously made decisions. We evaluate our system by comparing it with established psychological results and by examining how the performance of system changes as a function of the number of available cases.

We begin by summarizing relevant psychological results on quantity insensitivity and how an order of magnitude formalism can be used to capture this phenomenon. Next, we describe MoralDM and how it works. Then we show that MoralDM can account for results from two psychological studies, and that its performance can be improved by accumulating examples. Finally, we discuss related and future work.

Decision-Making and Quantity Insensitivity

In the presence of sacred values, people tend to be less sensitive to outcome utilities in their decision-making. This results in decisions which are contrary to utilitarian models. We claim that this can be accounted for using an existing qualitative reasoning formalism. After summarizing the relevant moral decision-making findings, we present a simplified version of Dague's (1993) ROM(R) qualitative order of magnitude formalism which we use to capture these results.

Sacred or protected values concern acts and not outcomes. When dealing with a case involving a protected value, people tend to be concerned with the nature of their action rather than the utility of the outcome. Baron and Spranca (1997) argue that when dealing with protected values people show insensitivity to quantity. That is, in trade-off situations involving protected values, they are less sensitive to the outcome utilities of the consequences. The amount of sensitivity (or insensitivity) towards outcomes vary with the context. Lim and Baron (1997) show that this effect varies across cultures.

In addition to contextual factors, the causal structure of the scenario affects people's decision-making. Waldmann and Dieterich (2007) show that people act more utilitarian, i.e., become more sensitive to the outcome utilities, if their action influences the patient of harm rather than the agent. They also suggest that people act less quantity sensitive when their action directly, rather than indirectly, causes harm. Bartels and Medin (2007) argue that the agent's sensitivity towards the outcome of a moral situation is dependent on the agent's focus of attention.

We model quantity sensitivity by using Dague's (1993) ROM(R) qualitative order of magnitude formalism. Order of magnitude reasoning is a form of commonsense reasoning which provides the kind of stratification that seems necessary for modeling the impact of sacred values on reasoning. Raiman (1991) uses the analogy of a coarse balance to describe the intuitions behind order of magnitude reasoning: a course balance can weigh quantities with more or less precision. This precision level depends on the order of magnitude scale used to map quantities onto course values. He uses two granularity levels *Small* and *Rough* to build a multitude of order of magnitude scales. These two granularity levels provide three qualitative relations between quantities which have been formally defined in FOG (Raiman, 1991). Both O(M) (Mavrovouniots and Stephanapoulos, 1987) and ROM(K) (Dague, 1993) are attempts to provide a more comprehensive order of magnitude formalism.

ROM(R), the mapping of ROM(K) onto $\Re$, is the only system that guarantees validity in $\Re$. Some order of magnitude representations (e.g. FOG) do not allow values at different levels to ever be comparable. One of the features of ROM(R) is that it includes two degrees of freedom, k_1 and k_2, which can be varied to capture differences in quantity sensitivity. Dague defines four classes of relationship between two numbers: "close to", "comparable to", "negligible with respect to" and "distant from". While FOG and O(M) fail to capture gradual change, the overlapping relations in ROM(K) allow a smooth, gradual transition between the states.

Although for engineering problems two degrees of freedom and four relations is quite useful, we believe for the task that we are interested in one degree of freedom and three binary relations are more plausible. Therefore, we implemented a simplified version of ROM(R) using one degree of freedom, k, resulting in three binary relations; almost equal, greater than, and orders of magnitude different. These three classes can be computed using the following rules:

- $A \approx_k B \Leftrightarrow |A-B| \leq k * \mathrm{Max}(|A|,|B|)$
- $A <_k B \Leftrightarrow |A| \leq k * |B|$
- $A \neq_k B \Leftrightarrow |A-B| > k * \mathrm{Max}(|A|,|B|)$

These relations respectively map to "close to", "greater than" and "distant from". k can take any value between 0 and 1. Figure 1 demonstrates the interval landmarks of the system. The analog in the above system of the parameter ε of ROM(K) depends on the value of k. When $k < ½$, ε is $k/(1 - k)$, and, when $k \geq ½$, ε is $(1 - k)/ k$. Quantity

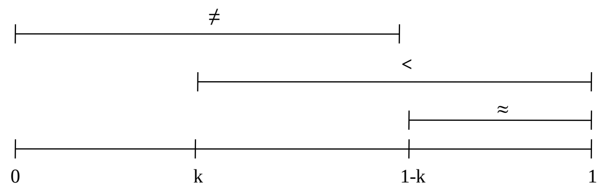

Figure 1: Interval Landmarks

sensitivity can be varied by changing k: setting k to $k - \varepsilon$ shifts the relationship between the compared values and moves it from ≈ to < or from < to ≠ resulting in higher quantity sensitivity. Depending on the sacred values involved and the causal structure of the scenario, we vary k to capture sensitivity towards the utility of the outcome.

MoralDM

Our model of moral decision-making, MoralDM, has been implemented using the FIRE reasoning engine and underlying knowledge base. The knowledge base contents are a 1.2 million fact subset of Cycorp's ResearchCyc knowledge base[1], which provides formal representations about everyday objects, people, events and relationships. The KB also includes representations we have developed to support qualitative and analogical reasoning. The scenarios, decisions and rule sets used in MoralDM are all represented uniformly and stored in this KB.

MoralDM operates in two mutually exclusive modes of decision-making: utilitarian and deontological. If there are no sacred values involved in the case being analyzed, MoralDM applies traditional rules of utilitarian decision-making by choosing the action which provides the highest outcome utility. On the other hand, if MoralDM determines that there are sacred values involved, it operates in deontological mode and becomes less sensitive to the outcome utility of actions, preferring inaction to actions that would cause harm.

To solve a given moral decision-making scenario, MoralDM begins by using a natural language understanding system, to automatically translate simplified English scenarios into predicate calculus. Given this representation, the Orders of Magnitude Reasoning (OMR) module calculates the relationship between the utility of each choice. Using the outcome of OMR, MoralDM utilizes a hybrid reasoning approach consisting of a First-

[1] http://research.cyc.com

Principles Reasoning (FPR) module and an Analogical Reasoning (AR) module to arrive at a decision. FPR suggests decisions based on rules of moral reasoning. AR compares a given scenario with previously solved decision cases to suggest a course of action. We believe using hybrid reasoning improves the robustness of the system and provides a more cognitively plausible approach to decision-making. Figure 2 depicts the MoralDM architecture.

FRP and AR work in parallel and complement each other by providing support (or disagreement) for a decision. If both succeed and agree, the decision is presented. When one module fails to arrive at a decision, the answer from the other module is used. If the modules do not agree, the system selects FPR's choice. If both fail, the system is incapable of making a decision. After a decision is made for a given scenario, it can be stored in the case library for future use. This enables the system to make decisions in more scenarios as it accumulates experience. Next, we discuss each module in detail.

Explanation Agent NLU

Our inputs are dilemmas from the psychological literature, expressed in natural language. To construct formal representations of these stimuli, we extended the Explanation Agent Natural Language Understanding system (EA NLU, Kuehne and Forbus, 2004). Unrestricted automatic natural language understanding is currently beyond the state of the art. Consequently, EA NLU uses a controlled language and operates semi-automatically, enabling experimenters to interactively translate natural language stimuli into simplified syntax and guide the generation of predicate calculus. This practical approach allows us to broadly handle syntactic and semantic ambiguities and to build deep formal representations suitable for complex reasoning. This is a significant advantage over having experimenters construct representations by hand for two reasons. First, constructing representations by hand is very time-consuming and requires substantial expertise. Second, hand-coding increases tailorability, i.e., the possibility that representation choices were made to get a particular example to work, as opposed to being uniform, independently motivated conventions. Since EA NLU is used by multiple projects and relies on an off-the-shelf knowledge base, tailorability is greatly reduced.

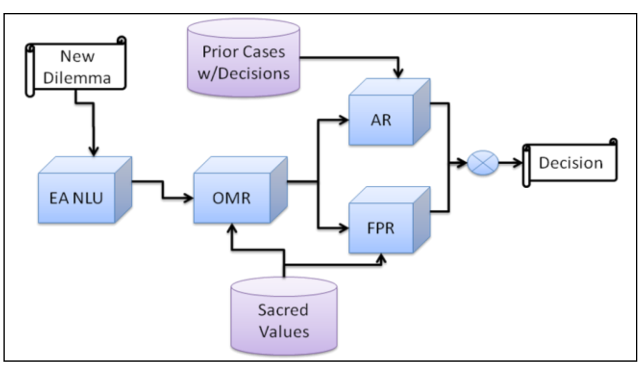

Figure 2: MoralDM Architecture

EA NLU uses Allen's bottom-up chart parser (Allen, 1995) in combination with the COMLEX lexicon (Macleod *et al.* 1998) and a simplified English grammar (Kuehne and Forbus, 2004). The parser uses subcategorization frames from ResearchCyc for word and common phrase semantics. Each frame represents a case for the term encoded as predicate calculus with syntactic/semantic role variables. Roles are filled during the parsing process. Frames are filtered according to both case constraints and syntactic requirements explicitly included in the frame.

Sentences within a stimulus are parsed separately. The resulting parse trees are presented, together with the semantic frames they entail, to the user in an interactive interface. The user can selectively include or exclude trees as well as individual frames. These selections serve as input to a transformation process using dynamic logic principles from Discourse Representation Theory (DRT) (Kamp & Reyle 1993) to construct a description of the sentence content. This description supports numerical and qualitative quantification, negation, implication, modal embedding and explicit and implicit utterance sub-sentences. Anaphoric references are resolved via selectional restrictions from the Cyc ontology guiding sentence attachment and thereby the integration into the representation of the discourse as a whole.

```
A convoy of trucks is transporting food
to a refugee camp during a famine in
Africa. 1000 people in a second refugee
camp will die. You can save them by
ordering the convoy to go to that
refugee camp. The order will cause 100
people to die in the first refugee camp.
```

Figure 3: Starvation scenario in simplified English

Figure 3 contains the controlled language for the starvation scenario. Given these statements, EA NLU identifies events of transporting, famine, dying (1000 people), saving, ordering, going and dying (100 people) together with the two quantified sets of people, the convoy, food, two refugee camps and the proper name Africa. There is also an explicit reference to the listener, "you". Figure 4 contains the frame-based interpretation of the order.

Causal links are explicitly stated between the order and the saving and the order and the second set of deaths. The abstraction of saving drives inferential attention to events in the description that the beneficiary may be being saved from. The expected future modality of the first set of deaths makes it a reasonable candidate. Based on the possible modality of the saving/ordering sequence,

```
(isa order131049 Ordering-CommunicationAct)
(performedBy order131049 you128898)
(recipientOfInfo order131049 convoy127246)

(infoTransferred order131049
 (and
  (isa refugee-camp129739 RefugeeCamp)
  (isa convoy127246 Convoy)
  (isa go129115 Movement-TranslationEvent)
  (primaryObjectMoving go129115 convoy127246)
  (toLocation go129115 refugee-camp129739)))
```

Figure 4: Predicate calculus for ordering

combined with the use of the explicit reference to the listener, the system infers an abstraction of choice being presented with known consequences resulting from both action and inaction. Figure 5 contains the inferred abstraction of choice and its causal consequences.

```
(isa Sel131949 SelectingSomething)
(choices Sel131949 order131049)
(choices Sel131949 Inaction131950)
(causes-PropSit
  (chosenItem Sel131949 Inaction131950)
  die128829)
(causes-PropSit
  (chosenItem Sel131949 order131049)
  save128937)
```

Figure 5: Predicate calculus for the choice presented

Order of Magnitude Reasoning Module

The inputs to OMR include the sacred values for the culture being modeled and the causal structure of the scenario. Using the predicate calculus produced by EA NLU, OMR calculates the expected utility of each choice by summing the utility of its consequences. For each consequence of a choice, OMR uses its rules to ascertain if the outcome is a positive or negative outcome, and to identify any sets whose cardinality matters in the decision (e.g., number of people at risk).

After computing utilities, OMR selects a k value based upon the context of the scenario. Assuming that the relationship between the utilities, a and b, are "comparable", MoralDM sets k to $1 - (|\ a\ /\ b\ |)$. This results in the relationship between the utilities falling within <, right between ≠ and ≈ (Fig 1). If the decision involves a sacred value for the modeled culture, setting k to $k + \varepsilon$ shifts the relationship between utilities from greater than to close to, resulting in the system being less sensitive to the numeric utility of the outcome. On the other hand, if the there are no sacred values involved, the system substitutes k with $k - \varepsilon$ thereby making the system more quantity sensitive to the computed utilities. In addition to sacred values, the causal structure of the scenario affects k. OMR checks to see if the scenario contains patient intervention or agent intervention. It uses low quantity insensitivity for the first case and high otherwise, consistent with psychological findings (Waldmann and Dieterich 2007). The system also checks for direct versus indirect causation. In the case of indirect causation, a higher degree of insensitivity is applied.

Returning to the starvation scenario, there are two choices: ordering and inaction. For ordering, there are two consequences, 1000 people in the second camp will be saved and 100 people in the first camp will die. Consulting the KB, the system determines that dying has negative utility and saving positive, resulting in a choice utility of 900 for the ordering choice. Using the same procedure, the utility for inaction is calculated to be -900. Using the formula given above, k is initially set to 0 with $\varepsilon = 1$. Given that both choices involve agent intervention and indirect causation, there are no structural differences between the two choices. Therefore, the k value is set solely by the existence of sacred values. In this case, causing someone to die is a sacred value resulting in k being set to $k + \varepsilon = 1$, therefore causing the system to act less quantity sensitive. Using ROM(R), the relationship between the utilities of the two choices is calculated to be ≈. On the other hand, if there had not been a sacred value, the value of k would have remained 0 causing the relationship between the utilities to be ≠. These utilities, 900 and -900, and the computed relationship, ≈, are provided to FPR and AR.

First-Principles Reasoning Module

Motivated by moral decision-making research, FPR makes decisions based upon the following factors: the orders of magnitude relationship between utilities, sacred values, computed utilities, and action vs. inaction. FPR uses three methods for making decisions. First, the utilitarian method, which selects the choice with the highest utility, is invoked when the choice does not involve a sacred value. Second, in situations with sacred values and without an order of magnitude difference between outcomes, the pure-deontological method selects the choice that does not violate a sacred value. Third, the utilitarian-deontological method operates when the scenario contains sacred values and an order of magnitude difference between outcomes, selecting the choice with the higher utility. Therefore, the pure-deontological method is the only method that makes decisions that violate utilitarian norms.

In the starvation scenario, there is a sacred value, people dying, and no order magnitude difference between the utility of the two choices. Therefore, the system uses the pure deontological method to select the inaction choice.

These methods are mutually exclusive, returning at most one choice per scenario. Given the breadth of moral reasoning scenarios, the rules implementing FPR are not complete. Therefore, FPR necessarily fails on some scenarios. These cases highlight the need for the hybrid-reasoning approach taken in MoralDM. The resulting choice is compared with the results of the analogical reasoning module of MoralDM.

Analogical Reasoning Module

An important role that analogy plays in decision-making is framing the situation. When making a choice, decision makers frequently use past experiences and draw inferences from their previous choices (Markman and Medin, 2002). For more details and examples about the use of analogy in decision-making please refer to Dehghani et al. (2008a). To model analogy in decision making, we use the Structure-Mapping Engine (SME) (Falkenhainer *et al.* 1989), a computational model of similarity and analogy based on Gentner's (1983) structure mapping theory of analogy in humans. SME operates over structured representations, consisting of entities, attributes of entities and relations. Given two descriptions, a *base case* and a *target case*, SME aligns their common structure to find a mapping between the cases. This mapping consists of a set of correspondences between entities and expressions in the two cases. SME produces mappings that maximize *systematicity*; i.e., it prefers mappings with higher-order relations and nested relational structure. The *structural evaluation score* of a mapping is a numerical measure of similarity between the base and target. SME identifies elements in the base that fail to map to the target and uses the common relational structure to calculate *candidate inferences* by filling in missing structures in target.

Running concurrently with FPR, AR uses comparisons between new cases and previously solved cases to suggest decisions. When faced with a moral decision scenario, AR uses SME to compare the new case with every previously solved scenario in its memory. The similarity score between the novel case and each solved scenario is calculated using SME by normalizing the structural evaluation score against the size of the scenario. If this score is higher than a certain threshold and both scenarios contain the same order of magnitude relationship between outcome utilities, then the candidate inferences are considered as valid analogical decisions. If the scenarios have different orders of magnitude relationships, it is likely that a different mode of reasoning should be used for the target scenario and AR rejects the analogical inference. After comparing against all of the solved scenarios, AR selects the choice in the new scenario with the highest number of analogical decisions. In the case of a tie, AR selects the choice with the highest average similarity score supporting it. Because analogical alignment is based upon similarities in structure, similar causal structures and/or sacred values align similar decisions. Therefore, the more structurally similar the scenarios are, the more likely the analogical decision is going to be the correct moral one.

Returning to our starvation example, AR can solve this decision problem through an analogy with a traffic scenario given below, in which the system chose to not transfer funds:

> A program to combat accidents saves 50 lives per year in a specific area. The same funds could be used to save 200 lives in another area, but the 50 lives in the first area would be lost.
>
> Do you transfer the funds?

The analogical decision is determined by the candidate inferences where the decision in the base, inaction, is mapped to the choice in the target representing inaction. Because the traffic scenario contains the same the order of magnitude relationship, almost equal, as in the starvation scenario, the system accepts the analogical decision.

Evaluation

We evaluated MoralDM by running it on 8 moral decision-making scenarios taken from two psychology studies (Waldmann and Dieterich 2007; Ritov and Baron 1999). In all the scenarios used, traditional utility theories fail to predict subjects' responses, as often the subjects choose the choice which provides a smaller overall outcome utility. We compare MoralDM's decisions to subjects' responses in these experiments. If the decision of MoralDM matched those of the subjects, as reported by the authors, we consider it a correct choice.

EU NLU translated all 8 cases into predicate calculus. MoralDM made the correct choice in each of the scenarios using the result from FPR. This illustrates MoralDM's ability to do complex reasoning from natural language input and provides evidence for its psychological fidelity.

One of the more difficult aspects in building the FPR module is the number of rules to handle the broad range of situations covered in moral decision making. The AR module is capable of making moral decisions in situations when gaps in the KB or rule set would prevent the FPR module from coming up with an answer. Therefore, we evaluated the AR module independently of the FPR module, to answer two questions: (1) Can we use analogy to do moral decision-making from natural language input? (2) How is AR performance affected as the number of previously solved cases stored in memory increases?

Given the 8 solved scenarios, we created case libraries of every combination of these scenarios. This provided us with 254 different case libraries (8 of size 1, 28 of size 2, 56 of size 3…). Then, with each case library, we tested the AR module by running it on each of the scenarios not in the case library. So for each of the 8 libraries of size 1, the test consisted of 7 decision scenarios for a total of 56 decision scenarios.

Figure 6 shows the performance of AR as a function of the number of available cases. There is a monotonic increase in the number of correct answers as the size of the library increases (Pearson's $r = .97$, $p < .0001$). Also, there is a significant decrease in the number of cases where AR does not come up with an answer ($r = -.95$, $p < .001$). The number of incorrect decisions changes insignificantly from 18% to 25% ($r = .53$, $p < .22$). The statistics reported have been computed by comparing each series against the size of the case library.

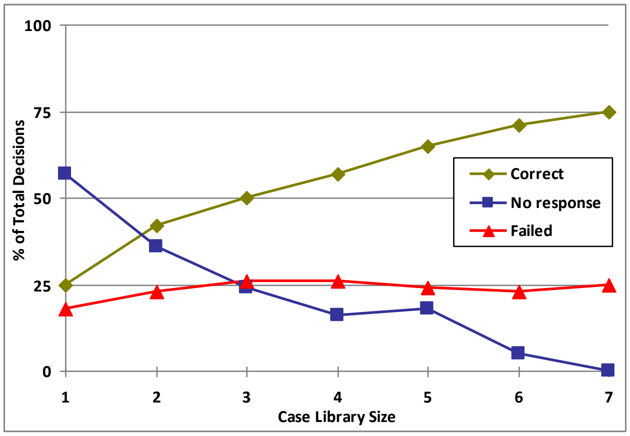

Figure 6: Analogical reasoning results

The results of these evaluations are very encouraging. First and foremost, our system matches human behavior on 8 decision-making scenarios provided in natural language. In addition to this result, we also found that there was a significant improvement in AR module performance as the number of cases in MoralDM's memory increased. For a more comprehensive analysis of evaluation of each of the modules please refer to Dehghani et al. (2008b).

Related Work

Reasoning with orders of magnitude is a form of commonsense reasoning which is mostly suitable when complete quantitative information is not available or when tackling problems involving complex physical systems. Order of magnitude reasoning has been used in several engineering tasks (e.g. Dague, 1994; Mavrovouniots and Stephanapoulos, 1988; Dague, Deves and Raiman 1987).

Several research projects have focused on building ethical advisors. The MedEthEx system uses ILP techniques to learn decision principles from training cases (Anderson *et al.* 2006). Mclaren's Truth-Teller and SIROCCO systems (2005) use case-based reasoning to highlight relevant ethical considerations and arguments to a human user. Like them, we use prior cases, but to guide the system's own reasoning, rather than give advice. They also were not designed to model the effects of sacred versus secular values that MoralDM captures.

Computational models of cultural reasoning are receiving increasing attention. For example, the CARA system (Subrahmanian *et al.* 2007) is part of a project to "understand how different cultural groups today make decisions and what factors those decisions are based upon". CARA uses semantic web technologies and opinion extraction from weblogs to build cultural decision models consisting of qualitative rules and utility evaluation. While we agree that qualitative reasoning must be integrated with traditional utility evaluation, we also believe that analogy plays a key role in moral reasoning. Moreover, we differ by evaluating our system against psychological studies, which helps ensure its judgments will be like those that people make.

Our combination of analogical and first-principles reasoning is inspired in part by Winston's (1982) use of both precedents and rules to reason about a situation. His work was hampered by the lack of off-the-shelf large-scale knowledge bases, and the technologies for NLU and analogical reasoning have improved since then.

Our use of simplified English is inspired by both CMU's KANT project (cf. Mitamura & Nyberg 1995) and Boeing's controlled language work (cf. Clark *et al.* 2005).

Conclusions and Future Work

MoralDM uses qualitative modeling to reason about utilities, capturing the differences between sacred and secular values via an order of magnitude representation. It uses a combination of first-principles logical reasoning and analogical reasoning to determine the utility of outcomes and make decisions based on this information. The hybrid approach produces answers in a wider range of circumstances than either alone. Natural language input of scenarios, in simplified English, reduces tailorability, a key problem in cognitive simulation research. We showed that MoralDM can be used to model psychological results from two studies. While there is still more to be done, we think MoralDM represents an important step in computational modeling of moral decision-making.

We plan to pursue several lines of investigation next. First, we plan to extend the valuation rules to model different cultures, based on existing collaborations with cognitive psychologists and anthropologists. This will require extending the first-principles reasoning rules to cover a broader range of scenarios. Constructing these rules is a time consuming and error prone process. One alternative is to automatically extract rules by generalizing over previously made decisions. By focusing on decisions from a specific culture, we can explore automatic model construction for making novel predictions about the behavior of a certain group (Dehghani et al. 2007). Second, we plan to extend the range of EA NLU coverage to handle a wide range of cultural stories. This will enable us to create story libraries for different cultural groups, and translate transcripts from interview data more easily. Third, we plan to incorporate a cognitively plausible model of similarity-based retrieval, MAC /FAC (Forbus et al., 1995), to make analogical reasoning more scalable as the story library grows. Finally, we plan to test MoralDM on a wider range of problems, using data gathered from participants from multiple cultural groups.

References

Allen, J. F. 1995. *Natural Language Understanding.* (2nd ed). Redwood City, CA.: Benjamin/Cummings

Anderson, M., Anderson, S., and Armen, C. 2006. An Approach to Computing Ethics. *IEEE Intelligent Systems*. 21(4): 56-63.

Bartels, D. M. & Medin, D. L. 2007. Are Morally-Motivated Decision Makers Insensitive to the Consequences of their Choices? *Psychological Science,* 18, 24-28.

Baron, J., and Spranca, M. 1997. Protected Values. *Organizational Behavior and Human Decision Processes* 70: 1–16.

Clark, P., Harrison, P., Jenkins, T., Thompson, J., and Wojcik, R. 2005. Acquiring and Using World Knowledge using a Restricted Subset of English. In *Proceedings of The 18th International FLAIRS Conference*.

Dauge, P. 1993. Symbolic Reasoning with Relative Orders of Magnitude. In *Proceedings of the 13th IJCAI.*

Dauge, P. 1993. Numeric Reasoning with Relative Orders of Magnitude. In *Proceedings of the 7th International Workshop on Qualitative Reasoning*.

Dague, P., Deves, P., and Raiman, O. 1987. Troubleshooting: when Modeling is the Trouble, In *Proceedings of AAAI87 Conference*, Seattle.

Dague, P. 1994. Model-based Diagnosis of Analog Electronic Circuits. *Annals of Mathematics and Artificial Intelligence*,11, 439-492

Dehghani, M., Tomai, E., Forbus, K., Iliev, R., Klenk, M. (2008a). MoralDM: A Computational Modal of Moral Decision-Making. To appear in Proceedings of the 30th Annual Conference of the Cognitive Science Society (CogSci), Washington, D.C

Dehghani, M., Tomai, E., Forbus, K., Klenk, M. (2008b). An Integrated Reasoning Approach to Moral Decision-Making. To appear in Proceedings of the Twenty-Third AAAI Conference on Artificial Intelligence (AAAI). Chicago, IL.

Dehghani, M., Unsworth, S., Lovett, A., Forbus. K. (2007) Capturing and Categorizing Mental Models of Food Webs using QCM. *21st International Workshop on Qualitative Reasoning*. Aberystwyth, U.K.

Falkenhainer, B., Forbus, K. and Gentner, D. 1989. The Structure-Mapping Engine: Algorithms and Examples. *Artificial Intelligence,* 41: 1-63.

Forbus, K., Gentner, D. and Law, K. 1995. MAC/FAC: A Model of Similarity-based Retrieval. *Cognitive Science*, 19(2), 141-205.

Gentner, D. 1983. Structure-Mapping: A Theoretical Framework for Analogy. *Cognitive Science*, 7: 155-170.

Kuehne, S. and Forbus, K. 2004. Capturing QP-Relevant Information from Natural Language Text. *Proceedings of QR04.*

Lim, C. S, and Baron, J. 1997, Protected values in Malaysia Singapore, and the United States. Manuscript, Department of Psychology, University of Pennsylvania.

Markman, A and Medin, D.L. 2002. *Decision Making.* Stevens Handbook of Experimental Psychology, 3rd edition: Volume 2, Memory and Cognitive Processes. New York: Wiley.

Macleod, C., Grishman, R., and Meyers, A. 1998. COMLEX Syntax Reference Manual, Version 3.0. Linguistic Data Consortium. University of Pennsylvania: Philadelphia, PA.

Mavrovouniotis M. L. and Stephanopoulos, G. 1987, Reasoning with orders of magnitude and approximate relations, In *Proceedings of AAAI 1987 Conference*, Seattle.

Mavrovouniotis M. L. and Stephanopoulos, G. 1987, Order of Magnitude Reasoning in Process Engineering, Computer. chem. Engng. 12, 1988 .

McLaren, B. 2005. Lessons in Machine Ethics from the Perspective of Two Computational Models. In Anderson, M. *et al.* (Eds.) *Machine ethics: Papers from the AAAI Fall Symposium*, Technical Report FS-05-06. Menlo Park, CA.

Mitamura, T., & Nyberg, E. H. 1995. Controlled English for Knowledge-Based MT: Experience with the KANT System. In *Proceedings of 6th International Conference on Theoretical and Methodological Issues in Machine Translation*, Leuven, Belgium.

Raiman, O. 1991. Order of magnitude reasoning. *Artificial Intelligence* 51.

Ritov, I. and Baron, J. 1999. Protected Values and Omission Bias. *Organizational Behavior and Human Decision Processes*, 79(2): 79-94.

Subrahmanian, VS., Albanese, M., Martinez, V., Nau, D., Reforgiato, D., Simari, G., Sliva, A., and WIlkenfeld, J. 2007. CARA: A Cultural Adversarial Reasoning Architecture. *IEEE Intelligent Systems.* 22(2): 12-16.

Waldmann, M. R., and Dieterich, J. 2007. Throwing a Bomb on a Person Versus Throwing a Person on a Bomb: Intervention Myopia in Moral Intuitions. *Psychological Science*, 18 (3), 247-253.

Winston, P.H. 1982. Learning New Principles from Precedents and Exercises. *Artificial Intelligence* 19(3), 321-350.