A Bayesian Model of Four-Factor Theory of Emotion: Case Study in Medical Shared Decision-Making

Abtin Mogharabin (abtin.mogharabin@metu.edu.tr)
Department of Statistics, Middle East Technical University

Abstract

Decision-making, particularly in high-stakes environments such as healthcare and emergency response situations, is heavily influenced by emotional and pragmatic factors. However, past methods in Bayesian models of Mind mainly focused on the pragmatic aspect of rational reasoning, often neglecting the critical role that emotional and contextual influences play in shaping decisions under pressure. As a result, they are unable to suitably represent phenomena such as life changing medical decisions under pressure. In this paper, we extend the rational speech acts (RSA) framework to model the dynamics of high-stake shared decision-making (SDM). For this purpose, we use various strategies to incorporate the effects of emotions, as introduced in the four-way theory of emotion in decision-making by Pfister and Böhm (2008), in the traditional RSA models. Particularly, we expand this framework by introducing the *multi-agent listener/observer* model for scenarios consisting of more than two agents. This way, we incorporate the effect of acquaintances' actions and suggestions in the decision maker's decisions – something that, to the best of our knowledge, has not been investigated by any of the previous literature. In the end, we provide a case study of medical SDM for a cancer patient about the post-surgery treatment options. We demonstrate how the suggestions of the doctor and family members and the patient's emotions might affect the patient's decision.

Keywords: Shared Decision Making, Bayesian Theory of Mind, Rational Speech Acts, Multi-Agent models

I. Introduction

"The heart has its reasons which reason knows nothing of" — Blaise Pascal, Pens´ees

Introducing the prospect theory, Kahneman and Tversk (1979) criticized the expected utility theory (Neumann and Morgenstern, 1947) by analyzing the human decisions under risk. By experimental methods, they rejected the necessity that rational agents need to maximize utility in all scenarios. They demonstrated that in every decision, individuals evaluate potential losses and gains relative to a reference point, in other terms, their perception, rather than in absolute terms. Just like how Richard Lazarus, in his book *Psychological Stress and the Coping Process* (1966) defined emotions as a relationship between the person and the environment, triggered and shaped by how individuals subjectively assess the personal significance of a situation, goal or an action. Thus, introducing the cognitive appraisal theory.

Building upon many of the previous literature, Pfister and Böhm (2008) analyzed the past theories of emotion. They argued that a majority of past literature on the effect of emotion in decision making suffer of three main problematic assumptions: (1) in the irrational view of emotions, decision making is a purely rational mental process that is jeopardized by emotions. However, evidence is accumulating that this conception might be false (Bechara & Damasio, 2005; Bechara, Damasio, Tranel, & Damasio, 1997). In fact, without emotional involvement, decision making might be far from optimal (Damasio, 1994). (2) in the positive-negative emotions, all emotional states are assumed to be mapped onto a one-dimensional scale of valence, characterized by contrasting labels such as positive versus negative (Barrett, 2006b; Russell, 2003). Pfister and Böhm (2008), similar to Solomon and Stone (2002), point out that classifying all emotions into pure positive/negative groups might be impossible. For example, what is good or beneficial need not be pleasurable, and what is harmful might nevertheless be satisfying. (3) The assumption of emotions as a homogenous category, where emotions are taken as a single category of something real and natural, not as a conceptual construction (Charland, 2002; Boyd, 1999; Griffiths, 2004). But more recent works, demonstrate evidence that generalizations are based on similarity by analogy, but not on homology, that is, no general regularities, physiological, neurological, or behavioral, can be reliably identified that are common to and essential for all emotions (as discussed in Griffiths, 1997, 2004). In addition, Barrett (2006a) demonstrates that particular emotions, such as anger and sadness, do not typically exhibit a homeostatic property cluster.

Taking the above evidence into account, Pfister and Böhm (2008) introduced a four-way theory of emotion in decision making that not only follows many of the strengths from previous theories, but also avoids the abovementioned three key criticism. Their approach distinguished the effect of emotional phenomena on decisions in four distinct functions:

The information function is associated with joys and (dis)likes and provides information about the pleasure and pain associated with certain actions or outcomes; the speed function allows quick decisions under time constraints, and is amplified by fear and disgust; the relevance function, affected by regret and disappointment, allows to focus attention on relevant aspects of a decision; and the commitment function, amplified by guilt, love, and anger, helps sustain decisions that align with social and moral values. In the computational modeling section, we will further elaborate on the four functions of emotion, and discuss our interpretations regarding them.

On the computational modeling side, Bayesian models of cognition have shown considerable success in modeling human learning and reasoning as inference in complex probabilistic models (Griffiths and Tenenbaum, 2007; Griffiths, Kemp and Tenenbaum, 2008). One of the most influential models on pragmatic reasoning and communicative decision making is Rational Speech Acts (RSA), a model defined to accurately replicate a wide variety of pragmatic phenomena using utility maximization methods. (Frank and Goodman, 2012; also see Degen 2023). As will be discussed in our subsequent sections, RSA provides one of the most suitable frameworks to represent the four-way theory of emotions. In fact, we demonstrate that many pre-existing RSA models seamlessly incorporate the effects of the information function.

In this paper, we extend upon the traditional methods and introduce a new RSA variation for shared decision making (SDM). SDM typically takes place in medical decision making between a physician and a patient. In SDM, an advisor and a patient (decision maker) collaborate towards a joint decision (Elwyn et al., 2012; Elwyn et al., 2015; Elwyn et al., 2021). The advisor has more knowledge about the (very uncertain) medical condition (e.g. a disease) and its outcomes but they don't completely know the patient's preferences (e.g. interests and emotions). Wrong decision might result in significant health issues for the patient in the future, even risking their lives in cases. Thus, SDM scenarios often involve time constraints and risks s, making it challenging to model them without constructing a detailed framework of human emotion, appraisal, and communication. In the following sections, we incorporate several relevant elements from the four-way theory of emotions (Pfister & Böhm, 2008) and construct a Bayesian model of SDM by extending the well-defined structure of the Rational Speech Acts (RSA) framework.

II. Relevant Works

Extending the previous applications of Bayesianism to object perception (Kersten et al., 2004) and language acquisition (Chater & Manning, 2006), a new interest to apply Bayesian models to human emotions have begun. Unfortunately, the vast majority of these models focus on third-person models, following the cognitive appraisal theory (e.g., Saxe & Houlihan, 2017; Ong et al., 2019), instead of focusing on the effects of one's own emotions on actions. Perhaps, the most relevant work to our model is the Bayesian Drift-Diffusion Model introduced by Ying et al. (2022), where they designed a model of Schachter-Singer's Two Factor Theory (1962).

Utilizing Drift-Diffusion Models, Ying et al. (2022) introduced a method to model emotions as the outcome of cognitive labeling, or attribution of a diffuse pattern of autonomic arousal. Their method, utilizing random walks, modeled a decision-maker who accumulates evidence until a relative decision value meets one of the two decision boundaries. A choice is made corresponding to the boundary being crossed. Then, the corresponding choice is selected to be the resulting decision (as discussed in Fudenberg et al., 2020). They investigated two model scenarios for each of the two aspects of Two Factor Theory. Unfortunately, their method separated the two components of the Two-Factor Theory of Emotion in two separate studies. Which made it unideal for more realistic models where arousal and context (to factors of the Two Factor Theory) interact.

III. Computational Models

Rational Speech Acts

The Rational Speech Act (RSA) framework, as defined by Goodman and Frank (2012), views communication as a process of recursive reasoning between a speaker and a listener. The listener interprets the speaker's utterance by reasoning about a cooperative speaker who is trying to inform a naive listener about some state of affairs. Using Bayesian inference, the listener reasons about what the state of the world is likely to be, given that the speaker produced an utterance, and knowing that the speaker is reasoning about how the listener is most likely to interpret that utterance. There are multiple levels of inference: at the top, the sophisticated, pragmatic listener reasons about the pragmatic speaker, who selects the best utterance by maximizing the a utility function over the literal listener.

The literal listener, L_0, interprets an utterance u based on its literal meaning. The probability that a literal listener L_0 infers a state s given an utterance u is given by

$$P_{L0}(s|u) \propto [\![u]\!](s) \cdot P(s) \quad (1)$$

where $[\![u]\!](s)$ denotes the truth conditions of the utterance u in state s, and $P(s)$ is the prior probability of state s.

The pragmatic speaker, S_1, chooses (makes a decision) an utterance u to convey a state s by maximizing the utility of u. The probability that a pragmatic speaker S_1 chooses an utterance u given a state s is defined as

$$P_{S1}(u|s) \propto \exp(\alpha \cdot U_{S1}(u; s)) \quad (2)$$

where α is the rationality parameter that controls the speaker's sensitivity to utility differences, and $U_{S1}(u;s)$ is the utility function representing the speaker's preference for utterance u given state s. Where the utility allows speaker to decide on the utterances are preferred (liked by the speaker) at communicating, while taking a cost for each choice (unlikability of the choice for the speaker). In other words:

$$U_{S1}(u; s) = \log L_0(s|u) - C(u) \quad (3)$$

The pragmatic listener, L_1, then interprets an utterance u by updating their beliefs about the possible states s based on the likelihood of u being chosen by a rational speaker.

$$P_{L1}(s|u) \propto P_{S1}(u|s) \cdot P(s) \quad (4)$$

It is worth noting that, despite its promise as a model of conversational reasoning, it has rarely been used to model more than a single conversation. As clear from the abovementioned definition, RSA, in its most basic form, is defined as a model of a single agent's recursive reasoning. One of the few efforts of adapting RSA for multi-agent conversations was done by Anderson (2021). The demonstrated a system that iteratively updates conversational states, including the common ground and participants' beliefs, as they interact over multiple conversational turns. In each iteration, the agents' roles switches between being the speaker and the listener.

In Anderson's framework (2021), the posterior output of each speaker model over the utterances was used to linearly update the prior distribution of the other agent's listener model over the world states (i.e., the common ground update), based on a learning rate. Mathematically, this common ground update could be represented as follows:

$$CG_{L1(i+1)}(s)=(1 - lr) \cdot CG_{L1(i)}(s) + lr \cdot P_{Li}(s \mid u) \quad (5)$$

where $CG_{L1(i+1)}$ represents the Common of the (i+1)th iteration's listener model for the current speaker agent, $CG_{L1(i)}$ represents the current ith iteration's prior Common Ground of the speaker agent. And P_{Li} represents the posterior of the listener model for the current iteration's speaker agent, and contains the probability distribution of the world states after processing the speaker's utterance u. The learning rate determines the updated posterior's weight compared to the prior.

Information Function

The information function of emotions (Pfister and Böhm, 2008) provides essential evaluative information for decision-making by indicating pleasure or pain associated with different options. Emotions, such as joy or distress, inform individuals about the desirability or undesirability of potential outcomes, guiding their preferences and choices. We consider an SDM scenario where an advisor (S_1) provides suggestions on the best treatment options to the patient (the listener), taking into account both the likability and unlikability of each option. In this case, we denote the treatment option as t. By representing the advisor's undesirability as $D_{S1}(t)$, we redefine the model to better represent the information function.

$$P_{L0}(s| t) \propto [\![t]\!](s) \cdot P(s) \quad (6)$$

$$P_{S1}(t| s) \propto \exp(\alpha \cdot U_{S1}(t; s)) \quad (7)$$

$$U_{S1}(t; s) = \log L_0(s| t) - D_{S1}(t) \quad (8)$$

$$P_{L1}(s| t) \propto P_{S1}(t| s) \cdot P(s) \quad (9)$$

$$CG_{L1(i+1)}(s)=(1 - lr) \cdot CG_{L1(i)}(s) + lr \cdot P_{L1(i)}(s \mid t) \quad (10)$$

Here, $P_{L0}(s| t)$ represents the likability of each treatment for the advisor, i.e., the probability of a treatment to positively improve the patient's state. The unlikability of each treatment, $D_{S1}(t)$, is evaluated as a linear combination of the weights assigned by the advisor to each potential side effect of the treatment.

Speed Function

The speed function (Pfister and Böhm, 2008) of emotions emphasizes the role of affective responses in enabling rapid decision-making under time constraints. It suggests that certain emotions, like fear and disgust, act as rapid response mechanisms to stimuli that require immediate action. Research on bounded rationality has similarly highlighted that humans often rely on heuristic methods and simplified decision processes when faced with limited time (Herbert, 1955, 1990, 1991). To replicate this effect, we theorize that under pressure and stress, agents assign greater weights to the most significant outcomes and side effects of each treatment, while disregarding less critical factors. Consequently, they seek a satisfactory option rather than the optimal one. In mathematical terms, the advisor under time pressure is assumed to have a utility of $W_{S1}(t) * U_{S1}(t; s)$ for treatments.

Relevance Function

The relevance function (Pfister and Böhm, 2008) is integral to how emotions influence decision making by focusing the decision maker's attention to the most significant aspects of a situation based on their personal appraisals and concerns.

In this context, the relevance function influences how both the advisor and the patient perceive the situation. The advisor might deem the situation irrelevant to themselves, leaving them unaffected by this function. Conversely, the patient would find this decision to be highly relevant. Depending on the patient's values, their decision-making process would be impacted. For example, if choosing the wrong treatment could endanger the patient's life, the relevance function might lead the patient to focus more on safer options, knowing they would regret selecting a riskier one, even if the risk is minimal. To quantify this function, we introduce the *relevance weight*. This weight, directly adjusts the side effects' cost terms for each agent based on what they consider most important and relevant.

Commitment Function

The commitment function (Pfister and Böhm, 2008) pertains to the role of emotions in sustaining long-term decisions and fostering social and moral behavior, even when these actions contradict their original self-interest. Emotions such as guilt, shame, love, and anger serve as commitment devices. These emotions provide the internal motivation to act ethically, even when it might be more beneficial in the short term to act selfishly.

To avoid overcomplication, we mainly incorporate some aspects of love and its associated social and moral norms. As discussed by Ho (2008), there have been observations of patients being influenced by their family members and changing their minds under the influence of their loved ones. Previous work has examined whether such effects are detrimental to patients (Breslin, 2005).

To model these effects, we introduce a third agent into our system, representing a family member of the patient. By generating a mind model of their family members, the patient assigns a weight to their family member in the decision-making process. The patient's mind model of the family member acts as a second pragmatic speaker model (S_2), making suggestions after considering the doctor's advice. To accommodate this, we extend the traditional RSA model with a new variation of pragmatic listener models, allowing each listener to consider the suggestions of multiple sources simultaneously. The family agent is defined as follows:

$$P_{S2}(s \mid t) \propto P_{S1}(t \mid s) \cdot P(s) \quad (11)$$

$$P_{S2}(t \mid s) \propto \exp(\alpha \cdot W_{S2}(t) * U_{S2}(t; s)) \quad (12)$$

$$U_{S2}(t; s) = (1-\phi) \cdot \log L_0(s \mid t) + \phi \cdot \log P_{S1}(s \mid t) - D_{S2}(t) \quad (13)$$

where ϕ is the weight family member puts on the advisor's suggestions.

The patient (DM) then incorporates their mind model of the doctor, the family member, as well as their own self-interests, to make the final decision by utility maximization.

$$P_{DM}(s \mid t) \propto P(s) \cdot P(t \mid s) \quad (14)$$

$$P_{DM}(t \mid s) \propto \exp(\alpha \cdot W_{DM}(t) * U_{DM}(t; s)) \quad (15)$$

$$U_{DM}(t; s) = (1-\delta) \cdot \log L_0(s \mid t) + \delta \cdot \log P_{S2}(s \mid t) - D_{DM}(t) \quad (16)$$

Setting the Speed Weight

As described, emotions such as fear affect the patient's actions under time constraints. Considering that these emotions are highly dependent on the severity of the patient's situation, we define a model to select a prior for the patient's stress level.

First, the advisor explains the surgery results to the patient but prefers to avoid lengthy descriptions, especially if they believe the patient has sufficient medical knowledge. Therefore, the advisor decides on the amount of detail (utterance's detailedness – UD) based on the patient's medical understanding (PMU). If the patient demonstrates a high level of medical understanding, the advisor provides a concise summary. Conversely, if the patient has a lower level of medical understanding, the advisor offers a more detailed explanation to ensure comprehension. This model follows the basic RSA format:

$$P_{L0}(UD \mid PMU) \propto [\![UD]\!](PMU) \cdot P(PMU) \quad (17)$$

$$P_{S1}(UD \mid PMU) \propto \exp(\alpha \cdot U_{S1}(UD; PMU)) \quad (18)$$

$$U_{S1}(t; s) = \log L_0(s \mid t) - D_{S1}(t) \quad (19)$$

After deciding on UD, the advisor then decides on the utterance itself, depending on the patient's state. The patient then appraises the situation and would get affected depending on their inference on the doctor's description.

$$P_{\mathrm{Patient}}(\mathrm{SR} \mid \mathrm{PMU}, \mathrm{DU1}, \mathrm{UD}) \propto P(\mathrm{SR}) \cdot P_{\mathrm{DS1}}(\mathrm{DU1}, \mathrm{UD} \mid \mathrm{PMU}, \mathrm{SR}) \quad (20)$$

The patient's speed weight then follows a truncated beta distribution (bounded between 0 and 1), decided based on the patient's inference of the state.

The pseudo code of the algorithm is available in figure 1.

IV. Experiments

For a better demonstration of the model, we work on a medical SDM scenario for a rectal cancer patient about the post-surgery treatment options, a decision complicated by emotional weight and potential life-changing outcomes. Colorectal cancer (CRC), prevalent in Western societies and increasingly in Asia, commonly forms in the rectum's tissues. The primary treatment for localized rectal cancer is surgical

```
# Initializations
initialize   relevance_func_doc,   relevance_func_fam,   relevance_func_pat,   commitment_func,   speed_func,
worldStates, learning_rate, true_state
initialize   initial_CG_doc, initial_CG_fam, initial_CG_pat = uniform(worldStates)
define models   patient_cure, doctor_speaker, family_listener, family_speaker, patient_listener, decision_model

# Agent 1: Doctor's Suggestion
doctorSpeaker_post = doctor_speaker(true_state, relevance_func_doc, initial_CG_doc)
treatment1 = sample(doctorSpeaker_post)  # doctor suggests
doctor_mind_posterior = post(patient_cure(treatment1))
updated_CG_pat = (1- learning_rate) * initial_CG_pat + learning_rate * doctor_mind_posterior
updated_CG_fam = (1- learning_rate) * initial_CG_fam + learning_rate * doctor_mind_posterior

# Agent 2: Family Member's Suggestion
familySpeaker_post = family_speaker(true_state, relevance_func_fam, initial_CG_fam, updated_CG_fam)
treatment2 = sample(familySpeaker)  # family suggests
family_mind_posterior = post(family_listener(treatment2))
updated_CG_pat = (1- learning_rate) * updated_CG_pat + learning_rate * family_mind_posterior

# Agent 3: Patient's Decision Model
patientDM   =   decision_model(updated_CG_patient,   commitment_func,   speed_func,   relevance_func_pat,
updated_CG_pat)
print(patientDM) # The patient's decision posterior
final_decision = sample(patientDM)
```

Figure 1. The complete decision model.

resection, removing the cancerous part of the rectum along with a margin of healthy tissue. Post surgery, multiple treatment options are available depending on the surgery's outcome. We investigated various aspects of three-agent SDM following the surgery to demonstrate the model's applicability. Although I was not able to access a complete data because of the patient's privacy issues, all prior distributions and parameters were taken as close to the literature as possible (Shaukat et al., 2021; El-Shami et al., 2015; Edwards et al., 2012; Kosmider & Lipton, 2007; Desch et al., 2005; Compton et al., 2000).

V. Results

The result posterior probabilities of our agents are available in figures 2-4. For the results of the doctor and patient's model, we can observe a consistent output with the information in the literature (Shaukat et al., 2021; El-Shami et al., 2015; Edwards et al., 2012; Kosmider & Lipton, 2007; Desch et al., 2005; Compton et al., 2000). i.e., for the doctor, when the state is complete resection with clean margins, the most likely suggested treatment is to only monitor the patient. But the doctor also takes some amount of probability for the rest of the options because taking an extra treatment like chemotherapy ensures that any possible issues are solved, and the patient is guaranteed to avoid any relapse later on. Similarly, for when resection leaves positive margins, the most likely options, as described in the past literature, often involves either a second surgery to achieve clear margin, or radiation therapy may be employed to target residual microscopic disease. In addition, the standard care for node-positive colorectal cancer, i.e., when lymph node is involved, is using adjuvant chemotherapy to target the micro metastatic disease to reduce recurrence risk. While Radiation Therapy and Targeted Therapies are also possible options for local control, and depending on whether biomarker-driven options are available. Lastly, for when metastasis is detected, decisions have some degree of uncertainty and depend on the extent of metastasis, resectability, and patient condition.

For the emotion aspect, we have 3 main observations: For the relevance weight, as the relevance weight increases, the weight of side effects also increases for all 3 agents. Each

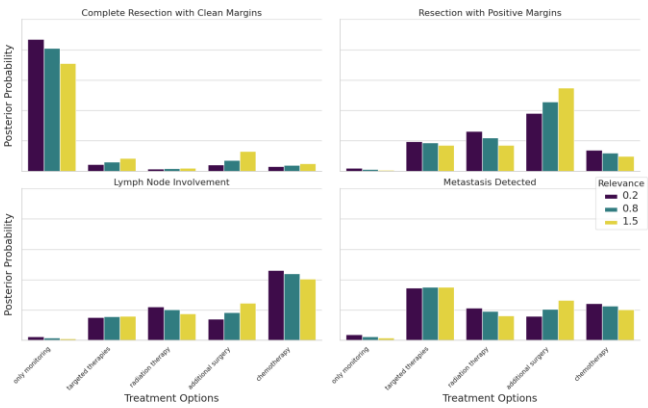


Figure 2. The resulting posterior on different treatment options for the doctor, demonstrated for different relevance weights and different states.

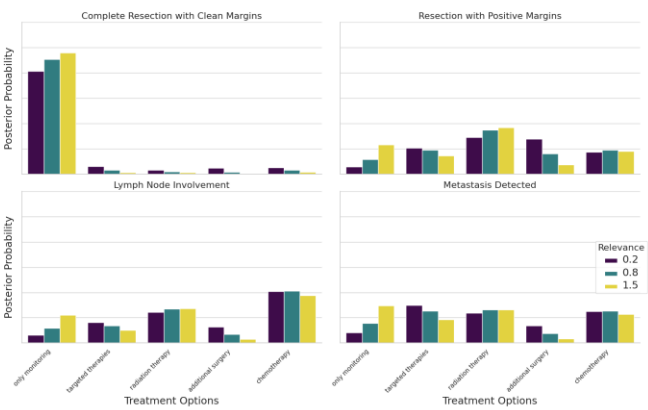


Figure 3. The resulting posterior on different treatment options for the family member, demonstrated for different relevance weights and different states.

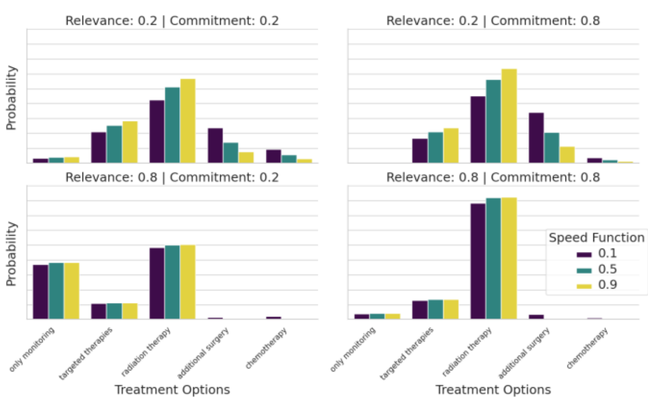


Figure 4. The resulting posterior on different treatment decision for the patient, demonstrated for different relevance, speed, and commitment weights and state 'Resection with Positive Margins'.

treatment has its own associated side effects, some of which are less extreme than others. For example, common associated issues of chemotherapy are nausea, hair loss, fatigue, and infection, while radiation therapy is associated with bladder issues, skin problems, and fatigue. It could be seen that each agent suitably incorporates these effects.

In addition, we observe that as the effects of the commitment function increase, the patient's selections get increasingly similar to the family member's possible suggestions, rather than the doctor's, which is as we expected from the commitment function. The speed function's effect also correctly incorporates the possible fears of the patient from specific side effects (e.g., hair loss in this case). Depending on the speed function's weight, the patient's actions might deviate from the patient's mind model of the doctor and the family member, and their suggestions.

VI. Conclusion

In this paper, we introduced a model of medical shared decision making based on the rational speech acts framework of communication. We demonstrated how the four aspects of emotional functions in Pfister and Böhm's four factor theory of emotion (2008) could be integrated in the available framework, and how these modifications suitably demonstrate the assumed effects of information, speed, commitment, and relevance functions from each agent.

We explored every emotional function is mapped differently onto the decision-making process, and the posterior probability. We confirmed that our finding were in line with multiple other past work on the theories of emotion, and Colorectal cancer. Unfortunately, due to the lack of individual experimental data, we were not able to fully assess the goodness of fit of our simulations. Nevertheless, our study proposes a Bayesian model for various emotional functions of Pfister and Böhm's (2008) proposed framework, and our experiments shows some encouraging baseline results. We hope that, by further experiments, future studies could further validate our models and improve the results. In addition, although we adapted the traditional RSA model of pragmatic reasoning for a scenario more following the theory of mind and emotions rather than pragmatics, it should be noted that as our method fully supports the integration of pragmatic aspects. In SDM alone, past literature has shown many interesting language games between doctors and patients. For example, for antibiotics prescription, patients frequently use various language strategies, such as using intensifiers, discourse markers, and metaphor, in order to lead the doctor to prescribe antibiotics for them. The doctors, on the other hand, are under pressure to prescribe antibiotics, and they need to align their medical observations, with their understanding of the patient, to decide on the best treatment (Stivers, 2007, 2021; Heritage et al., 2010). We hope that our model creates an optimal framework to model such language strategies under pressure and stressful scenarios.

References

Barrett, L. F. (2006a). Are emotions natural kinds? Perspectives on Psychological Science, 1, 28-58.

Barrett, L. F. (2006b). Valence as a basic building block of emotional life. Journal of Research in Personality, 40, 35-55.

Bechara, A., & Damasio, A. R. (2005). The somatic marker hypothesis: A neural theory of economic decision. Games and Economic Behavior, 52, 336-372

Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (1997). Deciding advantageously before knowing the advantageous strategy. Science, 275, 1293-1295.

Boyd, R. (1999). Homeostasis, species, and higher taxa. In R. A. Wilson (Ed.), Species: new interdisciplinary essays (pp. 141-185). Cambridge, MA: MIT Press.

Breslin J. Autonomy and the role of the family in making decisions at the end of life. J Clin Ethics 2005; 16: 11–19.

Charland, L. C. (2002). The natural kind status of emotion. British Journal for the Philosophy of Science, 53, 511-537.

Chater, N., & Manning, C. D. (2006). Probabilistic models of language processing and acquisition. Trends in cognitive sciences, 10(7), 335–344.

Compton, C. C., Fielding, L. P., Burgart, L. J., Conley, B., Cooper, H. S., Hamilton, S. R., ... & Willett, C. (2000). Prognostic factors in colorectal cancer: College of American Pathologists consensus statement 1999. Archives of pathology & laboratory medicine, 124(7), 979-994.

Damasio, A. R. (1994). Descartes' error: Emotion, rea son and the human brain. New York: Putnam.

Degen, Judith (2023). "The rational speech act framework". In: Annual Review of Linguistics 9, pp. 519–540.

Desch, C. E., Benson III, A. B., Somerfield, M. R., Flynn, P. J., Krause, C., Loprinzi, C. L., ... & Petrelli, N. J. (2005). Colorectal cancer surveillance: 2005 update of an American Society of Clinical Oncology practice guideline. Journal of Clinical Oncology, 23(33), 8512-8519.

Edwards, M. S., Chadda, S. D., Zhao, Z., Barber, B. L., & Sykes, D. P. (2012). A systematic review of treatment guidelines for metastatic colorectal cancer. *Colorectal Disease*, *14*(2), e31-e47.

Elwyn, G. (2021). Shared decision making: what is the work?. Patient Education and Counseling, 104(7), 1591-1595.

Elwyn, G., Frosch, D. L., & Kobrin, S. (2015). Implementing shared decision-making: consider all the consequences. Implementation science, 11, 1-10.

Elwyn, G., Frosch, D., Thomson, R., Joseph-Williams, N., Lloyd, A., Kinnersley, P., ... & Barry, M. (2012). Shared decision making: a model for clinical practice. Journal of general internal medicine, 27, 1361-1367.

Frank, Michael C. and Noah D. Goodman (2012). "Predicting Pragmatic Reasoning in Language Games". In: Science 336.6084, pp. 998–998.

Fudenberg, D., Newey, W., Strack, P., & Strzalecki, T. (2020). Testing the drift-diffusion model. Proceedings of the National Academy of Sciences, 117(52), 33141–33148.

Griffiths, P. E. (1997). What emotions really are: The problem of psychological categories. Chicago: University of Chicago Press.

Griffiths, P. E. (2004). Emotions as natural and normative kinds. Philosophy of Science, 71, 901-911.

Griffiths, T. L., Kemp, C., & Tenenbaum, J. B. (2008). Bayesian models of cognition. In R. Sun (Ed.), The Cambridge handbook of computational psychology (p. 59–100). Cambridge University Press.

Griffiths, T., & Tenenbaum, J. B. (2007). Two proposals for causal grammars. Causal learning: Psychology, philosophy, and computation, 323–345.

Heritage, J., Elliott, M. N., Stivers, T., Richardson, A., & Mangione-Smith, R. (2010). Reducing inappropriate antibiotics prescribing: the role of online commentary on physical examination findings. Patient education and counseling, 81(1), 119-125.

Ho, A. (2008). Relational autonomy or undue pressure? Family's role in medical decision-making. Scandinavian journal of caring sciences, 22(1), 128-135.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. Econometrica 47.2: 363-391.

Kersten, D., Mamassian, P., & Yuille, A. (2004). Object perception as Bayesian inference. Annu. Rev. Psychol., 55, 271–304.

Kosmider, S., & Lipton, L. (2007). Adjuvant therapies for colorectal cancer. World Journal of Gastroenterology: WJG, 13(28), 3799.

Lazarus, R. S. (1966). Psychological stress and the coping process.

Ong, D. C., Zaki, J., & Goodman, N. D. (2019). Computational models of emotion inference in theory of mind: A review and roadmap. Topics in cognitive science, 11(2), 338–357.

Pfister, H. R., & Böhm, G. (2008). The multiplicity of emotions: A framework of emotional functions in decision making. Judgment and decision making, 3(1), 5-17.

Russell, J. A. (2003). Core affect and the psychological construction of emotion. Psychological Review, 110, 145-172

Saxe, R., & Houlihan, S. D. (2017). Formalizing emotion concepts within a Bayesian model of theory of mind. Current opinion in Psychology, 17, 15–21.

Simon, H. A. (1955). A behavioral model of rational choice. The quarterly journal of economics, 99-118.

Simon, H. A. (1990). A mechanism for social selection and successful altruism. Science, 250(4988), 1665-1668.

Simon, H. A. (1991). Bounded rationality and organizational learning. Organization science, 2(1), 125-134.

Solomon, R. C., & Stone, L. D. (2002). On "positive" and "negative" emotions. Journal for the Theory of Social Behavior, 32, 417-435.

Stivers, T. (2007). Prescribing under pressure: Parent-physician conversations and antibiotics. Oxford University Press.

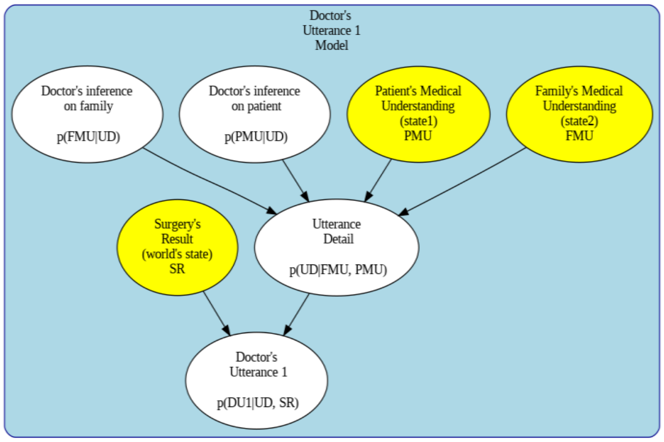
Stivers, T. (2021). Managing patient pressure to prescribe antibiotics in the clinic. Pediatric Drugs, 23(5), 437-443.

Von Neumann, J., & Morgenstern, O. (1947). Theory of games and economic behavior, 2nd rev.

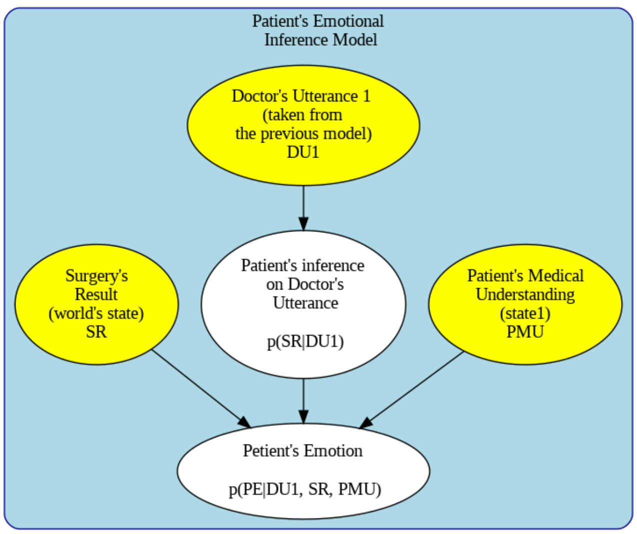
Ying, L., Michal, A., & Zhang, J. (2022). A Bayesian Drift-Diffusion Model of Schachter-Singer's Two Factor Theory of Emotion. Proceedings of the Annual Meeting of the Cognitive Science Society, 44(44).

VII. Appendix: Directed acyclic graphs

Directed acyclic graph for the doctor's detailedness:



Directed acyclic graph for the patient's emotion:



Directed acyclic graph for the patient's decision: