CPS: TTP Option: Medium: Multi-objective Control of Catoptric Systems

ROGER CHAMBERLAIN, CHANDLER AHRENS, CHRIS GILL
Dept. of Computer Science and Engineering, School of Engineering and Applied Science
College of Architecture, Sam Fox School of Design & Visual Arts
Washington University in St. Louis

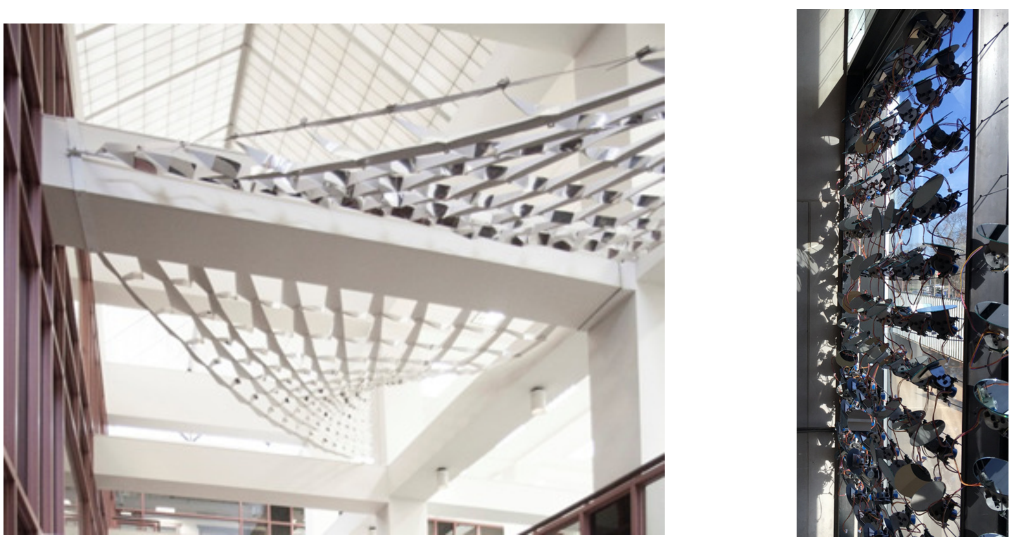
Email: roger@wustl.edu

1. INTRODUCTION

The energy consumption due to buildings (both residential and commercial) is estimated to be 20% to 40% of the total energy usage in developed countries [28], and lighting and heating are two significant components of this energy consumption [21]. Natural light (i.e., sunlight) is a readily available resource that can contribute to both the illumination [23] and heating [25] of structures, yet in the vast majority of circumstances, its use is limited to passive modalities. For example, daylighting (the use of natural light for illumination) design is dominated by passive window positioning and configuration [36] rather than active control mechanisms (see [20] for the few counterexamples). Heating systems that use sunlight do frequently use actively-controlled mirrors for tracking the relative position of the sun.

We propose to investigate the ability to effectively utilize actively controlled catoptric (mirror) surfaces to benefit the illumination and heating of buildings. Computer-based control of the dynamic positioning of individual mirrors, and computer-based management of the sunlight (as a resource), clearly put a system such as this within the scope of traditional cyber-physical systems.

Figure 1a is an image of a prototype catoptric surface (called AMP) that was designed, fabricated, and installed through an undergraduate architecture studio taught by Co-PI C. Ahrens. The installation redirects light from gable ends of an existing building into the darker recesses of the atrium to create better natural lighting where it is desired. In this installation, the mirror positions are fixed.



(a) (b)

Figure 1: Catoptric system prototypes. (a) *AMP*, TRex building, St. Louis, MO. (b) FIXME: name?, Steinberg Hall, St. Louis, MO.

In the next generation of this system, which is currently under construction, over 600 mirrors are under active, 2-axis, microprocessor-based control and therefore can be pointed in different directions dynamically as desired over time. This installation is on the south wall of the Steinberg Hall atrium (on the campus of Washington University in St. Louis), and a subset of the mirrors are shown in Figure 1b.

The ability to actively control the dynamic position of each mirror provides for unprecidented capacity to position the available natural light where it is desired. This can easily change over time, as the usage of the physical space changes.

The proposed research advances the investigation of reflected light by designing specific intensities in some areas and dissipated lighting conditions in other areas according to pre-determined, yet adjustable, image-based maps. The maps can consist of any raster image and generate the target positions for the reflected light in a space. The image is sampled according to intensity (black to white) to determine the density of target points where the higher the value, the denser the resulting field of points. Using an image-based map allows users to visualize zones of intensity in an interior space prior to the mirrors reflecting the daylight. Any user is capable of supplying or creating the image to be used for the map, thus encouraging user control and interaction with the system. The engagement of any member of a community on the creation of that image has an impact on the entire community [31] and encourages dialog about the quality and quantity of light within their environment.

Given the desire to control natural light (sunlight) via a catoptric surface, repurposing it for illumination and/or thermal management, a number of cruicial cyber-physical system issues must be addressed. This research will investigate the following questions:

1. ***What are the qualitative and quantitative benefits that can be achieved for bulding daylighting and thermal management through the use of catoptric systems?***

   Issues within this question include the ability to articulate the benefits and to quantify them effectively. Clearly, we are in the domain of multi-objective control, so the relationship between the competing goals must also be articulated and quantified. We intend to investigate the use of Markov Decision Processes (MDPs) as an approach to the multi-objective control problem, recognizing that maximization (of an objective function) in expectation is a robust way to acknowledge the inherent uncertainty of future events (whether it be sunlight availability, lighting demand, or any other effect that is stochastic in nature).

2. ***How do we provide for the safety, reliability, maintainability, and continued efficacy of these systems?***

   Even with an ideal multi-objective control system in place, the system as a whole has limited usefulness if these additional requirements are not dealt with in an effective way. Initially, just consider the issue of safety: highly concentrated sunlight aimed at a heat collector (important when harvesting energy for thermal management purposes) can be quite harmful if inadvertently aimed at humans.

   Each of these system-level requirements must ultimately be included within the optimization problem formulation, either as constraints (e.g., for safety) or as additional objectives (e.g., reliability and/or maintainability). Fortunately, the MDP formalism is well suited to the addition of concerns such as these (especially those with a stochastic nature, as reliability and maintainability tend to be).

3. ***Can we design abstractions that encapsulate subsystems for effective reuse?***

   A pair of immediate possibilities come to mind. Separating the concerns of low-level control (e.g., of mirror positions) and high-level system management (how the available light resource should be allocated) is one option. The low-level control subsystem can be encapsulated into a reusable component, applicable to any number of physical positioning applications. Similarly, a high-level management system (e.g., based upon MDP theory) could also be encapsulated in a resuable component, applicable to any number of other stochastic optimization problems. Ultimately, we would like to generalize the above into abstractions that can be leveraged more broadly for arbitrary cyber-physical systems development.

FIXME: Brief description of who we are and what we've done.

2. BACKGROUND AND RELATED WORK

FIXME: Describe first two installations.

FIXME: Literature review [3, 5, 12, 20, 23].

Markov Decision Processes (MDPs) [30] represent a general approach to modeling optimization problems and have been applied in a diverse set of application areas [38]. Examples include robotics [1], economics [4], experiment design [19], medical decisions [2], manufacturing [39], agriculture [22], and our own group's use in scheduling [16, 35] and wireless spectrum management [26].

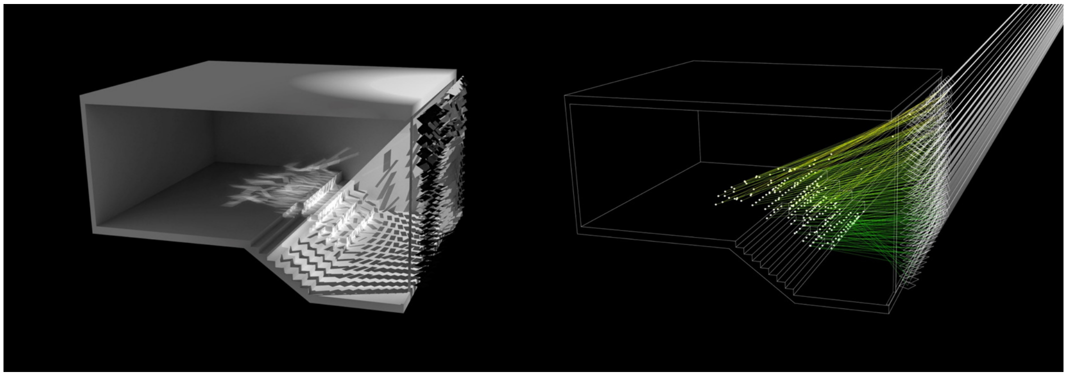


Figure 2: Ray tracing analysis of daylighting design.

Our prior research has used Markov decision process models [16] to generate resource management policies off-line [13] for non-preemptive sharing of a resource between multiple purposes at once on-line. For example, a meter-tall robot's camera (oriented by a pan-tilt unit similar to the ones we propose to use in our multi-mirror catoptric installations) may be directed downward to identify wire-frame chairs and other obstacles to navigation that other sensors on the robot may have difficulty detecting, or it may be directed upward to identify faces of people at a reception whose images it can then capture. Given distributions of the durations of intervals during which the camera would need to remain pointed in a given direction to complete an individual task, standard policy iteration techniques then can be used to generate run-time policies that in expectation maximize an objective such as adherence to a strictly proportional allocation of the resource over time [16], or even a more general definition of the utility of completing the different tasks at particular times [35]. We also showed that when different distributions of task completion intervals can occur in different modes (e.g., when a robot moves from room to room), it is possible to learn on-line what mode the system is in, or if the mode is known what the distributions are, but not both [14].

However, policy iteration is exponentially expensive, and even the memory requirements to store complete policies for on-line use may be prohibitive in resource-limited systems. We therefore focused next on the policies that were being generated from the models, and discovered consistent structure in those policies that allowed a reasonable heuristic approximation. For simple proportional sharing, a single geometric partition of a simplex could be calibrated to encode the appropriate policy accurately [15]. For utility-based resource sharing multiple disjoint heuristics were needed but the most effective one to use was clearly defined by problem parameters [34].

As a further illustration both of the applicability of MDP-based policy iteration to generate effective resource management policies, we applied similar techniques to manage a much different resource: the transmission spectrum in wireless networks [27]. Although the semantics of that resource differed radically from the pan-tilt camera, the MDP models were reasonably similar. We extended the basic model to include modulation as well as admission decisions, discovered and characterized common structure among the policies that were generated, and again obtained efficient and effective heuristic policies for on-line use [26].

In this proposal we adopt the definition used by Glaubius et al. [16] of a (discrete-time) Markov decision process as a 5-tuple $(\mathcal{X}, \mathcal{A}, T, R, \gamma)$, with ***states*** designated as $\chi \in \mathcal{X}$, ***actions*** designated as $a \in \mathcal{A}$, and a transition system, T, which gives the probability $P_T(\chi' \mid \chi, a)$ of transitioning from state χ to state χ' on action a. The reward function $R(\chi, a, \chi') \in \mathbb{R}_{\geq 0}$ describes the reward that accrues when transitioning from state χ to state χ' via action a, under a discount factor, γ, to ensure convergence of the long term reward.

FIXME: Add transition into the next section.

3. RESEARCH DESCRIPTION

3.1. Question 1 – What are the qualitative and quantitative benefits that can be achieved for building daylighting and thermal management through the use of catoptric systems?

There is ample literature that documents the benefits of natural light in human-occupied spaces [23, 24, 29], yet the thermal effects can also be significant [6] (i.e., too much sunlight can increase temperature to an unacceptable degree).

While many studies focus on the benefits of a homogeneous light levels within a space [3, 5], there is often differing need for light level based on user. The increased control due to the independent positioning of the mirrors can provide functional benefits for people that need more or less light due to age of the occupant or their task. For example, someone using a computer needs less light since the screen is a source of artificial light while someone reading from physical media like a book requires higher light levels. The light level required for vision varies with the age of the occupants. As people get older, they need higher light levels and increased luminous contrast. Furthermore, everyone regardless of age have different qualities of vision. Therefore, having a system that can accommodate these varying conditions using daylight rather than artificial light is beneficial.

For thermal management, the literature on harvesting sunlight for space heating is substantial (see, e.g., [11, 17, 18, 25, 33, 37], and we do not propose to contribute anything new to that field. We will simply leverage what are already well-understood techniques. It is worth acknowledging that many sunlight-based heating systems use catoptric systems (primarily for sun tracking). Our contributions here will be to integrate a sunlight harvesting system into dual-purpose use, both heating and illumination.

We will explore the use of Markov decision processes for multi-objective control of catoptric systems. We will start with a single objective, matching a provided image map that represents a desired illumination pattern, and then generalize to include the competing goal of thermal energy harvesting.

For the purpose of designing an MDP, we will assume that a low-level control mechanism exists that can position each individual mirror as desired. The MDP will model the higher-level management issues, e.g., which mirror should be pointed in which direction? With a low-level controller in place, the state space, $\mathcal{X}$, can encapsulate the set of mirrors pointed at each position in the image map. The set of actions, $\mathcal{A}$, will embody the movement of one or more mirrors from their present position to a new position, and the transition system, T, will encode the probabilistic variations present in the available sunlight. This allows the reward function, R, to be a quantified match between the desired image map and the resulting daylight that is directed to each position in the physical space.

After an initial exploration of the MDP formulated as described above, we will then proceed to expand the framework to include the additional objective of harvesting heat. We will quantify the benefits of each objective using general utility functions [35] in a manner previously explored by our group. One candidate set of utility functions we will investigate will be to prioritize daylighting performance, and only allocate excess sunlight to the HVAC subsystem. A second alternative will be to proportionally provide sunlight to each goal up to the point that the illumination can no longer benefit, at which point all additionally available sunlight goes into heating.

In each of the above MDPs under investigation, we will utilize the value-optimal solution (guaranteed to exist within Markov decision process theory); however, this solution typically must be computed off-line, since it is, in general, exponentially expensive to compute. In parallel, we will explore the space of value-optimal solutions and seek to formulate an inexpensive to compute heuristic that closely approximates the true value-optimal solution.

3.2. Question 2 – How do we provide for the safety, reliability, maintainability, and continued efficacy of these systems?

The ability to control mirror position is only useful over time if the additional considerations of safety, reliability, etc., are managed properly for the system as a whole. To address these issues, we will investigate the degree to which they can be handled within the context of the low-level controller versus being dealt with within the high-level complete system control.

We anticipate some of these considerations being incorporated as constraints within the optimization process and others as additional goals that we wish to maximize during the multi-objective optimization itself. Let us first consider those that will be addressed as constraints (e.g., safety).

Constraints. A commonly used safety constraint on a positioning subsystem is to specify hard limits on the range of motion (e.g., of each mirror, individually). This kind of limit can be imposed by the physical design of the pan-tilt mechanics, simply by adding physical stops at the limit positions.

The more interesting constraint systems occur when the safety considerations are no longer locally determinable, but are dependent upon context. An example that is relevant for our catoptric surfaces is a limit on the total light intensity that can be supported at various positions in the field of view of the mirrors. When providing heat into the HVAC system, we desire a high light concentration delivered to the heat transfer point. When illuminating a physical space for human occupants, the above levels of light concentration are not only undesirable, they are patently unsafe.

These context sensitive constraints add two additional requirements. First, the higher-level management system must be responsible for them. In our MDP formulation, we much adjust the feasible state space, $\mathcal{X}$, so that they are unreachable (or have sufficiently negative reward as to amount to the same thing). Second, we must also ensure that whatever choices are made by the optimization system (whether it be via MDPs or some other approach), the actual mirror positions are still constrained so as to not result in an unsafe condition. This will require feedback of some form on the actual mirror positions and the ability to determine whether or not the actual positions differ from the controlled positions. This might be accomplished either locally (e.g., via shaft encoders on the pan-tilt mechanism) or globally (e.g., via a visual monitoring system and appropriate image analysis software).

Goals. Many of the additional considerations are more appropriately addressed as additional (potentially competing) goals, adding to the desires of daylighting and heating. An example here would be the impact on reliability of the system when individual mirrors are moved. As with all electro-mechanical systems, the pan-tilt mechanism has a limited lifetime, which can be significantly impacted by its usage duty cycle (i.e., the more it is moved, the sooner it will fail).

This is precisely the set of circumstances in which Markov decision processes excel. FIXME: defend this statement.

3.3. Question 3 – Can we design abstractions that encapsulate subsystems for effective reuse?

FIXME: We will investigate the viability and utility of two candidate abstractions: direct mirror control and MDP control [10].

3.4. Intellectual Merit

The intellectual contributions of this project are FIXME: describe summary of intended intellectual merit.

4. EVALUATION/EXPERIMENTATION PLAN

FIXME: Two sites identified:

1. **VelociData, Inc.**, 10425 Old Olive Stree Rd., St. Louis, MO. FIXME: VelociData is located in the recently announced ***39 North*** innovation district, which has the Danforth Plant Science Center, Monsanto, Bio Research & Development Growth Park, and Heliz Center Biotech Incubator as anchors.

2. **BECS Technology, Inc.**, 10818 Midwest Industrial Dr., St. Louis, MO. FIXME: BECS Technology is a small manufacturer of electronic control systems in a number of markets (e.g., agriculture, aquatics, refrigeration). We will have access to the HVAC system in their building.

5. PROJECT MANAGEMENT AND COLLABORATION PLAN

6. BROADER IMPACTS

FIXME: Describe broader impacts: environmental benefits of energy savings and quality of life benefits to building occupants.

At the undergraduate education level, this work is closely related to FIXME: describe CSE 132 connection.

At the graduate education level, this work will support 4 graduate students at Washington Univ. in St. Louis. FIXME: Expand, including REUs, multidisciplinary angle.

We will leverage a pair of existing university programs to help us attract students from traditionally underrepresented groups. The Olin Fellowship Program (for women) and the Chancellor's Fellowship Program (aimed at underrepresented minority students) have had a successful track record of enabling individuals to pursue graduate study. In our experience, the most effective method for attracting students from underrepresented groups is by personal contact with a suitable role model. To facilitate this, we regularly ask the appropriately qualified individuals in our group to be actively involved in the recruiting process. This cohort currently includes two minority graduate students (one African-American student and one hispanic student). FIXME: Can we strengthen the BPC story? Maybe somehow with 132 and maker spaces?

7. RESULTS FROM PRIOR NSF SUPPORT

CSR: Small: Concurrent Accelerated Data Integration (CNS-1527510, PI R. Chamberlain, co-PI Ron Cytron), 10/2015–9/2019, $519,275.

Intellectual Merit – This project investigates the accelerated execution of data integration workflows, which increasingly are bottlenecks in data science. Execution platforms being targeted include both graphics engines and FPGAs.

Broader Impacts – This research project has supported 3 graduate students and 4 REU students. The applications investigated come from the fields of computational biology, astrophysics, and the Internet of Things, further expanding the scope of the students' experience.

Evidence of Research Products and their Availability – Publications resulting from this work include [8, 9, 26, 32]. A benchmark suite of the above workflows has been released as a community resource [7].

CPS: Medium: Collaborative: CyberMech, a Novel Run-Time Substrate for Cyber-Mechanical Systems (CNS-1136073 and CNS-1136075, Washington University: PI C. Gill, co-PIs Kunal Agrawal and Chenyang Lu; Purdue University: PI Arun Prakash, co-PI Shirley Dyke), 9/2011-8/2016, $1,800,000 total.

This research project developed novel foundations for parallel real-time computing, and used them to demonstrate the first ever real-time hybrid simulation involving a thousand-degree-of-freedom structure.

Intellectual Merit – Results of this research include new methods for parallel real-time execution of control and simulation computations, new parallel real-time scheduling techniques and analyses, and characterization and exploitation of trade-offs involving both high computational demand and stringent timing constraints.

Broader Impacts – This multi-university project involved 7 PhD, 3 masters, and 7 undergraduate students, and 2 visiting scholars in highly multi-disciplinary research collaborations. Results of this research appeared in 10 publications at top-tier conferences and journals.

Evidence of Research Products and their Availability – Data, experiment configurations, platform software, and simulation source-code have been published on-line at Washington University and Purdue University.

REFERENCES

[1] N. Achour and K. Braikia. An MDP-based approach oriented optimal policy for path planning. In *Proc. of International Conference on Machine and Web Intelligence*, pages 205–210, Oct. 2010.

[2] O. Alagoz, H. Hsu, A. J. Schaefer, and M. S. Roberts. Markov Decision Processes: A Tool for Sequential Decision Making Under Uncertainty. *Medical Decision Making*, 30(4):474–483, 2010.

[3] M. Alrubaih, M. Zain, M. Alghoul, N. Ibrahim, M. Shameri, and O. Elayeb. Research and development on aspects of daylighting fundamentals. *Renewable and Sustainable Energy Reviews*, 21:494–505, Feb. 2013.

[4] A. Briggs and M. Sculpher. An Introduction to Markov Modelling for Economic Evaluation. *PharmacoEconomics*, 13(4):397–409, 1998.

[5] B. Bueno, J. Wienold, A. Katsifaraki, and T. E. Kuhn. Fener: A radiance-based modelling approach to assess the thermal and daylighting performance of complex fenestration systems in office spaces. *Energy and Buildings*, 94:10–20, Feb. 2015.

[6] C. Buratti, E. Moretti, E. Bellone, and F. Cotana. Unsteady simulation of energy performance and thermal comfort in non-residential buildings. *Building and Environment*, 59:482–491, Jan. 2013.

[7] A. M. Cabrera, C. J. Faber, K. Cepeda, R. Derber, C. Epstein, J. Zheng, R. K. Cytron, and R. D. Chamberlain. Data Integration Benchmark Suite v1. DOI: `http://dx.doi.org/10.7936/K7NZ8715`, Apr. 2018.

[8] A. M. Cabrera, C. J. Faber, K. Cepeda, R. Derber, C. Epstein, J. Zheng, R. K. Cytron, and R. D. Chamberlain. DIBS: A data integration benchmark suite. In *Proc. of ACM/SPIE Int'l Conf. on Performance Engineering Companion*, pages 25–28, Apr. 2018.

[9] R. D. Chamberlain. Assessing user preferences in programming language design. In *Proc. ACM Int'l Symp. on New Ideas, New Paradigms, and Reflections on Programming and Software*, pages 18–29, Oct. 2017.

[10] R. D. Chamberlain, C. Ahrens, and C. Gill. Abstractions for cyber-physical systems development: An international opportunity. In *Visioning Workshop for International Networks to Advance CPS Research, Development, and Education Worldwide*, Apr. 2018. Available at `https://cps-vo.org/node/48624`.

[11] F. de Winter. Heat exchanger penalties in double-loop solar water heating systems. *Solar Energy*, 17(6):335–337, 1975.

[12] A. Galatioto and M. Beccali. Aspects and issues of daylighting assessment: A review study. *Renewable and Sustainable Energy Reviews*, 66:852–860, Sept. 2016.

[13] R. Glaubius, T. Tidwell, C. Gill, and W. Smart. Scheduling policy design for autonomic systems. *International Journal of Autonomous and Adaptive Communication Systems*, 2(3):276–296, June 2009.

[14] R. Glaubius, T. Tidwell, C. Gill, and W. Smart. Real-time scheduling via reinforcement learning. In *Proc. of 26th Conference on Uncertainty in Artificial Intelligence (UAI)*, July 2010.

[15] R. Glaubius, T. Tidwell, B. Sidoti, D. Pilla, J. Meden, C. Gill, and W. Smart. Scalable scheduling policy design for open soft real-time systems. In *Proc. of IEEE Real-Time and Embedded Technology and Applications Symposium (RTAS)*, Apr. 2010.

[16] R. Glaubius, T. Tidwell, W. Smart, and C. Gill. Scheduling design and verification for open soft real-time systems. In *Proc. of Real-Time Systems Symposium*, pages 505–514, Nov. 2008.

[17] A. J. Hunt. A new solar thermal receiver utilizing a small particle heat exchanger. In *Proc. of 14th Intersociety Energy Conversion Engineering Conference*, pages 159–163, Aug. 1979.

[18] S. A. Klein, W. A. Beckman, and J. A. Duffie. A design procedure for solar heating systems. *Solar Energy*, 18(2):113–127, 1976.

[19] M. Kolonko and H. Benzing. The sequential design of Bernoulli experiments including switching costs. *Operations Research*, 33(2):412–426, Apr. 1985.

[20] M. Konstantoglou and A. Tsangrassoulis. Dynamic operation of daylighting and shading systems: A literature review. *Renewable and Sustainable Energy Reviews*, 60:268–283, Feb. 2016.

[21] M. Krarti, P. M. Erickson, and T. C. Hillman. A simplified method to estimate energy savings of artificial lighting use from daylighting. *Building and Environment*, 40:747–754, 2005.

[22] A. R. Kristensen. A general software system for Markov decision processes in herd managment applications. *Computers and Electronics in Agriculture*, 38(3):199–215, Mar. 2003.

[23] R. Leslie. Capturing the daylight dividend in buildings: why and how? *Building and Environment*, 38:381–385, 2003.

[24] D. H. W. Li and J. C. Lam. Evaluation of lighting performance in office buildings with daylighting controls. *Energy and Buildings*, 33:793–803, 2001.

[25] P. J. Lunde. *Solar Thermal Engineering: Space Heating and Hot Water Systems*. John Wiley & Sons, Inc., New York, NY, USA, 1980.

[26] J. Meier, C. Gill, and R. D. Chamberlain. Combining admission and modulation decisions for wireless embedded systems. In *Proc. IEEE 19th Int'l Symp. Real-Time Distributed Computing*, pages 69–78, May 2016.

[27] J. Meier, S. Sistla, B. Karaus, C. Gill, R. Chamberlain, and T. Tidwell. Assessing the appropriateness of using Markov decision processes for RF spectrum management. In *Proc. of 16th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM)*, Nov. 2013.

[28] L. Pérez-Lombard, J. Ortiz, and C. Pout. A review on buildings energy consumption information. *Energy and Buildings*, 40:394–398, 2008.

[29] P. Plympton, S. Conway, and K. Epstein. Daylighting in schools: Improving student performance and health at a price schools can afford. In *Proc. of American Solar Energy Society Conference*, June 2000.

[30] M. L. Puterman. *Markov Decision Processes*. John Wiley & Sons, Inc., Hoboken, NJ, 1994.

[31] B. Schaban-Maurer. *Rise of the Citizen Practitioner: A Phronesis-Based Approach to Citizen Engagement and Social Policy*. Scholar's Press, 2013.

[32] J. A. Shidal. *Exploiting the Weak Generational Hypothesis for Write Reduction and Object Recycling*. PhD thesis, Dept. of Computer Science and Engineering, Washington University in St. Louis, May 2016.

[33] K. Szeibel, J. Mason, and V. Fthenakis. A solar grand plan. *Scientific American*, 298(1):64–73, Jan. 2008.

[34] T. Tidwell, C. Bass, E. Lasker, M. Wylde, C. Gill, and W. Smart. Scalable utility aware scheduling heuristics for real-time tasks with stochastic non-preemptive execution intervals. In *Proc. of 23rd Euromicro Conference on Real-Time Systems (ECRTS)*, pages 238–247, July 2011.

[35] T. Tidwell, R. Glaubius, C. D. Gill, and W. D. Smart. Optimizing expected time utility in cyber-physical systems schedulers. In *Proc. of IEEE 31st Real-Time Systems Symposium*, pages 193–201, Dec. 2010.

[36] D. Vázquez-Moliní, M. González-Montes, A. Álvarez Fernández-Balbuena, Ángel García-Botella, W. Pohl, T. Galan, and E. Bernabéu. Horizontal daylighting system for office buildings. *Energy and Buildings*, 67:525–530, 2013.

[37] R. Z. Wang and R. G. Oliveira. Adsorption refrigeration—An efficient way to make good use of waste heat and solar energy. *Progress in Energy and Combustion Science*, 32(4):424–458, 2006.

[38] D. J. White. A Survey of Applications of Markov Decision Processes. *The Journal of the Operational Research Society*, 44(11):1073–1096, Nov. 1993.

[39] K. K. Yin, G. G. Yin, and H. Liu. Stochastic modeling for inventory and production planning in the paper industry. *AIChE Journal*, 50(11):2877–2890, Nov. 2004.