MTHE 493 - Engineering Mathematics Project

List of Projects 2019-2020

Contents

A Projects in Control

Supervisors:

- Prof. Andrew Lewis (lewisa@queensu.ca)
- Prof. Serdar Yüksel (yuksel@queensu.ca)
- Prof. Abdol-Reza Mansouri (mansouri@queensu.ca)

A-1 Modelling and control of Ripstick® (Lewis)

A Ripstik® is a skateboard-like toy that was popular a few years ago. It is propelled by a rocking motion of one's feet and motion is achieved with no contact of the rider with the ground. The mechanism by which motion is achieved is not well-understood, at least not by me.

The first goal of the project is to come up with a working model of the Ripstik®. This model will necessarily include modelling the salient features of the human rider. Once a working model is obtained, it will be required to produce a working simulation of the model. Finally, a control law will be devised to propel the Ripstik® model.

The modelling will involve understanding and applying sophisticated methods from kinematics and dynamics. A significant part of the project will be producing a graphical representation of the Ripstik® for the purposes of evaluating the dynamical and control models.

Social/environmental/economic considerations

As with many of our projects, this is rather open-ended, and left to the students. An example of a direction students can go is this. There are likely significant changes that will arise in the transportation sector, with such things as electric cars, self-guided cars, etc. Some benefit is to be had by being open-minded about mechanisms for propulsion.

A-2 Using vortices for propulsion (Lewis)

As an airfoil moves through its surrounding fluid, moving forward and generating lift, an essential byproduct is vortex shedding. In this project, this physical mechanism will be, in some sense, reverse engineered and vortices will be used to affect motion of an airfoil, rather than resulting as a byproduct of airfoil motion. Comparatively simple ordinary differential equation models will be derived for the wing/vortex system, and vortex properties such as strength and position will be used as control parameters.

In the first part of this project, the modelling assumptions will have to be decided upon, and the associated mathematical model derived. With a model in hand, some experimentation will need to be done to understand what sorts of motion can be imparted by fiddling with vortex parameters. The final stage of the project will involve control design to achieve desired motion of the airfoil.

This is a quite mathematical project, involving a good understanding of differential equations and control theory.

Social/environmental/economic considerations

The prevalence of air travel makes extremely important any understanding of the mechanics of flight. There is great potential here to perform a detailed analysis of the social/environmental/economic impact.

A-3 Implementation of numerical optimal control package (Lewis)

Optimal control problems arise naturally in many applications of control theory. There are well-understood theoretical techniques for analysing these problems. However, it is only in rare cases that the theoretical methods lead directly to understandable solutions to a given problem. Typically, one arrives at a point where some numerical analysis is required to convert the theoretical results into an actual solution to the control problem. In this project, the emphasis is on using the joint mathematical/numerical tools to solve optimal control problems.

The problem will come in two components which will be done in concert. One component is the application of theoretical optimal control methods to concrete problems. These problems should be selected to achieve two things: (1) one problem should be a simple one that might serve to illustrate the dangers of using purely numerical methods to solve optimal control problems; (2) another problem will be a more complicated one illustrating that numerical methods are indispensable. The second component will be figuring out how to use computer packages to solve problems.

This project will involve a combination of using mathematical methods and intelligently using software to solve specific engineering problems in optimal control.

Social/environmental/economic considerations

There are many specific applications students can come up with that illustrate the benefits of coming up with a practical way of solving optimal control problems.

A-4 Pose Control (Lewis)

The control system considered in this project is a triple pendulum system with actuation in the first two joints. A “pose” is a configuration of the system that can be maintained by an application of control forces, i.e., a “controlled equilibrium.” The objective of the project is to design control laws to efficiently move the system from one pose to another. The matter of what “efficient” means is to be decided upon as part of the project objectives, but minimising energy and time are two common cost prescriptions, neither of which, taken by itself, is often completely satisfactory. Part of the project will be to devise an effective means of displaying the behaviour of the system in simulation.

The problem involves fairly simple mechanical modelling, and quite complicated control methods, as the control problem is not a “standard” one.

This is a project that has a lot of latitude. The numbers of three pendula and two actuators, for example, can easily be modified, as can the location of the actuators. Depending on the exact configuration chosen, the problem can be made more or less challenging.

Social/environmental/economic considerations

The fairly obvious direction to explore triple bottom line issues with this project is in terms of factory automation. But other avenues are also possible.

A-5 A simple robot that hops and flips (Lewis)

The system here is a simple three degree-of-freedom robot that can hop in place (so presenting a stabilisation problem) and that, between sufficiently high hops, can do a somersault. Of course, the robot should stabilise itself upon landing. Part of the problem will be modelling a system that operates in different regimes, i.e., in contact and not in contact with the ground. An impressive simulation tool will be an important part of the project.

The first part of the project will be to model the system, and set up a simulation framework. Then the control design can be undertaken.

There will be some nontrivial modelling required for this project, as the impacts with the ground will have to be modelled. The control methods will mesh “standard” stabilisation methods with nonstandard techniques for controlling mechanical systems.

Social/environmental/economic considerations

The idea here might be to make contact with the expanding world of humanoid robots, and the impacts these have, and may have in the future.

A-6 Approximation Methods and Robustness in Stochastic Control and Applications (Yüksel)

Stochastic control deals with control and optimization of dynamical systems in the presence of probabilistic uncertainty. It has a vast range of applications in very diverse areas engineering and applied mathematics. A general controlled Markov model is represented by the dynamics

$$x_{t+1} = f(x_t, u_t, w_t), \qquad t = 1, 2, \ldots,$$

where x_t is the state of the system, u_t is the control action, and w_t is some independent random noise sequence. The goal is to select the actions u_t so that the resulting state sequence has some desirable properties. One typical criterion is the minimization of the so-called expected cost of the system

$$\sum_{t=0}^{\infty} E[\beta^t c(x_t, u_t)],$$

where $\beta \in (0, 1)$ and $c(x, u)$ is a given a cost function. Although the existence and structural properties of optimal control policies have been studied extensively, computing and implementing such policies is generally a challenging problem for systems with "continuous" (uncountable) state or action spaces. One method to compute approximately optimal solutions is to construct an approximate model by quantizing (discretizing) the state and action spaces.

The goal in this project is to study such a problem and to develop an efficient approximation algorithm. There will be two sub-projects:

P1a. Fully Observed Systems. For such systems, the decision maker has access to the state information x_t prior to selecting an action u_t.

P1b. Partially Observed Systems and Approximations for Non-Linear Filtering. For such systems, a decision maker has access only to a measurement variable given by

$$y_t = g(x_t, v_t),$$

for some function g and i.i.d. measurement noise v_t.

The **design** problems associated with the project involve:

- How does one quantize the state space and the action space?
- Is there a trade-off between computational complexity and approximation quality?
- Can one guarantee asymptotic optimality of such models as the approximation gets finer?
- How can one deal with unbounded spaces when quantizing them?

- What iterative search algorithms yield good quantization methods?
- If only partial measurements are available, how could one design practical methods while also ensuring rigorous bounds on the performance loss?

In addition to the above, the project will also involve implementation of algorithms and policies using Matlab.

This project requires that the group is willing to acquire an advanced understanding of probability and stochastic control.

Social/environmental/economic considerations

This project has a wide range of applications that have become commonplace with the emergence of high-dimensional cyber-physical systems and big-data applications; these entail computationally challenging optimization problems. The economics, environmental, and social aspects of the project will be investigated in this context as well as in the specific contexts of the application area.

A-7 Reinforcement Learning Algorithms for Stochastic Control (Centralized or Decentralized Cases) (Yüksel)

In some stochastic control and optimization problems, one does not know the true transition kernel, or the cost function, and may not even know how many other decision makers there may be in a system, and may wish to use the past empirical data to obtain a sufficiently good solution. In some problems, this may be used as a en efficient numerical method to obtain approximately optimal solutions.

Reinforcement learning / stochastic approximation methods are used extensively for such problems. A typical stochastic approximation has the following form:

$$x_{t+1} = x_t + \alpha_t(F(x_t) - x_t + w_t) \tag{1}$$

where w_t is a noise variable. The goal is to arrive at a point x^* which satisfies $x^* = F(x^*)$. Variations of such dynamics form the core of many learning algorithms.

An increasingly important avenue of research is decentralized stochastic control or dynamic team theory, which involves multiple decision makers who strive for a common goal but who have access only to local information. As an example, consider a discrete-time system:

$$x_{t+1} = f(x_t, u_t^1, u_t^2, w_t),$$

with observations at two different control stations given by

$$y_t^1 = g^1(x_t, v_t^1), \qquad y_t^2 = g^2(x_t, v_t^2)$$

Here, v^1, v^2, w are disturbance/noise variables, $u^i, i = 1, 2$ are the control actions and x is the state variable. The controller i is required to use only its local observations while generating control actions at any given time, unlike a centralized system where the controllers can use all of the information available to each controller. The design goal may be to minimize some expected cost function.

If the decision makers have access to the system model and the exact probabilistic characterization of the environment, they may obtain solutions using various analytical methods. However, in many applications, one has to design optimal policies in an environment which is not completely known; for example, an exact probabilistic model may be missing. For such problems, a decision maker may try to utilize the empirical data to arrive at good control policies. In this case, if there is only one decision maker acting in an environment that is stationary (for example a single investor which takes part in an economy with invariant characteristics) then one can apply well-known algorithms which eventually converge to optimal policies. On the other hand, when there are multiple decision makers present in a system, since each of the multiple decision makers take part in learning, the environment seen by a single decision maker is not stationary. Learning for such contexts is challenging, but many physical systems admit such setups. The goal in this project is to design and implement empirical learning methods to compute optimal solutions for such decentralized problems. In particular, a variation of a popular method known as *Q-learning* can be utilized.

We may consider energy systems and the smart grid as our application area. In this context, the Autonomous Demand Response (ADR) problem is on temperature regulation with multiple (possibly a very large number of) units. In such a system, the real-time price of electricity p_t depends on the system load (that is, the total demand); there are two common approaches to determining p_t: In ex-post pricing, p_t is determined at the end of the time stage t; whereas, in ex-ante pricing, p_t is determined at the beginning of the time stage t. An example of ex-post pricing is $p_t = p\left(\sum_i u_t^i + L_t\right)$ where L_t is the (random) total power drawn by all n units during the period t (kW) and p is an increasing price function mapping the total system load to the unit price of energy. An example of ex-ante pricing would be $p_t = p\left(\sum_i u_{t-1}^i + L_{t-1}\right)$. In either case, since the real-time price of electricity depends on the decisions of all units, the units are facing coupled stochastic optimal control problems. Furthermore, in a realistic implementation each unit would have access only to its local information. In particular, a unit would not have access to any information about the other units, and may not even be aware of the presence of the other units.

There will be two sub-projects:

P1a. Centralized Systems. For such systems, there is a single decision maker acting on the system.

P1b. Decentralized Systems. For such systems, there are multiple decision makers who have only local information and who wish to arrive at an optimal solution under a decentralized information structure.

Some **design** problems will be as follows:

- First, one needs to understand the setup with a single decision maker. In this case, under what conditions can one ensure the convergence of good learning methods? Are there any trade-offs between experimentation and the use of suboptimal policies arrived at through experimentation?
- Suppose that the decision makers have access to identical information regarding measurements. Can the decision makers utilize this common information to arrive at optimal policies?
- Suppose that the decision makers only have access to local measurements, but also have access to a common clock. Could this be utilized to arrive at optimal solutions?
- In the decentralized case, are there trade-offs between experimentation lengths, the use of suboptimal policies and the number of decision makers?
- Suppose that the decision makers only have access to local measurements and are completely oblivious to even the presence of other decision makers (and thus a common clock is also missing)? Can convergence to optimal solutions be achieved?
- Suppose that there is a very large number of units participating in the system. Could one obtain a scalable and practical design for such a system? Could an ideal model with infinitely many units suggest approximately optimal solutions that are implementable for a finite, but large, system?

The project will also involve implementation of algorithms and policies using Matlab.

This project requires that the group is willing to acquire an advanced understanding of probability and stochastic control.

Social/environmental/economic considerations

This project has a wide range of applications that have become commonplace with the emergence of cyber-physical systems and wireless communication being embedded in almost every digital system and a physical need to ensure that such systems perform satisfactorily despite the presence of a decentralized information structure. The economics, environmental, and social aspects of the project will be investigated in this context.

A-8 Optimization and Design of Networked Control Systems and Interaction between Control and Information (Yüksel)

The *interaction between information and control* is a phenomenon that arises in every decision and control problem. Any performance-driven controller requires information on the unknowns that affect the operation of the underlying system; on the other hand, the transmission of information over communication channels and the process of shaping the source output and recovering the transmitted signal at the other end can be viewed as control design problems.

Networked control systems refer to decentralized systems in which the components are connected through communication channels. Such systems are becoming common with the advances in communications and digital technology and the connectedness of the physical world. Control systems are thus increasingly becoming decentralized and networked; *decentralized* because the information available at the controllers might be different and imperfect, and *networked* because there may be communication channels between various components of a control system (such as sensors, controllers and actuators). For such systems, a truly interdisciplinary view is required for optimal system design. One problem for such systems is the characterization of the minimum amount of information transfer needed for the stability of a system. Another problem is the construction of optimal encoding, decoding, and control policies under some performance criterion.

As an example, consider a sensor which observes a sequence of signals (such as a camera sensing the location of a submarine and transmitting this in real-time to a distant control station). The sensor, at each time stage, can use only a finite number of bits to encode its information. The control station, in turn, needs to arrive at control actions (to steer the submarine).

Suppose that the following model describes the system:

$$x_{t+1} = f(x_t, u_t, w_t),$$

where, $\{w_t\}$ is an independent noise sequence, x_t is the $\mathbb{X}$-valued state and u_t is the $\mathbb{U}-$valued control action variable at time t. An encoder at any given time has access to $\{x_0, \cdots, x_t; u_0, \cdots, u_{t-1}\}$

and selects a channel input $q_t = \gamma_t^e(x_0, \cdots, x_t; u_0, \cdots, u_{t-1})$ which takes values from a finite set $\mathcal{M} := \{1, 2, \ldots, M\}$, for $0 \leq t \leq T-1$. Call such an encoder policy $\gamma^e = \{\gamma_t^e\}$. A controller, upon receiving the information from the encoders, generates its control at time t: An admissible causal controller policy is a sequence of functions $\gamma^d = \{\gamma_t^d\}$ such that $u_t = \gamma_t^d(q_{[0,t]})$, with $\gamma_t^d : \mathcal{M}^{t+1} \to \mathbb{U}$. We call such encoding and control policies, *causal* or *admissible*. One goal is the computation of

$$\inf_{\gamma^e, \gamma^d} J(\gamma^e, \gamma^d, T),$$

over all admissible encoder and control policies with

$$J(\gamma^e, \gamma^d, T) := \frac{1}{T} E_{\nu_0}^{\gamma^e, \gamma^d}[\sum_{t=0}^{T-1} c(x_t, u_t)],$$

for some $c : \mathbb{X} \times \mathbb{U} \to \mathbb{R}_+$.

Another goal would be to establish conditions that ensure that the system is (stochastically) stable under some coding and control policy. Thus, there may be different sub-projects to be considered.

The goal in this overall project is to obtain optimal and/or efficient and practical coding and control designs. In particular, we will approach the design problem using a control theoretic and information theoretic formulation, and obtain results which are operationally useful in practice. Some **design** problems will be as follows:

- Structural results are very useful because these allow for the reduction of the search space for optimal policies. Can one obtain structural results on optimal encoders and controllers? Given a structural result/constraint, how could one select a good policy based on empirical performances?
- Can the design for finite horizon problems provide suggestions for infinite-horizon (that is, long-run) optimization problems?
- Can one obtain approximately optimal coding schemes through trading performance against practicality, while ensuring that optimality is preserved in the *limit of impracticality* as measured by the memory order in the encoders and the decoders?
- Could one use information theoretic methods to obtain good lower bounds to test the performance of the system?
- For the case when the system is linear driven by Gaussian noise under a quadratic cost criterion, could one obtain more explicit results?

There will be research and implementation components for this project. Background information and further references will be provided. The project will also involve implementation of algorithms and policies using Matlab.

This project requires that the group is willing to acquire an advanced understanding of stochastic control theory and information theory.

Social/environmental/economic considerations

This project has many applications and implications in the context of networked control systems. The economics, environmental, and social aspects of the project will be investigated.

A-9 Prosthetics Design Using Stochastic Control and Information Theory (Yüksel)

The project involves the development of methods to design an optimal brain-neural system-artificial limb interface system that controls a robotic prosthetic arm. The aim is to design an artificial limb that can approach the performance of a healthy human arm with the highest possible degree of accuracy given physical limitations.

Current limitations in literature and research towards such applications suggest that there is no mathematical consensus on a model of the brain as a communication channel and there exist open problems involving the control theoretic framework that needs to adapted given the decentralization in information in the brain and in the artificial limb.

This problem can be viewed as a specific, but very important and practically consequential, application of the problem posed in Problem 1 above since the questions involve a control theoretic formulation under information constraints with some critical differences in the formulation. **There is a very heavy mathematical, i.e., control theoretic and/or information theoretic, component for the project and students who wish to undertake this project should be aware of this nature**.

Some **design** problems will be as follows:

- How can one design such a system given that there are fundamental open problem with respect to our understanding on the human brain? Can one design a system which treats the brain as a source of commands and which has unlimited computational resources?
- How does one map tasks into signals transmitted over the neural channel transmission? Can a control theoretic formulation lead to a map between tasks and actions to be executed?
- How does the artificial limb interprets the received signals?

- Should one consider a continuum of messages under an estimation theoretic framework, or a finite, discrete number of messages per time stage under a probability of error framework?
- Structural results are very useful because these allow for the reduction of the search space for optimal policies. Can one obtain structural results on optimal encoders and controllers? Given a structural result/constraint, how could one select a good policy based on empirical data consisting of a sequence of experimental outcomes?
- Can the design for finite horizon problems provide suggestions for infinite-horizon (that is, long-run) optimization problems?
- Could one use information theoretic methods to obtain good lower bounds to test the performance of the system?
- For the case when the system is linear driven by Gaussian noise under a quadratic cost criterion, could one obtain more explicit results?

There will be research and implementation components for this project. This project requires that the group is willing to acquire an advanced understanding of control theory and, in part, information theory.

During this project, we will regularly interact with Prof. Jon Sensinger of UNB, who is a medical researcher as well as an engineering professor, with extensive experience working with patients.

Background information and further references will be provided. The project will also involve implementation of algorithms and policies using Matlab.

Social/environmental/economic considerations

This project, naturally, has many applications and implications affecting our society profoundly.

A-10 Optimal control of the heat equation in a non-homogeneous medium (Mansouri)

Many real-world systems of interest are governed by partial differential equations instead of ordinary differential equations. As a concrete example, Consider a (1-dimensional) metal rod of length L with one end (at $x = 0$) thermally insulated from the environment and suppose that by heating or cooling the other end (at $x = L$), you can control how much heat you add at that end, this being your only means of influencing the system, i.e. your only control. The governing equation for this system is given by the partial differential equation

$$\begin{aligned}
\frac{\partial T}{\partial t}(x,t) &= \frac{\partial}{\partial x}(c(x)\frac{\partial T}{\partial x})(x,t), \quad 0 < x < L,\ t > 0, \\
\frac{\partial T}{\partial x}(0,t) &= 0, \quad t \geq 0, \\
T(L,t) &= u(t), \quad t \geq 0, \\
T(x,0) &= f(x), \quad 0 \leq x \leq L,
\end{aligned}$$

where $T(x,t)$ is the temperature of the rod at position x and time t, c is the (spatially varying) coefficient of thermal diffusion, u is the control function, and f is the initial temperature profile of the rod. Note that the state of the system at time t is the function $x \mapsto T(x,t)$, i.e. an element of an *infinite-dimensional* real vector space. The problem we wish to consider is that of finding a control function $u\,[0,\tau] \to \mathbb{R}$ which would steer the temperature function T from f at time $t = 0$ to some other temperature profile g at time $t = \tau$ while minimizing a cost function such as

$$\eta(u) = \int_0^\tau u^2(t)dt.$$

The main goal of this project is to approach this problem using tools from finite-dimensional linear control theory by "approximating" the above system by linear control systems of the form

$$\dot{\mathbf{x}}_h = \mathbf{A}_h\mathbf{x}_h + \mathbf{B}_h u_h,$$

with the approximation getting "finer" as $h \to 0$. The key questions then are whether the optimal controls u_h thus obtained do converge (in some sense) as $h \to 0$, and, if they do, whether their limit yields an optimal control for the original system. The specific goals of this project are meant to address these questions by

(i) understanding the mathematics of this problem (control of partial differential equations),

(ii) designing an optimal controller for this problem by adapting tools from finite-dimensional linear control theory,

(iii) performing computer simulations to verify the theoretical observations and validate the design.

Social/environmental/economic considerations

There are numerous application domains for this project, as most real world systems are governed by partial differential equations. In the specific case of the heat equation, the main application domain is process control; The economics, environmental, and social aspects of this design project will be studied in the context of this particular application domain.

B Projects in Computer Vision and Image Processing

Supervisors:

- Prof. Abdol-Reza Mansouri (mansouri@queensu.ca)

B-1 Stochastic Image Models, Stochastic Gradient Descent, and Applications to Image Restoration (Mansouri)

The stochastic approach to image modeling (and, more generally, signal modeling) consists in modeling an image (or signal) as a family of random variables indexed by the vertices of an undirected graph; the graph structure defines the neighborhood relation between these random variables, and the dependencies between these random variables are modelled by conditional probabilities. For a certain class of probabilistic image models known as Markov/Gibbs Random Fields, these conditional probabilities take a relatively simple form and lead to tractable algorithms. As an illustration, suppose given a particular noisy image O, say a page of text rendered unreadable through noise or a noisy satellite image: We wish to remove the noise and restore the image by finding the "most likely" uncorrupted image that gave rise to that noisy observation O. Modeling an observed image as a family of random variables $\mathbf{O}$ and similarly modeling an uncorrupted image as a family of random variables $\mathbf{I}$, and modeling their relation by an image formation process of the form:

$$\mathbf{O} = f(\mathbf{I}, \eta),$$

with η a family of random variables modeling noise, and the function f modeling possible deterministic image transformations (such as blur), the mathematical problem becomes that of finding I which maximizes the "posterior" conditional probability $P(\mathbf{I} = I | \mathbf{O} = O)$, which in turn is equivalent to finding I which maximizes the product

$$P(\mathbf{O} = O | \mathbf{I} = I) P(\mathbf{I} = I).$$

The probability function $P(\mathbf{O} = O | \mathbf{I} = I)$ is determined by the image formation model above, whereas the probability function $P(\mathbf{I} = I)$ (also called image prior) depends on the class of images one is considering (e.g. satellite images versus images of urban scenery). The problem is now two-fold:

- How to "learn" the prior probability distribution $P(\mathbf{I} = I)$ for a given class of images ?
- Once the prior probability distribution has been established, how to find the global maximizer of the posterior probability $P(\mathbf{I} = I | \mathbf{O} = O)$?

Note that these problems are far from trivial since the space of images one considers is typically huge and hence brute-force search is not feasible (even restricting ourselves to binary images of size 64×64, the set of all possible values that the random variables $\mathbf{I}$ and $\mathbf{O}$ can take – i.e. the set of all images – has cardinality 2^{4096}). It happens however that, under the assumption that $\mathbf{I}$ is a Markov Random Field and with a simplified image formation model, both of these problems can be approached using simple but powerful tools from Markov Chain theory.

The main goal of this project is to use this probabilistic approach in order to design an algorithm for restoring noisy or degraded images/signals. More specifically, the key objectives are to:

(i) Learn the required mathematical tools (Homogeneous and Non-Homogeneous Markov Chains, Markov Random Fields, Gibbs Sampling, Stochastic Gradient Descent) that form the foundations of this project,

(ii) design an algorithm for learning the prior probability distribution for a given class of images and applying it to image restoration,

(iii) validate the designed algorithm through computer implementation.

Social/environmental/economic considerations

There are numerous application domains for this project. In past years, applications such as the restoration of noisy images, the restoration of degraded music recordings, and the generation of images under different painting styles, have been adopted, and new application ideas can certainly be found. The economics, environmental, and social aspects of this design project will be studied in the context of the chosen application.

B-2 Region Tracking in an Image Sequence (Mansouri)

Suppose given an image sequence with thousands of frames, and suppose a region of interest is defined in the first frame of the sequence (say a car); the goal is to have the computer "track" this region as accurately as possible in the remaining frames. This is the problem of automated region tracking, and it has numerous applications, e.g. in biomedical imaging, where the heart is tracked from frame to frame to analyze its behaviour, or in traffic monitoring, where individual cars are tracked to identify possible traffic violations (just Google "region tracking" to see the diversity of applications). The variational approach consists in formulating this problem as a minimization problem over a suitable space. More precisely, letting $I_0, I_1 : \Omega \to \mathbb{R}$ be two images, and letting $R_0 \subset \Omega$ be the subset of the image domain Ω corresponding to the region of interest in image I_0, we define a function (also called functional) $E_{(R_0,I_0,I_1)} : \mathcal{C} \to \mathbb{R}$, where

$\mathcal{C}$ is a suitable subset of the set of all subsets of Ω (i.e. the set of all candidate regions), and we declare the global minimizer of this functional to be the desired region $R_1 \in \mathcal{C}$. The problem of "tracking" region R_0 from frame I_0 to frame I_1 is then reformulated as the problem of minimizing the functional $E_{(R_0,I_0,I_1)}$ over the set $\mathcal{C}$. Clearly, for this approach to be of any use, this functional must embed some information on the image characteristics and the regions' invariants. Once the mathematical issues of existence and uniqueness of a minimizer for the functional $E_{(R_0,I_0,I_1)}$ are resolved, the search for such a minimizer can proceed using tools from the calculus of variations, leading to partial differential equations which are then solved numerically.

The objective of this project is to design and implement a region tracking algorithm by:

(i) Learning the required mathematical tools (Calculus of Variations and Partial Differential Equations),

(ii) designing a region tracking functional and deriving the partial differential equations leading to a minimizer using the calculus of variations,

(iii) validating the tracking functional through computer implementation and evaluation of tracking performance on various image sequences.

Social/environmental/economic considerations

There are numerous application domains for this project. In past years, applications such as movie editing (e.g. retouching, inpainting, etc.), automated surveillance, medical imaging, and hockey analytics have been considered. The economics, environmental, and social aspects of this design project will be studied in the context of the chosen application.

B-3 Automated Face Recognition (Mansouri)

Face recognition by computer has seen numerous applications in recent years. An N by N gray scale image can be considered as a mapping $\mathbf{x} : \{1, 2, \cdots, N\}^2 \to \mathbb{R}$, and, equivalently, as a vector $\mathbf{x} \in \mathbb{R}^{N^2}$. Note that for $N = 256$ (a typical value for a small image) $\mathbb{R}^{N^2}$ is already a huge-dimensional real vector space. Assume now given a set

$$\mathcal{C} = \{(\mathbf{x}_1, l_1), (\mathbf{x}_2, l_2), \cdots, (\mathbf{x}_K, l_K)\}$$

of images of faces of various people (taken possibly under various lighting conditions or with various facial expressions), where for each $k \in \{1, 2, \cdots, K\}$, l_k denotes the label (i.e. identity) attached to image $\mathbf{x}_k$. Suppose now a new face image $\mathbf{z} \in \mathbb{R}^{N^2}$ is given; which individual does it correspond to ? Most approaches to face recognition assume that face images cluster around a low-dimensional linear subspace of $\mathbb{R}^{N^2}$ and that the non-essential features (e.g. lighting condition, facial expression) are in the image component orthogonal to that subspace. Face recognition can then proceed by dimensionality reduction and subspace projection. Although such face recognition algorithms have shown very good performance, the linear subspace assumption is clearly an idealization, and there is no reason for face images to cluster in such a linear way.

The main objective of this project is to design a face recognition algorithm that relaxes the linear subspace assumption; more specifically, the specific objectives are to:

(i) Understand the required mathematical tools for solving this problem (Principal Component Analysis, Fisher Discriminant),

(ii) design an algorithm for automated face recognition by formulating it as a problem of dimensionality reduction via a general (not necessarily linear) submanifold of $\mathbb{R}^{N^2}$,

(iii) validate the designed algorithm through computer implementation.

Social/environmental/economic considerations

The application domain for this project is automated surveillance. The economics, environmental, and social aspects of this design project will be studied in the context of this particular application.

C Projects in Information Theory and Communications

Supervisors:

- Prof. Fady Alajaji (fa@queensu.ca)

C-1 Information Bottleneck and Machine Learning Applications (Alajaji)

The *information bottleneck* (IB) method is an information-theoretic technique that optimizes the tradeoff between the compression length of a potentially *complex* (high-entropy) signal and the accuracy of predicting from the compressed signal a *relevant* (low-entropy) signal that is correlated to the original (high-entropy) one. More specifically, consider two correlated random variables X and Y with joint distribution P_{XY}, where X and Y play the role of the complex and relevant signals, respectively. The IB problem consists of determining a (stochastic) mapping $P_{T|X}$, where T is the compressed representation of X, such that $T \to X \to Y$ form a Markov chain in that order and $I(T;X)$ is minimized (to optimize the lossy compression of X) while maximizing $I(T;Y)$ (to ensure that the relevant information Y about X is preserved and captured by T), where $I(\cdot;\cdot)$ stands for mutual information. The IB compression-relevance problem is then to solve $\min_{P_{T|X}:T\to X\to Y} i_{\mathrm{IB}}(P_{XY},\beta)$, where $i_{\mathrm{IB}}(P_{XY},\beta) := I(T;X) - \beta I(T;Y)$ is a Lagrangian functional and β is a tradeoff parameter that balances between the importance of compression and preserving relevance (e.g., when $\beta = 0$, only compression counts and the minimum is achieved by mapping all X values to a single T value; also, when β is large, relevance is paramount and the compression becomes highly detailed by choosing $T = X$). Note that by the data processing theorem, $I(T;Y) \leq I(T;X)$; thus maximizing $I(T;Y)$ while simultaneously minimizing $I(T;X)$ generates a *bottleneck effect*: the relevant information Y is squeezed through the bottleneck caused by the compression of X. The IB principle provides an interesting information-theoretic connection with machine learning including deep neural networks and data privacy systems. The objective of this project is to identify a pertinent application of the IB principle in machine learning problems (e.g., geometric clustering, quantization for noiseless/noisy channels, privacy-preserving algorithms, classification, natural language processing, etc) design, analyze and implement the IB-centric system and systematically compare its performance and complexity with standard methods in the literature.

Students should possess basic knowledge of information theory, optimization techniques and good programming skills.

Social/environmental/economic considerations

This project has pertinent applications in machine learning and deep neural networks. The economics, environmental, and social aspects of the project will be investigated in the context of these applications.

C-2 Problems in Network Epidemics (Alajaji)

This project concerns the study of how epidemics affect the dynamics and behaviour of networks. Here an epidemic can represent the spread of disease through a population, the propagation of noise in communication networks, computer viruses or malware in smartphones or on the web, or the diffusion of an innovation, brand, rumour or idea or the dynamics of competing opinions in a social network. The project will build on prior recent work, where a mathematical network epidemics model based on a generalized Polya contagion urn scheme was developed and analyzed in terms of its stochastic properties, asymptotic evolution and curing strategies.

Two problems are proposed in this project:

(a) *Capturing real-world epidemics:* Study the model's effectiveness on real-world data, such as digital malware datasets, or biological or social networks data. This includes the design of techniques for judiciously fitting the datasets via the Polya network contagion model or its approximate models, developing appropriate metrics for measuring modelling accuracy and perform systematic performance comparisons with other epidemics networks in the literature, such as the susceptible-infected-susceptible (SIS) infection model.

(b) *An infection-curing game:* Study and analyze the model from a game theoretic viewpoint. Here, consider that two agents are competing against each other in a zero-sum game, where one agent is spreading infection in the networks (e.g., via malware attacks in digital networks) and another agent attempts to design curing strategies to control and mitigate the epidemic. One question to investigate is under what conditions on the Polya network contagion model parameters, on the budget constraints for each agent and one the network topography, does there exist a Nash equilibrium on the agents' policies. The design, implementation and evaluation of the agents' strategies in a simulated network will also be conducted.

Students should possess basic knowledge of probability theory, optimization techniques and good programming skills.

Social/environmental/economic considerations

This project has pertinent applications to the design and management of any epidemic-prone data network, such as social, computer, communication or biological networks. The economics, environmental, and social aspects of the project will be investigated in the context of these applications.

C-3 Bidirectional Communication (Alajaji)

This project concerns the efficient and reliable transmission of data over *two-way* (or *bidirectional*) communication channels, where two users exchange messages simultaneously over a common channel. This allows each encoder to *interactively* adapt the current input to its own message and all previously received signals, hence rendering it more resilient to channel distortions. The objective is to design, analyze and implement a high-performing joint source-channel coding/decoding system that competes well with and preferably outperforms coding systems that ignore interactive coding and/or employ separate source and channel codes. The communication channel will be described via a Markov channel model with time-correlated error and erasure bursts which represents well the behavior of wireless channels. The project's tasks include the following:

- Characterize and quantify the amount of statistical redundancy inherent in discrete (or quantized) information sources.
- Considering a low-complexity *baseline joint source-channel coding* system that sends the source uncoded over the channel, design an optimal receiver for the system that judiciously exploits both the source's redundancy and channel memory.
- Design an improved *interactive joint source-channel* coding system that allows each user to adapt their inputs to previously received signals to increase their robustness against transmission errors and erasures.
- Construct competing traditional *tandem source-channel coding* systems based on separately designed source and channel codes (with and without user interaction).
- Implement the different systems for a wide variety of data sources and channel conditions, and compare them in terms of both performance and complexity.

Students should possess basic knowledge of probability theory, information theory and good programming skills.

Social/environmental/economic considerations

This project has pertinent applications to the design of effective wireless communication devices. The economics, environmental, and social aspects of the project will be investigated in the context of these applications.

C-4 Generative Adversarial Networks for Deep Learning (Alajaji)

A generative adversarial network (GAN) is a powerful learning technique that mimics a given data distribution. It consists of pitting two neural networks, one generative and one discriminative, against each other. The generative network produces from random noise candidates of the data that is to be imitated, while the discriminative network evaluates them by distinguishing realizations produced by the generative network from the real training samples. The two networks play a minimax game for which a unique solution exists given an unconstrained discriminative function: the minimax game reduces to minimizing the Jensen-Shannon divergence between the generative distribution and the true data distribution, with the optimal generative network recovering the true data distribution. More recently, an information-theoretic extension to GAN, information maximizing GAN (InfoGAN), was introduced by adding a mutual information regularization term in order to learn disentangled representations of the data in a completely unsupervised manner.

The objective of this project is to design, analyze and implement a GAN or InfoGAN system for a specific machine learning application (e.g., image processing, classification and representation, data privacy, etc).

Students should possess basic knowledge of information theory, optimization techniques and good programming skills.

Social/environmental/economic considerations

This project has pertinent applications in machine learning and deep neural networks. The economics, environmental, and social aspects of the project will be investigated in the context of these applications.

C-5 A Comparative Study of Network Contagion Models (Alajaji)

This project concerns the study of how epidemics affect the dynamics and behavior of networks. Here an epidemic can represent the spread of disease through a population, the propagation of noise in communication networks, computer viruses or malware in smartphones or on the web, or the diffusion of an innovation, brand, rumor or idea or the dynamics of competing opinions in a social network. The project will build on prior recent work, where an infinite-memory mathematical network epidemics model based on a generalized Polya contagion urn scheme was developed and analyzed in terms of its stochastic properties, asymptotic evolution and curing strategies. It will focus on the investigation of the *finite-memory* version of the network Polya contagion model and its application to real-world epidemics. In particular, its asymptotic behavior will be analyzed and compared with those of the well-known susceptible-infected-susceptible (SIS) infection model. Furthermore, the model's effectiveness on real-world data will be examined, such as digital malware datasets, or biological or social networks data. This includes the design of techniques for judiciously fitting the datasets via the Polya network contagion model or its approximate models, developing appropriate metrics for measuring modeling accuracy and perform systematic performance comparisons with the SIS model.

Social/environmental/economic considerations

This project has pertinent applications to the design and management of any epidemic-prone data network, such as social, computer, communication or biological networks. The economics, environmental, and social aspects of the project will be investigated in the context of these applications.