REPORT DOCUMENTATION PAGE			*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
05-03-2015	Final	07-03-2012 – 06-03-2015

4. TITLE AND SUBTITLE Combining Offline and Online Computation for Solving Partially Observable Markov Decision Process	**5a. CONTRACT NUMBER** FA2386-12-1-4031 **5b. GRANT NUMBER** Grant AOARD-124031 **5c. PROGRAM ELEMENT NUMBER** 61102F
6. AUTHOR(S) Wee Sun Lee	**5d. PROJECT NUMBER** **5e. TASK NUMBER** **5f. WORK UNIT NUMBER**
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Department of Computer Science, National University of Singapore Computing 1, 13 Computing Drive, Singapore 117417, Republic of Singapore	**8. PERFORMING ORGANIZATION REPORT NUMBER** N/A
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) AOARD UNIT 45002 APO AP 96338-5002	**10. SPONSOR/MONITOR'S ACRONYM(S)** AFRL/AFOSR/IOA(AOARD) **11. SPONSOR/MONITOR'S REPORT NUMBER(S)** AOARD-124031

12. DISTRIBUTION/AVAILABILITY STATEMENT

Distribution Code A: Approved for public release, distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Partially observable Markov decision process (POMDP) provides a general and mathematically elegant way of formulating planning and control problems under uncertainty. Unfortunately, POMDPs are computationally intractable to solve in the worst case, prompting the development of approximation algorithms. In this project, they explored the use of online algorithms for approximately solving large-scale POMDPs. They developed a new online POMDP solver, DESPOT, with good theoretical and practical properties. The DESPOT algorithm was used as part of our entry that finished in first place at the ICAPS 2014 International Probabilistic Planning Competition (IPPC) POMDP track. They also applied the DESPOT algorithm on the problem of autonomous vehicle navigation through crowded locations, demonstrating its use in a real application. Although POMDPs are intractable in the worst case, there are subclasses of POMDPs that can be tractably approximated and are at the same time practically interesting. They applied online methods to two such special cases of POMDPs, specifically adaptive informative path planning and active learning, obtaining practical polynomial-time algorithms with guaranteed approximation bounds.

15. SUBJECT TERMS

Partial observable markov decision process, Online decision making, Belief tree, Sampling, Active learning, Path finding, Autonomous vehicle navigation

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON** Hiroshi Motoda, Ph. D.
U	U	U	SAR	9	**19b. TELEPHONE NUMBER** *(Include area code)* +81-42-511-2011

Report Documentation Page	*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
25 MAR 2015	**Final**	**07-03-2012 to 06-03-2015**

4. TITLE AND SUBTITLE **Combining Offline and Online Computation for Solving Partially Observable Markov Decision Process**	5a. CONTRACT NUMBER **FA2386-12-1-4031**
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER **61102F**
6. AUTHOR(S) **Wee Sun Lee**	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Department of Computer Science,,National University of Singapore,Computing 1, 13 Computing Drive,,Singapore 117417, Republic of Singapore,NA,NA**	8. PERFORMING ORGANIZATION REPORT NUMBER **N/A**
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) **AOARD, UNIT 45002, APO, AP, 96338-5002**	10. SPONSOR/MONITOR'S ACRONYM(S) **AFRL/AFOSR/IOA(AOARD)**
	11. SPONSOR/MONITOR'S REPORT NUMBER(S) **AOARD-124031**

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Partially observable Markov decision process (POMDP) provides a general and mathematically elegant way of formulating planning and control problems under uncertainty. Unfortunately, POMDPs are computationally intractable to solve in the worst case, prompting the development of approximation algorithms. In this project, they explored the use of online algorithms for approximately solving large-scale POMDPs. They developed a new online POMDP solver, DESPOT, with good theoretical and practical properties. The DESPOT algorithm was used as part of our entry that finished in first place at the ICAPS 2014 International Probabilistic Planning Competition (IPPC) POMDP track. They also applied the DESPOT algorithm on the problem of autonomous vehicle navigation through crowded locations, demonstrating its use in a real application. Although POMDPs are intractable in the worst case there are subclasses of POMDPs that can be tractably approximated and are at the same time practically interesting. They applied online methods to two such special cases of POMDPs, specifically adaptive informative path planning and active learning, obtaining practical polynomial-time algorithms with guaranteed approximation bounds.

15. SUBJECT TERMS

Partial observable markov decision process, Online decision making, Belief tree, Sampling, Active learning, Path finding, Autonomous vehicle navigation

<table>
<tr><td colspan="3">16. SECURITY CLASSIFICATION OF:</td><td rowspan="2">17. LIMITATION OF ABSTRACT
Same as Report (SAR)</td><td rowspan="2">18. NUMBER OF PAGES
9</td><td rowspan="2">19a. NAME OF RESPONSIBLE PERSON</td></tr>
<tr><td>a. REPORT
unclassified</td><td>b. ABSTRACT
unclassified</td><td>c. THIS PAGE
unclassified</td></tr>
</table>

Final Report for AOARD Grant 124031

Combining Offline and Online Computation for Solving Partially Observable Markov Decision Process

Date: 6 March 2015

Name of Principal Investigators (PI and Co-PIs): Wee Sun LEE

- E-mail address : leews@comp.nus.edu.sg
- Institution : National University of Singapore
- Mailing Address : Department of Computer Science, Computing 1, 13 Computing Drive, Singapore 117417, Republic of Singapore
- Phone : +65 65164526
- Fax : +65 67794580

Period of Performance: 03/07/2012 – 03/06/2015

Abstract: Partially observable Markov decision process (POMDP) provides a general and mathematically elegant way of formulating planning and control problems under uncertainty. Unfortunately, POMDPs are computationally intractable to solve in the worst case, prompting the development of approximation algorithms. In this project, we explore the use online algorithms for approximately solving large-scale POMDPs. We developed a new online POMDP solver, DESPOT, with good theoretical and practical properties. The DESPOT algorithm was used as part of our entry that finished in first place at the ICAPS 2014 International Probabilistic Planning Competition (IPPC) POMDP track. We also applied the DESPOT algorithm on the problem of autonomous vehicle navigation through crowded locations, demonstrating its use in a real application. Although POMDPs are intractable in the worst case, there are subclasses of POMDPs that can be tractably approximated and are at the same time practically interesting. We applied online methods to two such special cases of POMDPs, specifically adaptive informative path planning and active learning, obtaining practical polynomial-time algorithms with guaranteed approximation bounds.

Introduction: Partially observable Markov Decision Processes (POMDP) have been shown to provide useful models for dialog systems [20], assistive technologies [9], autonomous vehicle navigation [2] and other practical problems of control and planning under uncertainty. POMDP allows the specification of actions with probabilistic outcomes, probabilistic observation functions, and reward functions that depends on the state and action; it provides an elegant and general method for formulating planning and control problems under uncertainty. However, the generality of the method comes at a cost. POMDPs are intractable to solve in the worst case, prompting the need for approximation techniques.

Fifteen years ago, only small POMDP problems of the order of tens of states can be solved. Approximation methods, particularly those based on point-based methods [10, 15, 17] pushed the size of solvable problems to thousands of states in the last ten years. Further work, by our group and others pushed the size of solvable problems much larger, allow some practical problems to be feasibly tackled using POMDPs [2, 9, 20]. Naturally, not all large problems can be efficiently approximated, and our aim in this project is to develop tools that allow a wider range of POMDPs to be solved.

We take two approaches in developing effective approximation methods for POMDPs in this project:

- We developed *online sampling-based anytime* algorithms for approximating POMDP problems. Most previous state-of-the-art POMDP algorithms are offline algorithms. Offline algorithms have the drawback of needing to pre-compute all possible situations in advance before the plan is actually put into practice. For more difficult problems, pre-computing all possible situations in advance is not possible. In online algorithms, only the space reachable from the current state of the world needs to be considered by the algorithm. Often, considering a relatively small space in the neighbourhood of the current state is sufficient to decide on an adequate action. Online algorithms work well for these problems, expanding the set of POMDP problems that can be effectively approximated in practice.

- Not all problems that can be modeled using POMDPs are equally difficult to solve. We examined subclasses of POMDPs that are of practical interest and developed algorithms specifically for those subclasses. In particular, we examined Bayesian reinforcement learning, adaptive informative path planning and active learning. For adaptive informative path planning and active learning, we obtained polynomial time algorithms with guaranteed approximation by exploiting properties of those problems. It turned out that these algorithms are online algorithms as well.

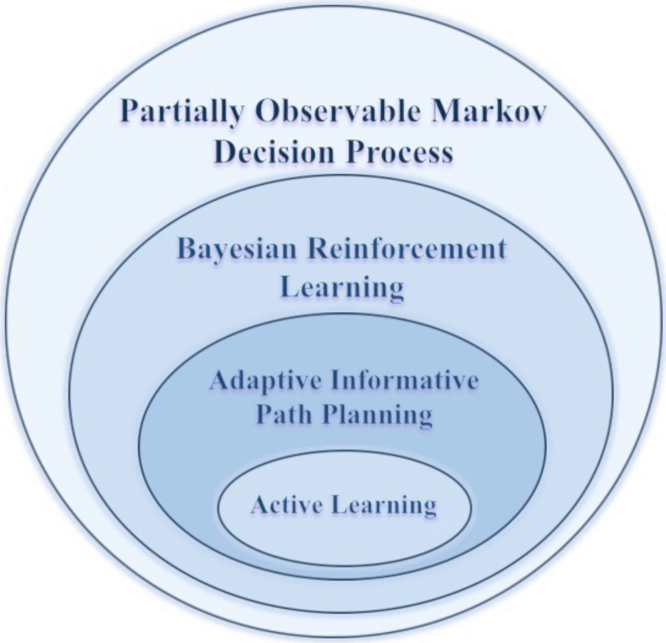

Figure 1: Subclasses of POMDP examined in this project

Figure 1 shows the subclasses of POMDP examined in this project. We now describe the subclasses and results obtained for them in this project.

Our initial work focused on extending our offline POMDP solver to handle more general POMDP classes. In particular, we developed a sampling based POMDP solver to handle continuous space and continuous observation POMDPs [4]. These models are particularly useful for solving problems that involve interaction with the physical world, such as robotic problems, as such problems are often naturally modeled with continuous space and observation. We also developed an offline solver for a subclass of POMDP: Bayesian reinforcement learning [19]. Bayesian reinforcement learning is a special case of POMDP where the state consists of both the system state as well as the unknown system *parameters*. We developed a new method for solving Bayesian reinforcement learning by sampling the parameters from its prior distribution and solving the resulting sampled system as a simpler discrete POMDP.

Based on the sampling techniques used in developing the offline solvers, we developed a new highly scalable online POMDP solver called DESPOT (Determinized Sparse Partially Observable Tree) [18]. DESPOT is an *anytime* algorithm that will output the best action found so far whenever the computation is stopped. The solver uses only state simulations, which is far more efficient than manipulation of probability distributions, in its search. This allows the solver to run on almost any problem, even extremely large ones. We are also able to show theoretically that, DESPOT's performance depends on the size of the optimal policy. This property of the DESPOT allows it to do well on easier problems (smaller optimal policies), instead of assuming that all problems are equally difficult and practically intractable.

The DESPOT algorithm was used as part of our entry that finished in first place at the ICAPS 2014 International Probabilistic Planning Competition (IPPC) POMDP track. We also applied the DESPOT algorithm on the problem of autonomous vehicle navigation through crowded locations to demonstrate its ability to scale to very large problems [1]. The video of the autonomous vehicle driving through a crowd can be viewed at https://www.youtube.com/watch?v=y_9VMD_sQhw. Finally, we are in the process of releasing the DESPOT solver as open source software for the benefit of the scientific community.

The design of the DESPOT solver allows it to run for almost all problems. However, the performance of the solver on a particular problem depends on the difficulty of the problem and the appropriateness of the solver for the problem. By appropriately restricting the subset of POMDP problems considered, we are able to develop polynomial time algorithms with approximation guarantees. In particular, we developed such algorithms for adaptive informative path planning and Bayesian active learning.

In informative path planning, a robot needs to move to different locations in order to gather information to identify some unknown parameters. In *adaptive* informative path planning, the robot may re-plan its movement after every observation and the aim is to minimize the expected movement cost required to identify the parameters. This is a subclass of Bayesian reinforcement learning where the action of the robot is restricted to moving from location to location. For this problem, we are able to develop a polynomial time approximation algorithm with performance within a polylog factor of the performance of the optimal algorithm [12]. The algorithm is an online algorithm which uses a group Steiner tree approximation algorithm to compute a path assuming that the most likely observation is received at each step, and recomputes the path whenever a different observation is received. Experiments show that the algorithm is practical and works better than competing algorithms on problems that require both adaptivity and planning in order to do well.

We also studied Bayesian active learning, which is a special case of adaptive informative path planning when the cost of moving from any location to any other location is the same [6, 7]. In particular, we examined when online greedy algorithms are effective. We showed a commonly used active learning method that selects the least confident example for labeling achieves a constant factor approximation to the optimal algorithm in the worst case in terms of eliminating the version space, when only a fixed number of queries is allowed. We also show that another commonly used algorithm that selects the example with the highest entropy is not able to achieve a constant factor approximation for an appropriately defined entropy objective function. We proposed an approximation to the entropy objective, the Gibbs error, which can be approximated within a constant factor by a online greedy algorithm and generalized the objective function to use general loss functions.

In the following section, we describe selected experiments done as part of the project and the results obtained. For details of the algorithms, including theoretical properties and proofs, and other experiments, we refer the reader to the publications [1, 4, 5, 6, 7, 12, 13, 18, 19].

Experiments and Results:

Continuous state and observation POMDP [4]: In the experiment, we examine an intersection crossing problem. The autonomous vehicle **B**, in blue, stops at the intersection and waits for the other vehicle **G**, in green to clear before proceeding (Figure 2). **B** wants to go through the intersection as fast as possible, while maintaining safety. **B** does not know **G**'s position initially but is equipped with laser that gives noisy readings.

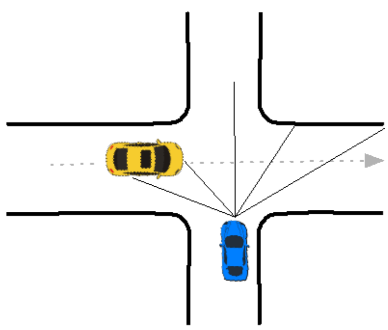

Figure 2: Intersection crossing.

We compare our new POMDP solver with continuous state and continuous observation with the continuous state discrete observation POMDP solver from [3]. The discrete observations are created by discretizing the possible locations of **G** and selecting the most likely location based on the current observation. The results can be found in Table 1, where the continuous observation model outperforms the discrete one in terms of both the crossing time and the accident rate.

Table 1: Results of intersection crossing for POMDP with continuous observations and discrete observations.

Observation Model	Crossing Time	Accident Rate
Laser beam (continuous)	2.61 ± 0.0095	0.0029 ± 0.00053
Most likely position (discrete)	4.27 ± 0.0012	0.0093 ± 0.00010

Bayesian reinforcement learning [19]: The experiment in this work is motivated by an accident in the 2007 DARPA Urban Challenge [11]. In that event, two autonomous vehicles, **R** and **A**, approached an uncontrolled traffic intersection as shown in Figure 3. **R** had the right-of-way and proceeded. However, possibly due to sensor failure or imperfect driving strategy, **A** did not yield to **R** and almost caused a collision. This situation is quite common and occurs frequently even with human drivers. Crossing the intersection safely and efficiently without knowing the driving strategy of **A** poses a significant challenge.

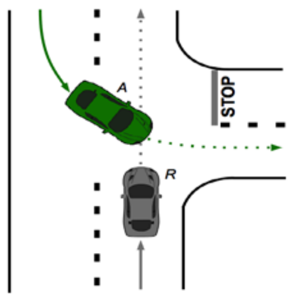

Figure 3: Model of near collision at DARPA Urban Challenge.

The driving strategy of **A** is unknown to the agent. We parameterize the driving strategy using 4 parameters [8]: (1) driver imperfection, (2) driver reaction time, (3) acceleration, and (4) deceleration. This parameterization can model a variety of drivers such as a reckless driver who never slows down at the intersection and an impatient driver who performs a rolling stop near the intersection. Bayesian reinforcement learning is challenging because the agent needs to both learn the parameters of A and cross the intersection, all within a small time window.

We compare our algorithm, MC-BRL, to a hand-crafted intersection policy that is commonly used in the traffic modeling community [14]. Our algorithm samples K parameters and reduce the problem to a discrete POMDP problem with the sampled parameters. The results are shown in Figure 4. With K = 150 and above, MC-BRL significantly outperforms that policy. While the hand-crafted policy is not designed to handle noisy observations, we think that the performance gap between the hand-crafted policy and MC-BRL is more likely to be caused by insufficient adaptivity of the hand-crafted policy in learning the driving strategy of A.

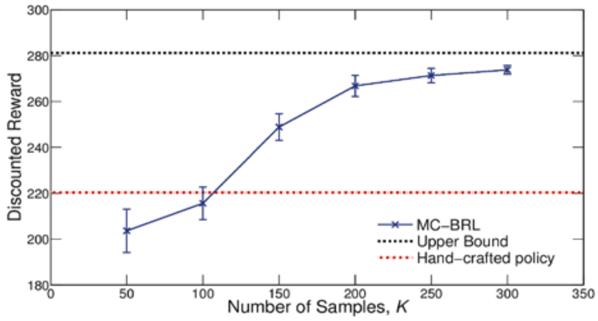

Figure 4: Average performance of MC-BRL at intersection navi gation, compared to an upper bound and a hand-crafted policy.

Online POMDP (DESPOT) [18]: In the experiment for this work, we created a target finding problem, LaserTag. In LaserTag, the agent's goal is to find and tag a target that intentionally moves away. Both the agent and target operate in a grid. The agent knows its own position and is equipped with a noisy laser to observe the target position. The agent can either stay in the same position or move to the four adjacent positions, paying a cost for each move. It can also perform the tag action and is rewarded if it successfully tags the target, but is penalized if it fails.

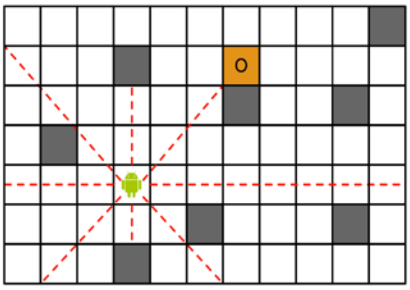

Figure 5: LaserTag. The robot needs t o tag the target with the help of laser observation.

We compare our new algorithm, DESPOT, to a state-of-the-art online POMDP algorithm, POMCP [16]. For the LaserTag problem, the average total discounted reward is -9.34 ± 0.26 for DESPOT and -19.58 ±

0.06 for POMCP.

For the ICAPS 2014 International Probabilistic Planning Competition (IPPC) POMDP track, we experimented with both DESPOT and POMCP using the problems from the 2011 competition. We found that DESPOT performs better on some problems and POMCP on others. For the 2014 competition, we ran both algorithms for a few rounds for each problem in order to select the better algorithm to run for the rest of rounds. In the 2014 competition, 8 domains were used: Traffic Control, Elevator Control, Crossing Traffic, Skill Teaching, Wildfire, Academic Advising, Tamarisk, and Triangle Tireworld. Our entry won the 2014 competition.

Figure 6; Autonomous golf cart driving through a crowd on campus.

We build a POMDP for autonomous driving among many pedestrians and implemented it physically on an autonomous golf cart [1]. The video of the golf cart driving through a crowd can be viewed at https://www.youtube.com/watch?v=y_9VMD_sQhw. We model the intention of a pedestrian as a subgoal location and assume that the pedestrian's behavior depends on the position of the subgoal location. The model captures uncertainty in pedestrian goal estimation as well as uncertainties in vehicle control and sensing. To achieve real-time performance with limited computational resources, we adopt a two-level hierarchical planning approach. This allows us to use the computationally expensive POMDP planning only in the critical part of the system that hedge against the uncertainty in predicting pedestrian behaviors. At the top level of the system, we apply the A* algorithm to search for a path through less crowded regions, based on a simplified predictive model of pedestrian motions. We then perform online POMDP planning in real time to control the speed of the vehicle along the planned path. We replan at both levels in each time step in order to handle dynamic changes in the environment. We tested the system extensively in a plaza on our university campus. Experiments show that the vehicle is capable of driving safely and smoothly in a crowded unstructured environment.

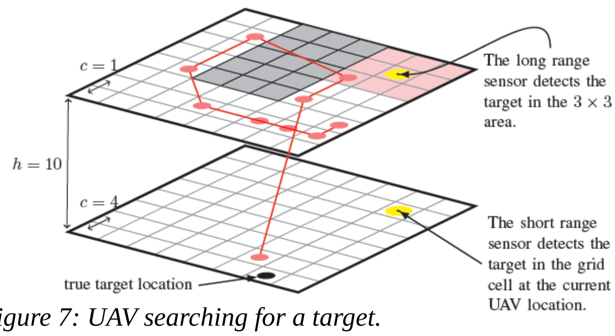

Figure 7: UAV searching for a target.

Adaptive informative path planning [12]: In this experiment, we simulated a UAV searching for a stationary target in an area modeled as a 9 by 9 grid (Figure 7). Initially the target lies in any of the cells with equal probabilities. The UAV can operate at two different altitudes. At the high altitude, it uses a long range sensor that determines whether the 3 by 3 grid around its current location contains the target. At the low altitude, the UAV uses a more accurate short-range sensor that determines whether the current grid cell contains the target. Moving between the different heights is expensive. Some grid cells are not visible from the high altitude because of occlusion, and the UAV must descend to the low altitude in order to search these cells.

We compare our algorithm, RAId with a replanning algorithm that reconstruct an open-loop plan at every time step using the latest information. Note that replanning gives the algorithm some adaptivity, but the adaptivity is limited due to the use of open-loop (non-adaptive) planning. For this problem, RAId has an average cost of 83.6 while the replanning algorithm has an average cost of 151.4.

Discussion: For a long time, POMDP has been considered mathematically elegant but impractical. We have been developing theoretical understanding as well as practical approximation algorithms to scale up POMDP solvers and make POMDP a practical technology. In this project, we have made various advances.

The online POMDP solver, DESPOT [18], developed in this project is highly scalable and can run on a wide range of problems, as demonstrated by our entry that finished first in the POMDP track of the ICAPS 2014 International Probabilistic Planning Competition (IPPC). In the problem of autonomous vehicle navigation through crowded locations, we further demonstrated that the DESPOT algorithm can be successfully implemented on a large problem of practical interest by appropriately exploiting domain knowledge [1]. DESPOT is being released as open source software for the benefit of the scientific community.

An offline POMDP solver for continuous state and observation was also developed in the project [4]. The paper reporting the result is an invited submission at the International Journal of Robotics Research (IJRR) Special Issue on RSS 2013 [5]. In addition, we developed new methods for special cases of POMDPs, in particular Bayesian reinforcement learning [19], adaptive informative path planning [12], and Bayesian active learning [6, 7]. Notably, in adaptive informative path planning and Bayesian active learning, we are able to obtain polynomial time algorithms with guaranteed approximation bounds. Furthermore, the algorithms developed are online algorithms. The paper reporting the result on adaptive informative path planning was invited for submission at the International Journal of Robotics Research (IJRR) Special Issue on WAFR 2014 [13]. One of the papers on active learning [6] won a best student paper award at UAI 2014, with the PI as faculty co-author.

List of Publications and Significant Collaborations that resulted from your AOARD supported project: In standard format showing authors, title, journal, issue, pages, and date, for each category list the following:

a) papers published in peer-reviewed journals,

- Haoyu Bai, David Hsu, and Wee Sun Lee, *Integrated Perception and Planning in the Continuous Space: A POMDP Approach,* Int. J. Robotics Research (Special issue on RSS 2013), 33(9):1288–1302, 2014.

b) papers published in peer-reviewed conference proceedings,

- Yi Wang and Kok Sung Won, David Hsu and Wee Sun Lee, *Monte Carlo Bayesian Reinforcement Learning*, International Conference on Machine Learning (ICML), 2012.
- Haoyu Bai, David Hsu, and Wee Sun Lee, *Integrated Perception and Planning in Continuous Space: A POMDP Approach,* Proc. Robotics: Science and Systems (RSS), 2013.
- Adhiraj Somani, Nan Ye, David Hsu, and Wee Sun Lee. *DESPOT: Online POMDP Planning with Regularization,* Neural Information Processing Systems (NIPS), 2013.
- Nguyen Viet Cuong, Wee Sun Lee, Nan Ye, Kian Ming A. Chai, and Hai Leong Chieu, *Active Learning for Probabilistic Hypotheses Using the Maximum Gibbs Error Criterion*, Neural Information Processing Systems (NIPS), 2013.
- Zhan Wei Lim, David Hsu, and Wee Sun Lee, *Adaptive Informative Path Planning in Metric Spaces.* Workshop on the Algorithmic Foundations of Robotics (WAFR), 2014.
- Nguyen Viet Cuong, Wee Sun Lee, and Nan Ye, *Near-optimal Adaptive Pool-based Active Learning with General Loss,* Uncertainty in Artificial Intelligence (UAI), 2014.

c) papers published in non-peer-reviewed journals and conference proceedings,

Nil

d) conference presentations without papers, Nil

e) manuscripts submitted but not yet published, and

- Haoyu Bai, Shaojun Cai, Nan Ye, David Hsu, and Wee Sun Lee, *Intention-Aware Online POMDP Planning for Autonomous Driving in a Crowd*, accepted for publication at International Conference on Robotics and Automation (ICRA), 2015.
- Zhan Wei Lim, David Hsu, and Wee Sun Lee, *Adaptive Informative Path Planning in Metric Spaces.* Submitted to Int. J. Robotics Research (Special issue on WAFR 2014), 2014.

f) provide a list any interactions with industry or with Air Force Research Laboratory scientists or significant collaborations that resulted from this work.

Nil

References:

1. Bai, H., Cai, S., Ye, N., Hsu, D., and Lee, W. S., *Intention-Aware Online POMDP Planning for Autonomous Driving in a Crowd*, accepted for publication at International Conference on Robotics and Automation (ICRA), 2015.
2. Bai, H., Hsu, D., Kochenderfer, M. J., and Lee, W. S., *Unmanned aircraft collision avoidance using continuous state POMDPs*, Robotics: Science and Systems (RSS), 2012.
3. Bai, H., Hsu, D., Lee, W. S., and Ngo, V., *Monte Carlo value iteration for continuous-state POMDPs*, Workshop on the Algorithmic Foundations of Robotics (WAFR), 2010.
4. Bai, H., Hsu, D., and Lee, W. S., *Integrated Perception and Planning in Continuous Space: A POMDP Approach*, Proc. Robotics: Science and Systems (RSS), 2013.
5. Bai, H., Hsu, D., and Lee, W. S., *Integrated Perception and Planning in the Continuous Space: A POMDP Approach*, Int. J. Robotics Research (Special issue on RSS 2013), 33(9):1288–1302, 2014.
6. Cuong, N. V., Lee, W. S., and Ye, N., *Near-optimal Adaptive Pool-based Active Learning with General Loss*, Uncertainty in Artificial Intelligence (UAI), 2014.
7. Cuong, N. V., Lee W. S. , Ye, N., Chai, K. M., and Chieu, H. L., *Active Learning for Probabilistic Hypotheses Using the Maximum Gibbs Error Criterion*, Neural Information Processing Systems (NIPS), 2013.
8. Gipps, P. G. *A behavioural car following model for computer simulation*. Transportation Research B, 15:105–111, 1981.
9. Hoey, J., Bertoldi, A., Poupart, P., and Mihailidis, A., *Assisting persons with dementia during handwashing using a partially observable Markov decision process*, International Conference on Vision Systems, (ICVS) 2007.
10. Kurniawati, H., Hsu, D., Lee, W. S., *SARSOP: Efficient Point-based POMDP Planning by Approximating Optimally Reachable Belief Spaces*, Robotics: Science and Systems (RSS) 2008.
11. Leonard, J., How, J., and Teller, S. *A perception driven autonomous urban vehicle*. Journal of Field Robotics, 25(10):727–774, 2008.
12. Lim, Z. W., Hsu, D., and Lee, W. S., *Adaptive Informative Path Planning in Metric Spaces.* Workshop on the Algorithmic Foundations of Robotics (WAFR), 2014.
13. Lim, Z. W., Hsu, D., and Lee, W. S., *Adaptive Informative Path Planning in Metric Spaces.* Submitted to Int. J. Robotics Research (Special issue on WAFR 2014), 2014.
14. Liu, Y. and Ozguner, U. *Human driver model and driver decision making for intersection driving*. IEEE Intelligent Vehicles Symposium, pp. 642–647, 2007.
15. Pineau, J., Gordon, G., and Thrun, S., *Point-based value iteration: An anytime algorithm for POMDPs*. International Joint Conference on Artificial Intelligence (IJCAI) 2003.
16. Silver, D., and Veness, J., *Monte-Carlo planning in large POMDPs*. Advances in Neural Information Processing Systems (NIPS). 2010.
17. Smith, T., and Simmons, R., *Heuristic search value iteration for POMDPs*, Uncertainty in AI (UAI) 2004.
18. Somani, A., Ye, N., Hsu, D., and Lee, W. S., *DESPOT: Online POMDP Planning with Regularization*, Neural Information Processing Systems (NIPS), 2013.

19. Wang, Y, Won, K. S., Hsu, D., and Lee, W. S., *Monte Carlo Bayesian Reinforcement Learning*, International Conference on Machine Learning (ICML), 2012.
20. Williams, J., and Young, S., *Partially observable Markov decision processes for spoken dialog systems*, Computer Speech & Language, 21(2), 2007.