NIPS 2013
Neural Information Processing Systems
December 5 - 10, Lake Tahoe, Nevada, USA

Reviews For Paper

Paper ID	1352
Title	Mental Rotation as Bayesian Quadrature

Masked Reviewer ID: Assigned_Reviewer_3

Review:

Question	
Comments to author(s). First provide a summary of the paper, and then address the following criteria: Quality, clarity, originality and significance. (For detailed reviewing guidelines, see http://nips.cc/PaperInformation/ReviewerInstructions)	This paper considers a task where one needs to judge whether two depicted objects are distinct or identical subject to a rotation. They analyzed several models of mental rotation and found that a model based on an optimal experimental design using Bayesian quadrature is more accurate than the others. Quality: Major question: they show that Bayesian Quadrature (BQ) outperforms other models in accuracy, and I see this is well supported by their analysis and experiment results. Nonetheless, does the current paper answer the question it poses in the beginning: how do people use mental simulation? Perhaps I'm missing something, but is there comparison to human data here? It seems the paper only compares the three models based on their simulation results. If I am correct on this matter, the paper only shows that BQ is a more accurate model than the others, but not that BQ is similar to what people might potentially use. Other than this, the paper is technically sound. Clarity: The paper is pretty clear, except for the confusion brought about by the aforementioned problem. A minor comment is regarding the description of the task. For example, I'm not clear how the shapes are generated, so that on Page 4 Section 4, I'm not sure about prior of the stimuli.

	Originality: The paper is original. It incorporates a state-of-art method for evaluating integration, BQ, in models of mental rotation. This is a novel combination of existing techniques. Significance: This paper may be of interest to a general NIPS audience. Nonetheless I find it might lack impact in its current form, because it seems not able to address whether the model captures human data, rather, the message seems to be that BQ is a better computational method in order to achieve better model accuracy.
Please summarize your review in 1-2 sentences	Sound paper; but no comparison to human data as it claims?
Quality Score - Does the paper deserves to be published?	5: Marginally below the acceptance threshold
Impact Score - Independently of the Quality Score above, this is your opportunity to identify papers that are very different, original, or otherwise potentially impactful for the NIPS community.	1: This work is incremental and unlikely to have much impact even though it may be technically correct and well executed.
Confidence	3: Reviewer is fairly confident

Masked Reviewer ID: Assigned_Reviewer_4

Review:

Question	
	SUMMARY: In this paper, the authors propose a rational account for how humans perform mental rotation tasks (in this case determining if one shape is a rotated version of a second shape). At a computational level of analysis, it is proposed that people go about this task by comparing the probability of the null hypothesis (h_0) that the two objects are not the same shape versus the probability of the alternative hypothesis (h_1) that they are the same shape (i.e., declare them the same if p(h_1) > p(h_0)). The challenge is integrating over all rotations to compute p(h_1). The authors propose an algorithmic solution to this problem that involves continually making small rotations and

Comments to author(s). First provide a summary of the paper, and then address the following criteria: Quality, clarity, originality and significance. (For detailed reviewing guidelines, see http://nips.cc/PaperInformation/ReviewerInstructions)	evaluating a similarity function that is defined to approximate the probability of the actual shape given the "mental" rotation. To compute the full likelihood ratio, the similarity function (S) needs to be approximated across all rotations. The authors present three methods for approximating S. The first approach (Naïve) performs simple hill-climbing search until a local maximum is found and then estimates the function via linear interpolation across the sampled rotations. The second approach (parametric) assumes a circular Gaussian shape of S and estimates the parameters of that distribution using the same hill-climbing samples from the previous approach. The third approach (nonparametric) uses Bayesian Quadrature to determine the shape of S. The authors demonstrate the superiority of the nonparametric approach and conclude that this provides a rational explanation of mental rotation. COMMENTS TO AUTHORS: This paper is well written and presents a nice model for mental rotation. However, it falls short of providing a convincing rational explanation of mental rotation. At the computational level, I don't see the need for integrating across the whole set of rotations. It seems more logical that humans just rotate until they find a match that is easy to recognize or exhaust the set of potential rotations. Perhaps the integration approach is more appropriate in difficult cases where a match can't be determined for sure or in cases where there is noise in the shapes. At the algorithmic level, the models that Bayesian Quadrature is compared to are very simplistic and obviously not going to perform as well. Both models take a hill-climbing approach that will get stuck in the first local maximum. This hypothesis is akin to saying that humans perform mental rotations of the object until it is somewhat more similar to the original object and then quit even

	if it's obviously not a match at that rotation. It seems that there could be other simpler models that match the human data as well. I'm a bit skeptical of the 95% confidence level used to select the hypothesis for the Bayesian Quadrature model (line 345). It seems that adjusting this would directly affect the mean error (ME) measure used to assess the model. Perhaps the 4% model error is close to the human 3.2% error only because the threshold was 95%. I don't see how the analysis described around line 372 matches up with the experimental results in [6]. They just happen to be two patterns that are linear for different reasons. If RT happened to be non-linear, you'd still expect the plot of mean mental rotation to true rotation to be linear because the goal is to find the rotation that produces a match. It seems that the number of steps to find a match would better compare to RT. Is there evidence that the psychological results are the same for the type of objects considered in this paper? It seems that this task is easier than the pseudo 3D version in Figure 1. Probably the linear dependence of RT on angle would still hold. Minor comments: Line 18: missing "be" in "should BE used in …" Line 382: if MSE is adjusted so that 1 is maximum error, why is it greater than 1 here?
Please summarize your review in 1-2 sentences	Interesting paper with a well developed, nice mathematical model. The evidence for the model being the correct account of mental rotation, however, is a bit lacking.
Quality Score - Does the paper deserves to be published?	5: Marginally below the acceptance threshold
Impact Score - Independently of the Quality Score above, this is your opportunity to identify papers that are very different, original, or otherwise potentially	1: This work is incremental and unlikely to have much impact even though it may be technically correct and well

impactful for the NIPS community.	executed.
Confidence	3: Reviewer is fairly confident

Masked Reviewer ID: Assigned_Reviewer_5

Review:

Question	
Comments to author(s). First provide a summary of the paper, and then address the following criteria: Quality, clarity, originality and significance. (For detailed reviewing guidelines, see http://nips.cc/PaperInformation/ReviewerInstructions)	The paper studies a task in which one needs to decide if two images depict the same object, albeit at different orientations, or if the images depict different objects. This task has been studied extensively in the psychology literature because of the insights it provides about "mental rotation". The paper focuses on the problem of deciding how much the object in one image should be rotated to try to align it with the object in the other image. The paper has a lot of strengths. I like that it considers the task both from a "rational viewpoint" (what is the optimal solution to the problem), as well as the "algorithmic viewpoint" (how could an approximation to the optimal solution be efficiently computed). I like how the paper formalizes the problem as a decision task. I like that the paper examines a number of different algorithms, each of which is interesting. And I like the simulation results demonstrating that a Bayesian quadrature algorithm seems to work best. To some extent, the paper feels premature. My hope is that the current work is a great foundation for a future, longer article that also includes data from experiments with human subjects. In the current manuscript (and in the hoped for future manuscript), the authors need to make a claim about people's mental representations and operations. Is the claim that people perform Bayesian quadrature in their heads? Are these computations "psychologically plausible"? If the authors' answer is yes, why should the reader believe this?
Please summarize your review in 1-2 sentences	In summary, the manuscript studies an unusual but interesting task, and does so in appealing way.
Quality Score - Does the paper deserves to be	8: Top 50% of accepted NIPS papers

published?	
Impact Score - Independently of the Quality Score above, this is your opportunity to identify papers that are very different, original, or otherwise potentially impactful for the NIPS community.	2: This work is different enough from typical submissions to potentially have a major impact on a subset of the NIPS community.
Confidence	3: Reviewer is fairly confident