Reinforcement Learning-based Control of Building Subsystems

Omid Ardakanian
Department of Computing Science
University of Alberta

March 2021

UNIVERSITY OF ALBERTA

Joint work with Tianyu Zhang, Gaby Baasch, and Ralph Evins**

SCRG

Why do we even care?

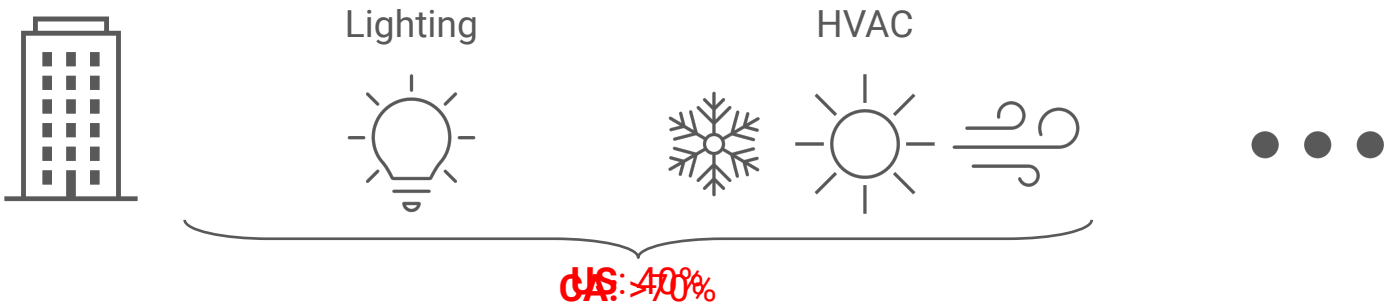
- The commercial sector in the United States used 9.3 quads (**9.8×10^{18}** joules) in 2019. HVAC and lighting systems accounted for ~40% of this energy consumption

Energy Consumption of Different Sectors

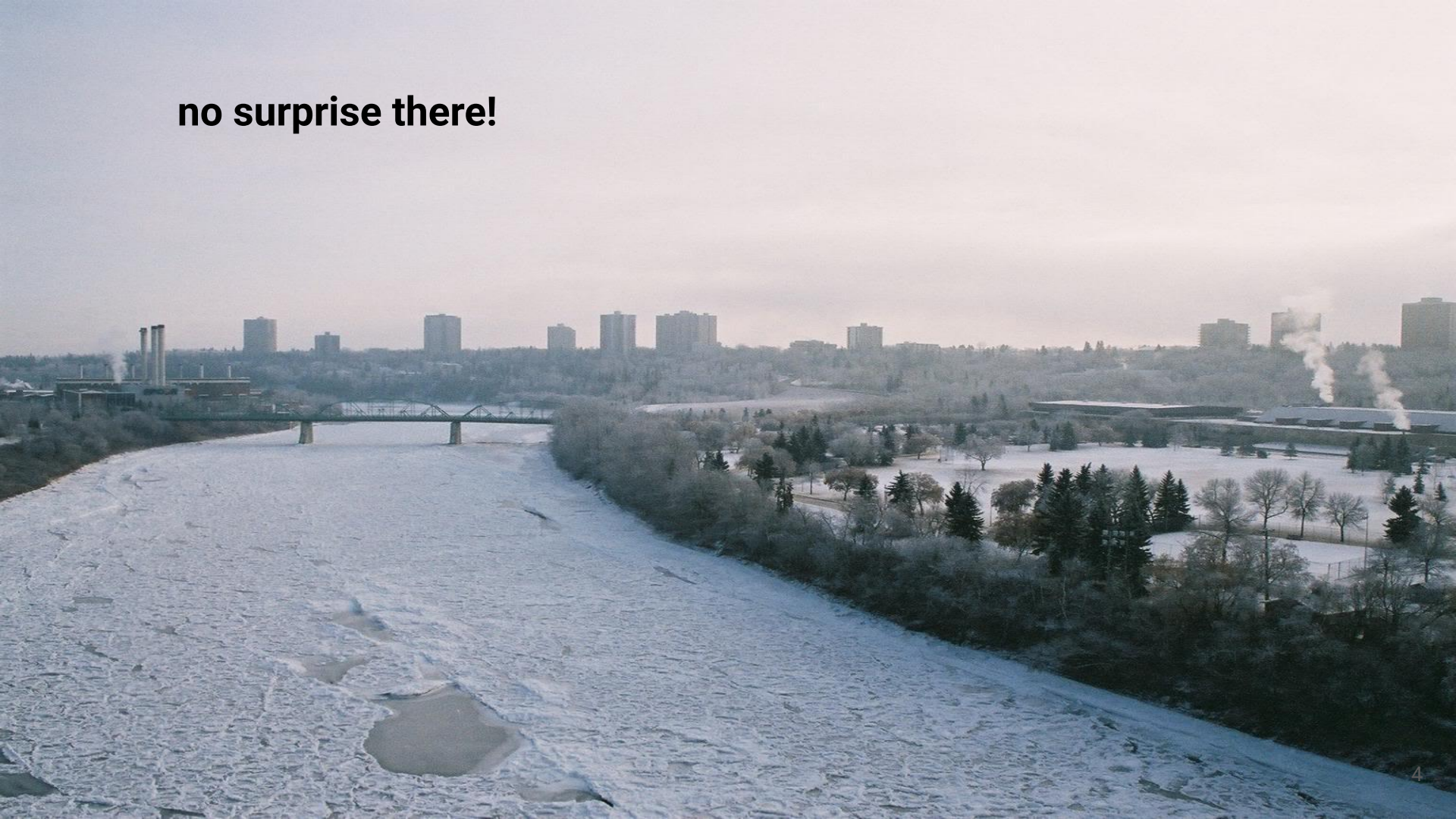
US: 12.3%

What are the most energy consuming building systems?

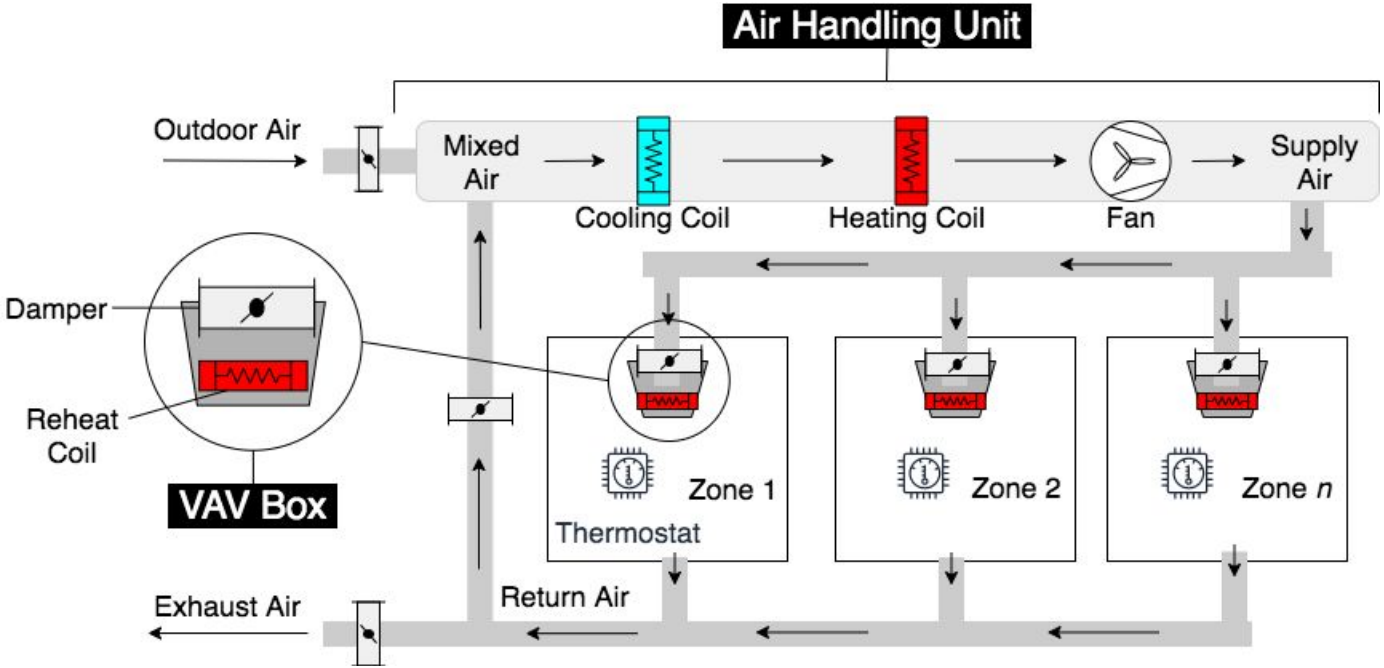
- The commercial sector in the United States used 9.3 quads (9.8 exajoules) in 2019. HVAC and lighting systems accounted for ~40% of this energy consumption
- In countries with a colder climate, HVAC and lighting are typically responsible for a larger percentage of building energy use
  - in Canada more than **70%** of building energy use can be attributed to HVAC and lighting



no surprise there!



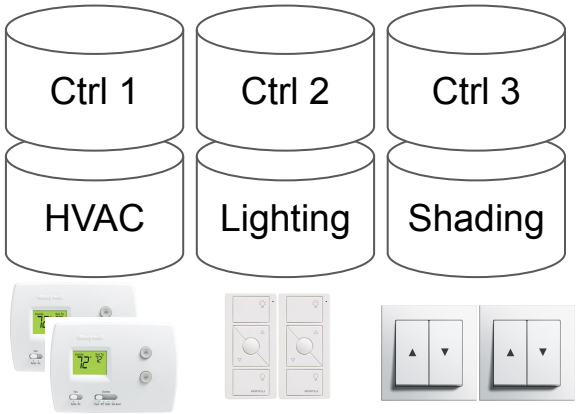
Control points in the HVAC system

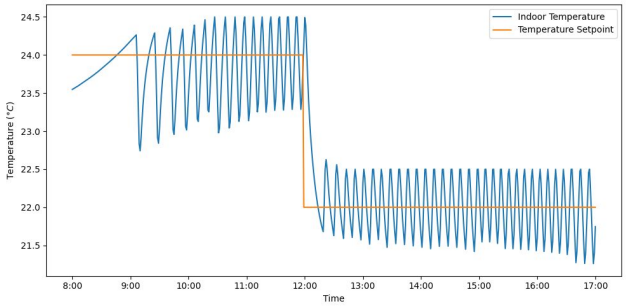


How are they controlled today?

- Building subsystems are controlled **separately** using rule-based and model-based controllers
  - **Why?** they are vertically integrated
  - **Goal:** solve a multi-criterion optimization problem to reduce energy use while maintaining comfort (and indoor air quality and ...)

MPC,
PID, etc.

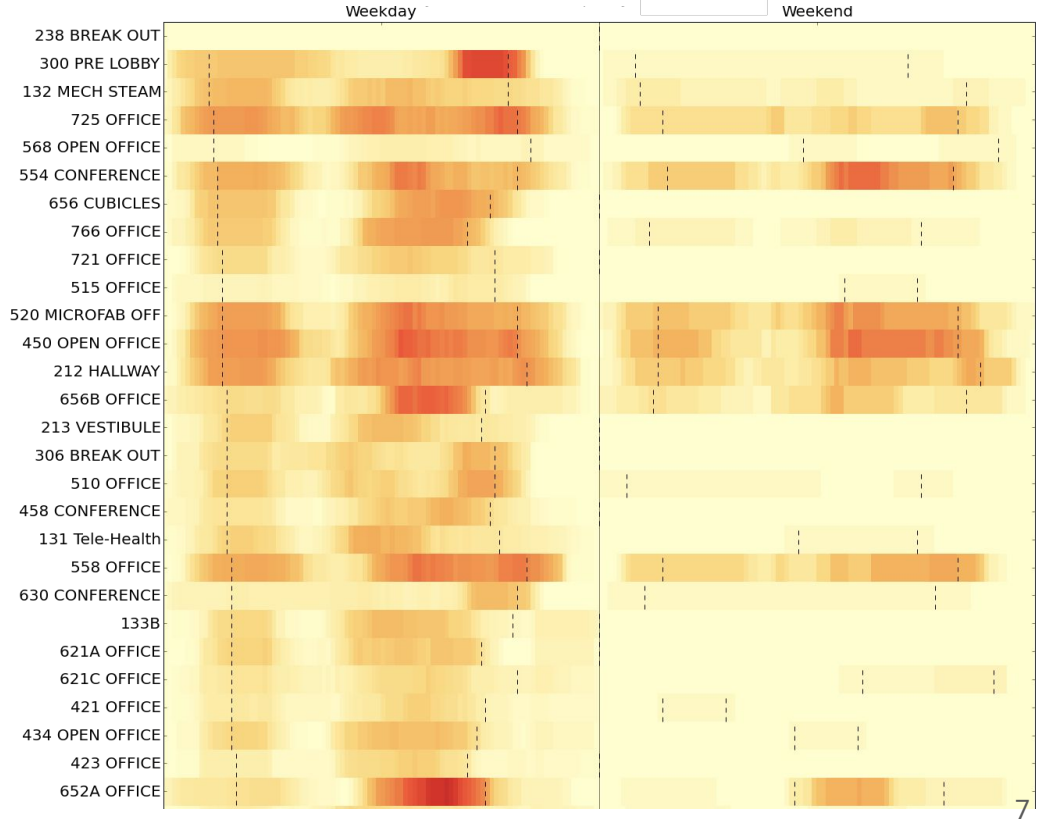




Recent efforts

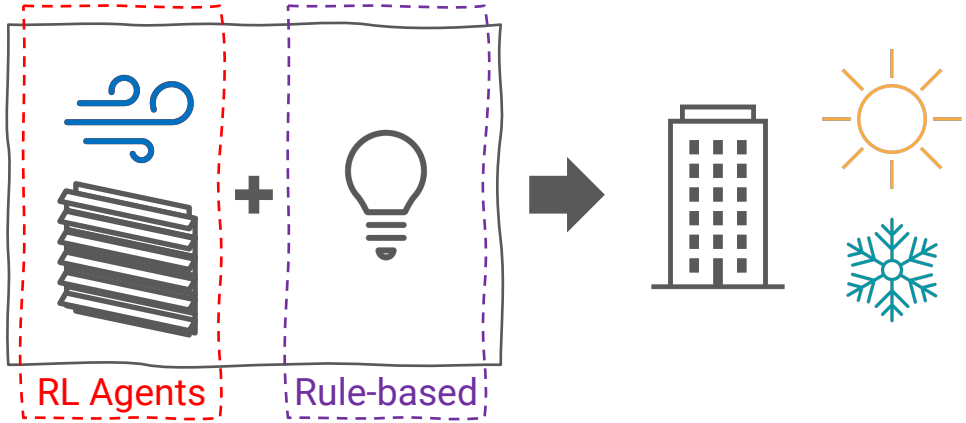
- A lot of work has been done in recent years to estimate and predict the occupancy state of a zone and to incorporate this information in control loops

- But there is still room for saving energy (and reducing costs) without sacrificing comfort...



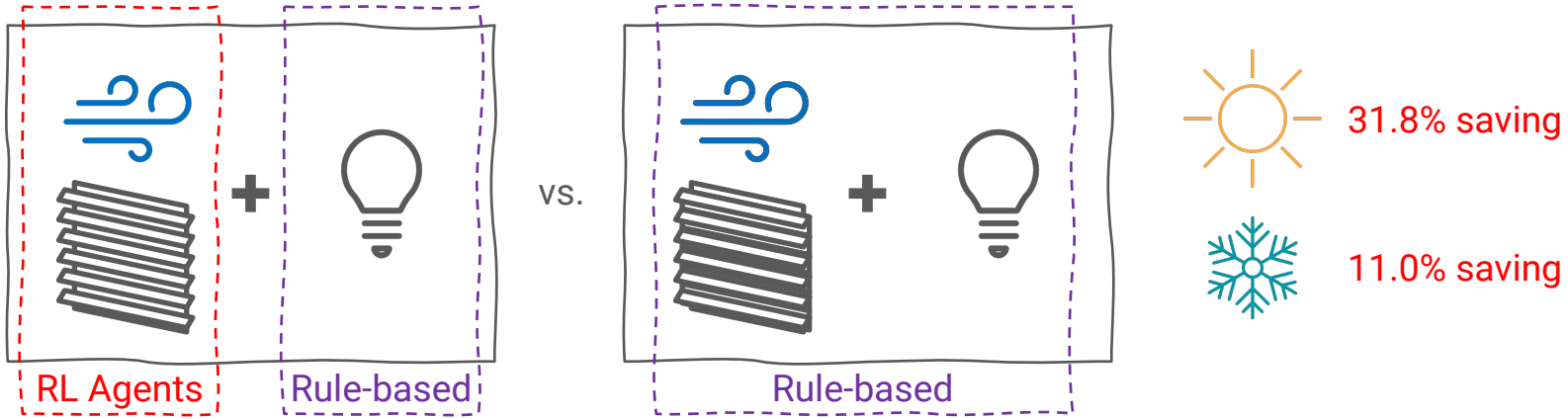
Proposed Solution

- Jointly control the HVAC system supply air temperature and blind angle setpoints using model-free RL algorithms with and without auto dimming of lights



Proposed Solution

- Jointly control the HVAC system supply air temperature and blind angle setpoints using model-free RL algorithms with and without auto dimming of lights
- Our results show that we can save 11% more energy in winter and 31.8% more energy in summer over existing rule-based control strategies that rely on zone-level occupancy information



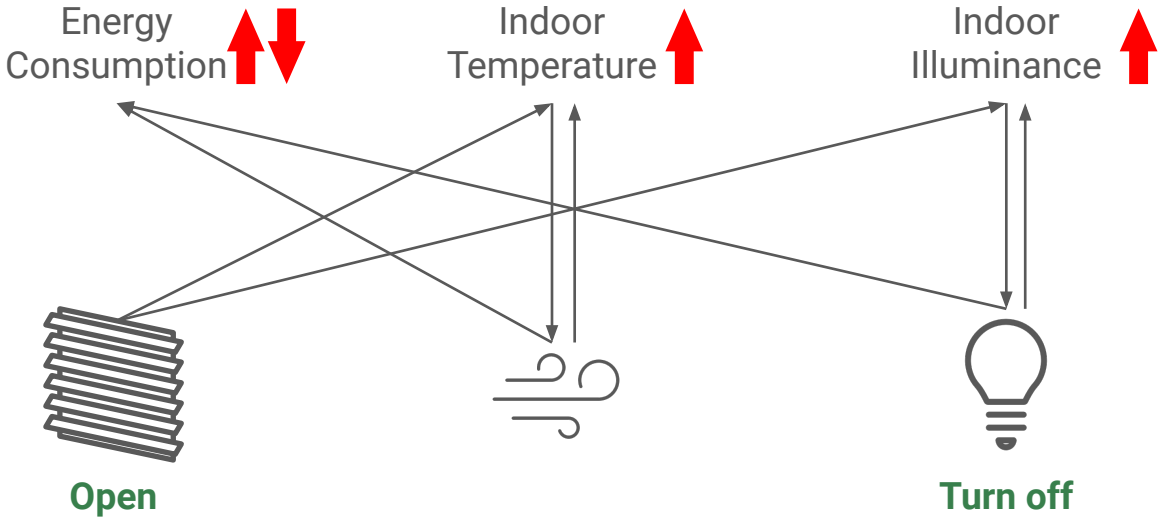
We hope to address these research questions

- How does the joint control of building systems affect the **whole-building energy use**?
- What are the best trade-offs between energy use, thermal comfort, and visual comfort?
  - facilities managers may need to trade energy savings for extra comfort
- Will incorporating zone-level occupancy information noticeably change the performance of a control policy?
  - is it worth the money or the hassle to install occupancy sensors in each zone?
- How does the performance of a given control policy vary across seasons?
  - the outside air temperature can affect the control performance

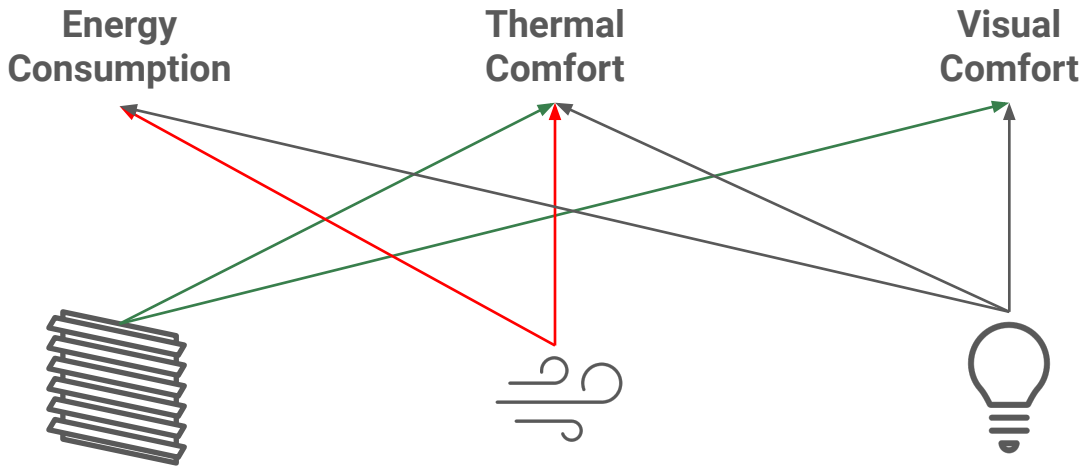
Outline

- Challenges
  - why model-based control strategies don’t work that well?
- Model-free RL-based control of building subsystems
  - Scenarios
  - Baselines
- Example results
- Takeaways

Building subsystems have complex interactions



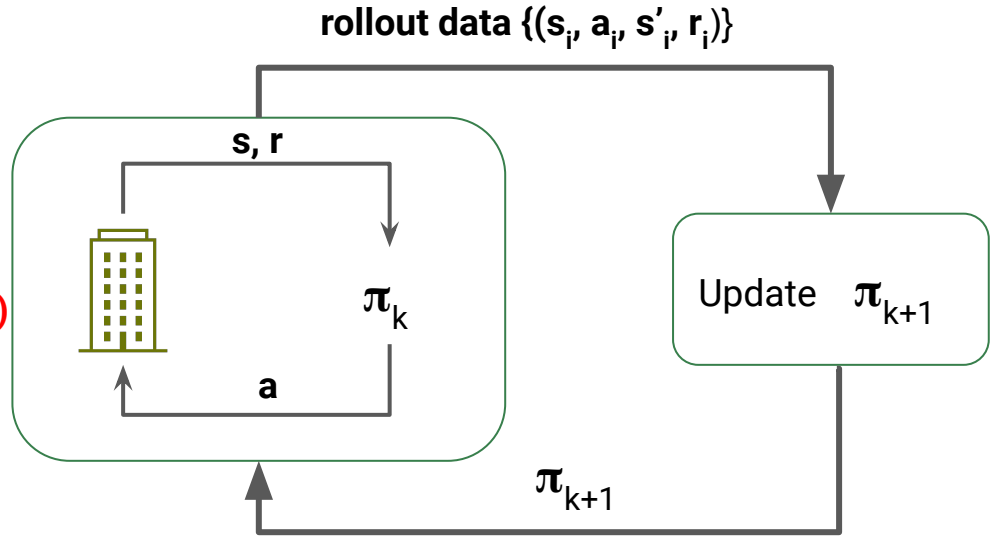
Building subsystems have complex interactions



it is difficult to model these interactions and understand their impacts on our objectives

Model-free control

- **Basic idea:** learn a control **policy** (a sequence of setpoint values) through interactions with building systems or a simulated environment (*EnergyPlus+COBS**)

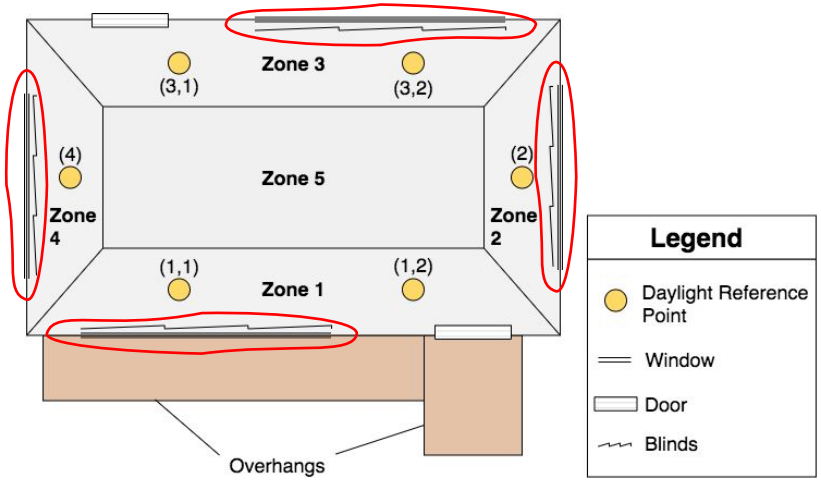


generate samples (i.e., run the policy)

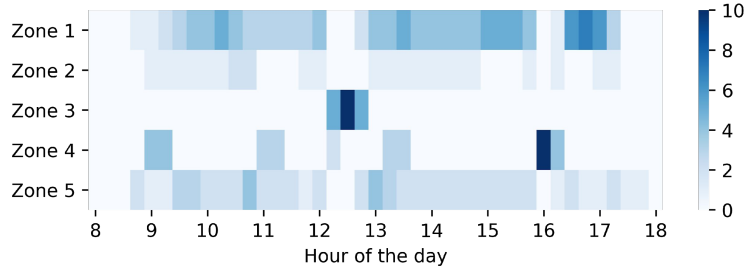
fit a model to estimate return and improve the policy

** Zhang, et al. 2020. COBS: Comprehensive Building Simulator. In Proceedings of the 7th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation (BuildSys '20). ACM, 314–315. DOI:https://doi.org/10.1145/3408308.3431119*

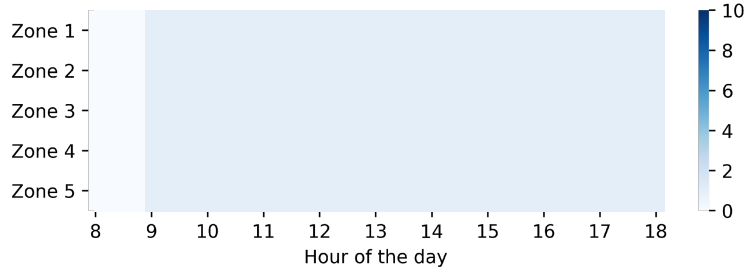
Environment



5-zone test building



Zone-level occupancy



Building-level occupancy

RL Problem Formulation

- State
  - Temperature in each zone
  - Number of occupants in each zone
  - Hour of the day
  - Slat angle of each blind
  - Ambient temperature (+ forecasts for the next 3 hours)
  - Solar radiation (+ forecasts for the next 3 hours)
- Action
  - Supply air temperature setpoint
  - Blind angles

RL Problem Formulation

- Reward
  - Weighted sum of three objectives
    - E: Energy Consumption
    - T_c: Thermal Comfort
    - V_c: Visual Comfort

$$R = -\rho_E \mathrm{Norm}(E) - \rho_T \mathrm{Norm}(T_c) - \rho_V \mathrm{Norm}(V_c)$$

reward parameters

$$T_{ci} = \begin{cases} 0, & |PMV_i| \le 0.5 \\ |PMV_i| - 0.5, & otherwise. \end{cases}$$

$$V_{ci} = \begin{cases} 0 & 300 \le \mathbb{E}[I_i] \le 750 \\ 300 - \mathbb{E}[I_i], & \mathbb{E}[I_i] < 300 \\ \mathbb{E}[I_i] - 750, & \mathbb{E}[I_i] > 750, \end{cases}$$

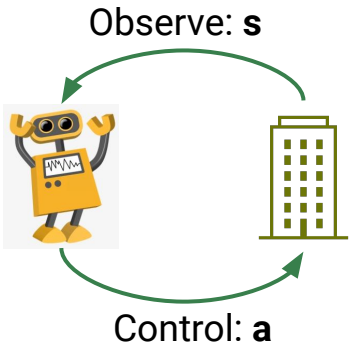
Control Scenarios

HVAC	Blinds	Lights
SAT setpoint	Always open	Not controlled
SAT setpoint	Always open	Auto Dimming
SAT setpoint	Using the same setpoint	Not controlled
SAT setpoint	Using the same setpoint	Auto Dimming
SAT setpoint	Using different setpoints	Not controlled
SAT setpoint	Using different setpoints	Auto Dimming

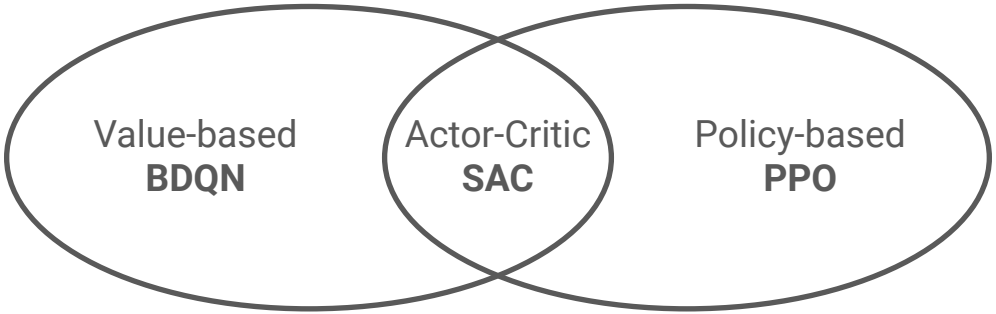
Baselines

- 4 rule-based baselines
  - Controlling HVAC
  - Controlling HVAC and blinds
  - Controlling HVAC with auto-dimming of lights
  - Controlling HVAC and blinds with auto-dimming of lights
- HVAC setpoint is controlled by **a reactive controller** that estimates the heating/cooling load in each zone and adjusts setpoints accordingly

RL Agents



- We used three state-of-the-art model-free RL algorithms
  - Branching Dueling Q-Network (BDQN)
  - Soft Actor-Critic (SAC)
  - Proximal Policy Optimization (PPO)



Research Questions and Results

- How does the joint control of building systems affect the whole-building energy use?
  - Adding motorized blinds makes possible:
    - 15.2% additional energy savings in winter
    - 11.7% additional energy savings in summer
  - Using auto-dimming with HVAC control, we can save:
    - -0.1% on whole-building energy use in winter
    - 27.9% on whole-building energy use in summer
  - By taking advantage of both we can reduce the whole-building energy consumption by:
    - 18.3% in winter
    - 32.0% in summer

Control Scenario	Baseline Number	Month	Baseline (MWh)
SAT setpoint Blinds always open	(1)	January July	8.34 4.44
SAT setpoint Blinds always open Auto dimming	(3)	January July	8.4 3.2
SAT setpoint Single blind setpoint	(2)	January July	7.07 3.92
SAT setpoint Single blind setpoint Auto-dimming	(4)	January July	6.81 3.02

Research Questions and Results

- What are the best trade-offs between energy use, thermal comfort, and visual comfort?

We found that the PPO agent is highly sensitive to the reward weights, whereas BDQN and SAC are not that sensitive

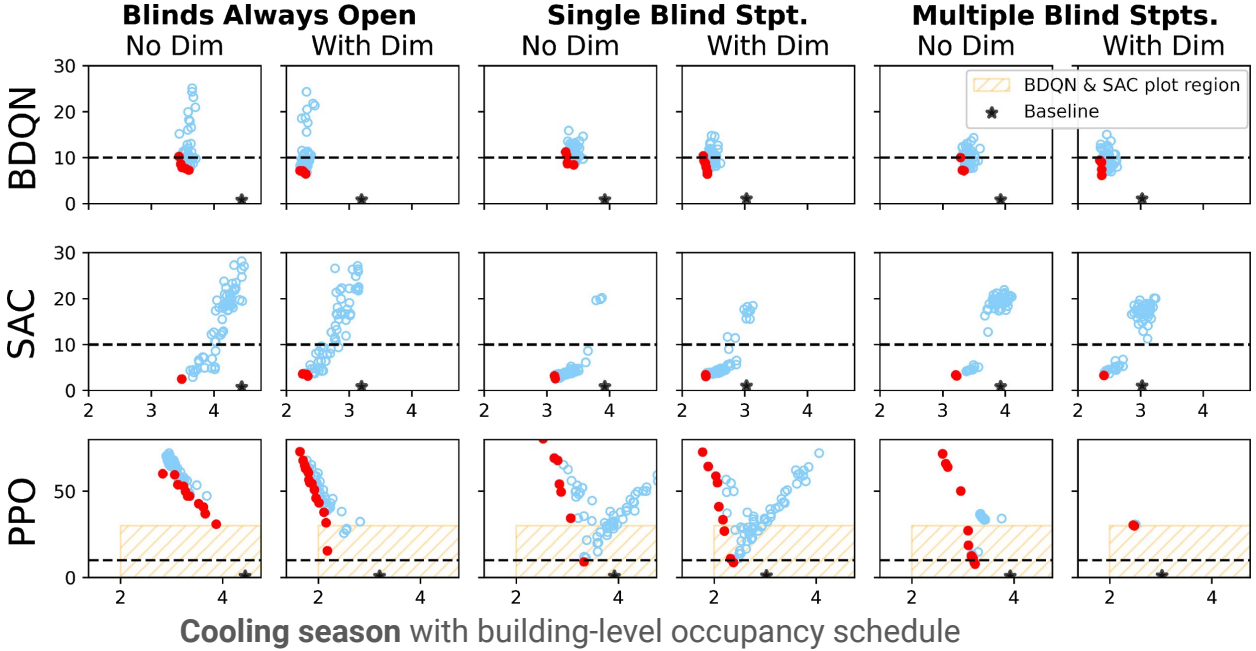
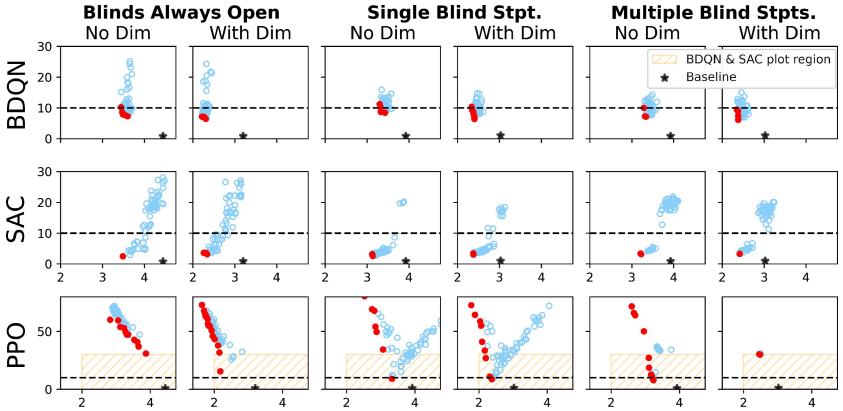


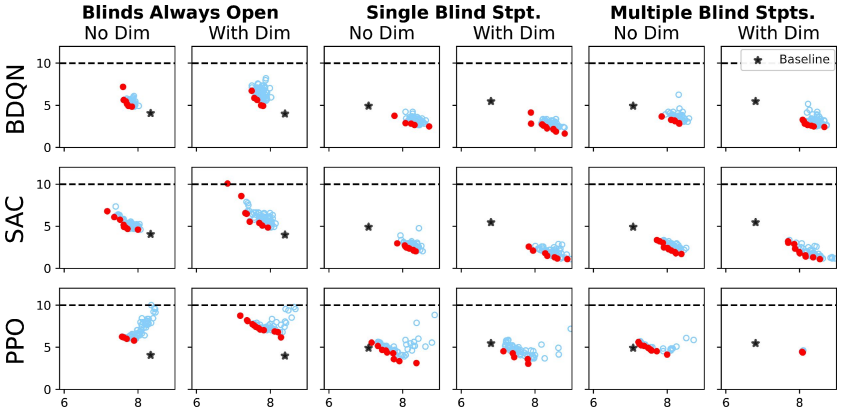
Figure: the PMV violation rate (y-axis) versus the monthly electricity consumption in MWh (x-axis) for different reward parameters

Research Questions and Results

- How does the performance of a given control policy vary across seasons?



Cooling season with building-level occupancy schedule



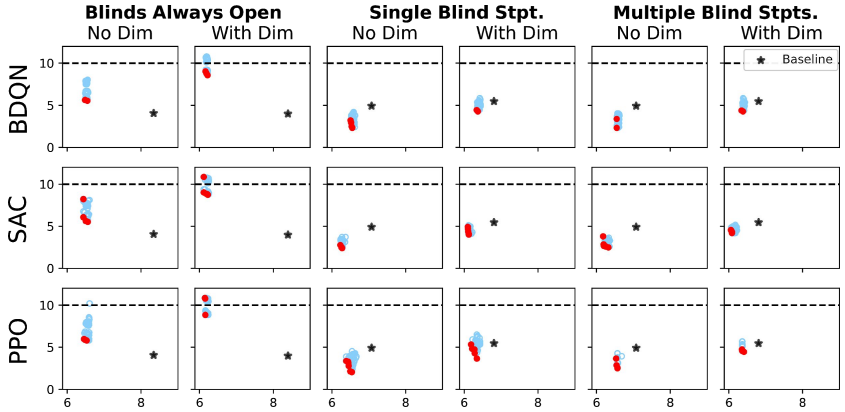
Heating season with building-level occupancy schedule

Research Questions and Results

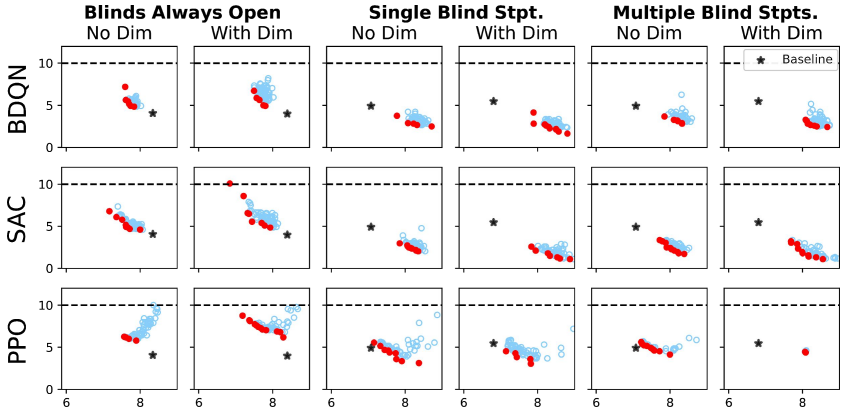
- Will incorporating zone-level occupancy information noticeably change the performance of a control policy?

  Incorporating zone-level occupancy info. offers additional energy savings of 6.24MWh ($437) per annum

  This can offset the cost of buying and installing occupancy sensors in the 5 zones



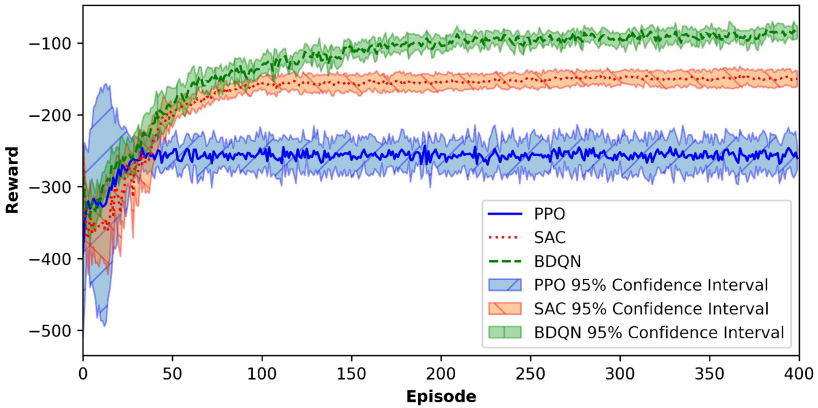
Heating season with **zone-level** occupancy schedule



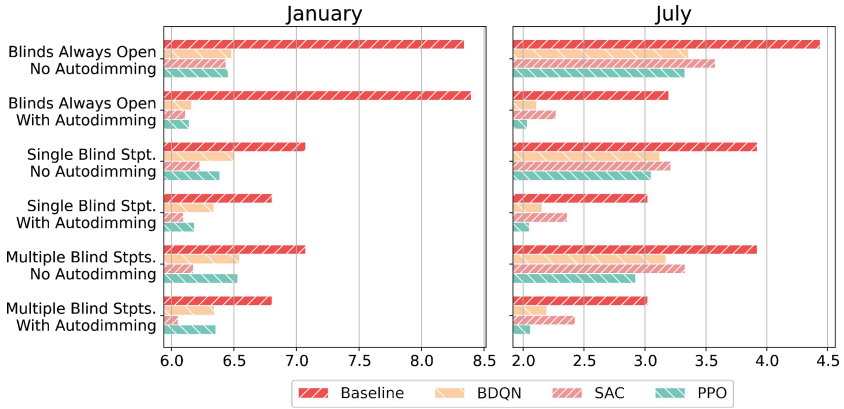
Heating season with **building-level** occupancy schedule

Research Questions and Results

- Comparing RL agents in terms of convergence speed and achieve reward



Performance comparison of three RL algorithms on the building control domain



Energy consumption of different RL agents using zone-level occupancy information

Takeaways

- The joint control of building subsystems could provide a better trade-off compared to when they are controlled separately
    - this is true even when we adopt rule-based baselines
    - it's essential to adopt a model-free reinforcement learning-based control strategy
- Incorporating high resolution occupancy data would greatly benefit model-free RL agents
    - they can always outperform baselines if zone-level occupancy information is available
- The facilities manager can navigate the three-way trade-off between energy use, thermal comfort, and visual comfort by tweaking the reward parameters
    - there is no one-size-fits-all approach

ardakanian@ualberta.ca