אילב האגודה הישראלית לחקר בצועים
OPERATIONS RESEARCH SOCIETY OF ISRAEL

ORSIS Conference – June 2010
Abstracts Book

Naor Lecture : Dick van Hertog

Two examples of *Optimization at Work*

Optimization plays a crucial role in many important applications in the Netherlands. In this talk I would like to present two of such applications. The first project is on optimization of dike heights. Protection against flooding is an important issue in the Netherlands, since 75% of this country is submersible.In 1953 a big flooding disaster occurred in the south-western parts of the Netherlands, in which 2.000 people were killed, and the economic damage was enormous. Recently the second Delta Committee reported that 19% of all Dutch dikes do not satisfy the standards. Moreover, this committee argued that the norms (which are stated in the Dutch law) are too low, not only because of unforeseen rises in the sealevel, but also because economic growth was not taken into account in the cost-benefit analysis. The government asked us to develop a new model to determine the right safety standards. We proved that for homogeneous dikes the optimal solution is periodic, and explicit expressions for the optimal updating times and heights can be derived. For the inhomogeneous dikes we developed a MINLP model to find optimal dike heights. Numerical results for several dikes will be presented.
This is joint work with Ruud Brekelmans (Tilburg University, Netherlands), Carel Eijgenraam (CPB, Netherlands), Kees Roos (TU Delft, Netherlands).
The second project I would like to discuss is on high dose rate brachytherapy (HDR-BT). Interstitial HDR-BT is a form of radiotherapy where a high-activity sealed source is inserted through a treatment catheter into a tumour for a short time period. Based on the shape and size of the tumour (planning target volume, PTV), multiple catheters are implanted in the tumour. Each catheter offers a range of positions (dwell positions) where the source can stay for some predefined time (dwell time). The objective is to optimally position the catheters (phase 1) and to determine the optimal dwell times (phase 2) such that the PTV receives a prescribed dose distribution while the surrounding organs (organs at risk, OAR) receive as little dose as possible. This problem is basically a MIP problem. When a catheter is implanted, it is likely to deviate from its intended position. Furthermore, a source cannot be positioned exactly at a dwell position. In this presentation we show that Robust Optimization can be used to find solutions that are robust against these uncertainties. Some preliminary numerical results will be presented.
This is joint work with Bram Gorissen (Tilburg University, Netherlands) and Aswin Hoffmann (Radboud University Nijmegen Medical Centre, Netherlands).

(W1A) Reinforcement Learning

Online learning in Competitive Dynamic Environments

Shie Mannor

We consider decision-making problems in Markov-style decision processes where both the rewards and the transition probabilities vary in an arbitrary (e.g., nonstationary) fashion. We propose online learning algorithms and provide guarantees on their performance evaluated in retrospect. The performance guarantees depend critically on the variability of the dynamics, but hold regardless of arbitrary changes in rewards and transition probabilities. Our basic approach is based on robust dynamic programming and we extend it to the case where reward observation is limited to the actual state-action trajectory. We then present a computationally efficient simulation-based Q-learning style algorithm that requires neither prior knowledge nor estimation of the transition probabilities. We argue that online learning in dynamic and competitive setup can be the foundation of adaptation and learning in autonomous systems, especially when the environment is expected to behave in a non-stationary manner.

Adaptive Bases in Reinforcement Learning

Dotan Di Castro* and Shie Mannor

We consider reinforcement learning algorithms in large state and action spaces. In order to cope with the size of the spaces, a function approximation approach to the state-action value functions is needed. We generalize several classical algorithms (such as Q-learning, Actor-Critic, etc.) to algorithms where the basis of the linear function approximation change dynamically while interacting with the environment. A motivation for such an approach is maximizing the state-action value function fitness to the problem faced, thus obtaining better performance. The algorithms are shown to converge using multiple time scales stochastic approximation. Finally, we discuss how this technique can be applied to a rich family of RL algorithms with linear function approximation.

Adaptive-Resolution Reinforcement Learning

Andrey Bernstein* and Nahum Shimkin

We propose a model-based learning algorithm, the Adaptive-resolution Reinforcement Learning (ARL) algorithm, that aims to solve the online, continuous state space reinforcement learning problem in a deterministic domain. Our goal is to combine adaptive-resolution approximation schemes with efficient exploration in order to obtain polynomial learning rates. The proposed algorithm uses an adaptive approximation of the optimal value function using kernel-based averaging, going from coarse to fine kernel-based representation of the state space, which enables us to use finer resolution in the "important" areas of the state space, and coarser resolution elsewhere. We consider an online learning approach, in which we discover these important areas online, using an uncertainty intervals exploration technique. In addition, we introduce an incremental variant of the ARL (IARL), which is a more practical version of the original algorithm with reduced computational complexity at each stage. Polynomial learning rates in terms of mistake bound (in a PAC framework) are established for these algorithms, under appropriate continuity assumptions.

(W1B) Continuous Optimization 1

A Convex Optimization Approach for a Class of Nonconvex Quadratic Estimation Problems

Marc Teboulle* and Amir Beck

We consider the nonconvex problem (RQ) of minimizing the ratio of two nonconvex quadratic functions over a possibly degenerate ellipsoid. This formulation is motivated by data fitting and estimation problems with data contaminated by noise, and arise in many scientific applications. We first briefly review such class of estimation problems. We then focus on the theoretical and algorithmic analysis for solving (RQ). We prove that under a certain mild assumption on the problem's data, problem (RQ) admits an exact semidefinite convex relaxation. We then study a simple iterative procedure which is proven to converge superlinearly to a global solution of (RQ), and show that the dependency of the number of iterations on the optimality tolerance epsilon, grows as O(sqrt(log(1/epsilon))).

A Moving Balls Approximation Method for a Class of Smooth constrained Minimization Problems
Ron Shefi*, Alfred Auslender and Marc Teboulle

We introduce a new algorithm for a class of smooth constrained minimization problems which is an iterative scheme that generates a sequence of feasible points that approximates the constraints set by a sequence of balls, and is accordingly called the Moving Balls Approximation algorithm (MBA). The computational simplicity of MBA, which uses first order data information, makes it suitable for large scale problems. Theoretical and computational properties of MBA in its primal and dual forms are studied, and convergence and global rate of convergence results are established for nonconvex and convex problems. We then introduce a variant of MBA that includes an active set technique, and is particularly suitable for problems with a large number of constraints. This variant is as simple as MBA, and is proven to preserve the convergence properties of MBA. Initial numerical experiments on quadratically constrained problems will demonstrate the viability and performance of our methods when compared to some existing state-of-the-art optimization methods/software.

Isotonic Recursive Partitioning
Ronny Luss*, Saharon Rosset and Moni Shahar

We present a new algorithm for isotonic regression based on recursively partitioning the space into isotonic regions. This creates a sequence of isotonic models of increasing model complexity that converges to the global isotonic regression solution. Our new algorithm generates a full path of regularized isotonic regression models which are often statistically superior to the global solution. This path offers a highly efficient method for generating the global solution through a sequence of structured subproblems, each equivalent to a network flow problem for which efficient algorithms exist. Success of the new algorithm in both prediction and its improved computational properties are demonstrated through a series of simulated and real examples.

(W1C) Stochastic Models 1

A Model for Network Evolution
Gail Gilboa-Freedman* and Refael Hassin

The computerization of data acquisition in the past decades, has led to the emergence of large data bases on the topology of real world networks. Scientists from mathematics, biology, economics and other sciences, search for commonality in the structure of real world networks. While focusing on the characteristics of a static network, some questions have largely been neglected. For example, what is the right way to describe a network evolution during time units? How to describe a network by identical behavior of its element and still capture the desired properties of real networks, like the non-symmetric degree distribution? We suggest a novel approach to modeling network evolution. In general, *Network evolution model* is based on an initial network-structure; it continues with a description of links propagation; and it ends when certain conditions are satisfied. In our model, the evolution is represented by the dynamic of N identical independent Markov chains, each represents an element of the network. We assume that the chains move between a common set of states. The time a chain spends at a given state is exponential with a known parameter. There are two types of states, L and M. If at any given time two chains are both in the same M-state, these chains meet. Starting with a network of no edges, we add an edge between chains when they meet. We explore the characteristics of the network (such as the clustering coefficient or the diameter) during the evolution. We examine characteristics Vs. t (time) or p (the proportion of connected edges in comparison with a clique). As a first step, we consider 2-state Markov chains with states L and M. We investigate the evolution of such networks under several models. For example, one of the chains is a leader and a link is added when a non-leader chain meets the leader; as above but when a chain meets a leader it also become a leader (spreading a disease). Some of our results: We compute |*T_1*|, the expected time until all *N* non-leaders are acquainted with the leader. We show |*T_1*| is linear as a function of \log(*N*); Despite the simplicity of the model and the weak assumptions, interesting features shows up: the distribution is highly skewed and the network is a small world for a large range of *p* values; the proportion of connected nodes is a power law function of *p*, ex' ; We compute |*T_2*|, the expected time until *N* healthy become sick. Interestingly, |*T_2*| may increase or decrease with *N* in dependency of a single parameter *rho*, which is the

ratio of transition rates in and out of the non-meeting state. Moreover, for most \rho values, |T_2| is a power-law function of N.

A Sex Talk: The matchmaking paradox

Iddo Eliazar* and Igor Sokolov

Medical surveys regarding the number of heterosexual partners per person yield different female and male averages – a result which, from a physical standpoint, is impossible. In this talk we establish a statistical model explaining this "matchmaking paradox". We consider a bipartite graph with N male and N female nodes (N>>1), and B bonds connecting them (B>>1). Each node is associated a random "attractiveness level", and the bonds connect to the nodes randomly –with probabilities which are proportionate to the nodes' attractiveness levels. The population's average bonds-per-nodes B/N is estimated via a sample average calculated from a survey of size n (n>>1). A comprehensive statistical analysis of this model is carried out, asserting that: (i) the sample average well estimates the population average if and only if the attractiveness levels possess a finite mean; (ii) if the attractiveness levels are governed by a "fat-tailed" probability law then the sample average displays wild fluctuations and strong skew – thus providing a statistical explanation to the matchmaking paradox.

Waiting Time in the Sequential Occupancy Problem

Rachel Ravid* and Tamar Gadrich

Occupancy Problems Models are constructed and used to analyze randomized phenomena. The models can be classified as one of three categories, each of which corresponds to a well-known physical system: Maxwell-Boltzmann (MB), Bose-Einstein (BE) or Fermi Dirac (FD) statistics. We studied the case of sequential trials using a BE model in each trial, (i.e.: a group of indistinguishable balls were thrown into distinguishable cells with unlimited capacity). We focused on the waiting time until each cell will be occupied by at least one ball. Decomposition of the waiting time in terms of the number of new fulfilled cells per trial is achieved by the usage of recursive generating probability functions. An application of SQC to the machine's arms that produced the same products gathered in the same collecting zone (products have no mark identifying the machine arms that produced them) is given.

(W2A) Simulation and Stochastic Optimization

Particle Filters, Hidden Markov Models and the Stochastic Enumeration Method for Counting and Combinatorial Optimization

Reuven Rubinstein

We briefly discuss several well established Monte Carlo methods, such as particle filters, hidden Markov models, sequential Monte Carlo with re-sampling and randomized algorithms for counting on complex discrete sets and related combinatorial optimization problems. We show that typically all those methods fail, while counting on the sets of feasible solutions of the integer programming constraints and in particular while counting the number of satisfiable assignments in a SAT model. To overcome this difficulty we present a new method for called stochastic enumeration (SE), a stochastic replica of the naive (computationally intractable) full enumeration method. Also presented is a new approach, called OSLA-SPLIT, a combination of classic splitting with importance sampling. The latter is based on the well-known method, one-step-look-ahead (OSLA) method. Polynomial convergence is proved for both SE and OSLA-SPLIT. We show how to implement SE and OSLA-SPLIT for some well known counting problems, such as self-avoiding walks and satisfiability, prove their convergence, and present simulation studies.

Randomized Algorithms with Splitting

Radislav Vaisman* and Andrey Dolgin

We present a new generic randomized algorithm, called the *splitting* or *cloning* algorithm, for combinatorial optimization, counting and sampling uniformly on complex sets, such as the set defined by the constraints of an integer program. Similar to the original classic randomized algorithms our one uses a sequential sampling plan

to decompose a "difficult" problem into a sequence of "easy" ones. It presents, in fact, a combination of MCMC (Markov Chain Monte Carlo), like the Gibbs sampler, with a specially designed splitting mechanism. The latter runs in parallel multiple Markov chains by making sure that all of them run in steady-state at each iteration. Main results obtained are: 1. We show that the splitting algorithm can be efficiently used not only for optimization and counting but for generating uniform samples on discrete sets. Without introducing a proper splitting mechanism, MCMC fails. Thus, in spite of the common consensus on the classic MCMC as being a universal tool for generating samples on complex sets, we found that it fails to generate points uniformly distributed on discrete sets, such as the set defined by the constraints of an integer program. We provide valid statistical tests supporting the uniformity of generated samples by the splitting method. 2. We show how to incorporate the classic *capture-recapture* method into the splitting algorithm in order to obtain a low variance estimator for the counting quantity representing, say the number of feasible solutions on the set of the constraints of an integer program. 3. We present a combined version of the splitting and cross-entropy (CE) algorithms and provide some complexity results. 4. We present supportive numerical results, while solving quite general counting problems, like counting the number of satisfability assignments in a SAT problem, counting the number of feasible colorings in a graph, calculating the permanent, the number of Hamiltonian cycles, the number of 0-1 tables, and calculating the volume of a polytope, as well as solving integer and combinatorial optimization, like TSP, knapsack and set-covering problems.

Partially Observable Markov Decision Processes for Intruder Recognition

Tal Ben-Zvi* and Jeffrey Nickerson

Many decisions have this nature: Something potentially dangerous is approaching, and someone needs to decide quickly what to do. If the potential threat is harmless, then stopping it wastes resources, or harms the well intentioned. If the potential threat is real, then the closer its proximity, the more damage is done. We consider a situation in which sensors gather information in a noisy environment. That is, we imagine a realistic sensor network that may produce false positives (features that are not present) or false negatives (not recognizing an intruder as such). We assume that the sensor network aims to protect a stationary target (e.g., a large battleship at anchor) and detects signals from approaching objects (e.g., small boats traveling in a harbor). Once an object enters a certain perimeter protection range, it is monitored. Then, if the network detects a sufficient number of signals from the object (for example, through video surveillance), it may classify the object's intention as hostile and act accordingly by informing responders. We formulate the detection problem as a finite-time Partially Observable Markov Decision Process (POMDP). POMDP is considered a generalization of a Stochastic Decision Process (SDP). The model we develop is the primary contribution of this paper. It specifies a way to make decisions in situations of relative uncertainty and time pressure. Here we consider a more realistic environment with two types of object intentions: hostile and benign. Beyond the model, we discuss a way to calculate the optimal policy. We show that the optimal policy is of the Control Limit Threshold (CLT) type: defenders should intercept only if the subjective probability that the object is hostile is larger than the CLT. In the study, we introduce a procedure for computing the CLT value and show some of its characteristics. The secondary contribution of the paper is counter-intuitive: control limits, as a function of distance, are not necessarily monotone. One might expect that as an object approaches the target, the control limit will decrease. Given this expectation, a hostile object that is further away will be judged less dangerous than the same object that is closer. Surprisingly, we find that this is not always the case, and we explain why. These results and their associated analysis are of interest to those involved in the design of systems. There also may be a broader application of this work. Although the system we discuss is designed for physical security and is based on Euclidean distance, our approach may extend to cases where there is a range of topological distances or any other multifaceted metric associated with the increasing danger of threat.

(W2B) Game Theory 1

Efficiency Levels in Sequential Auctions with Dynamic Arrivals

Ella Segev* and Ron Lavi

In an environment with dynamic arrivals of players who wish to purchase only one of multiple identical objects for which they have a private value, we analyze a sequential auction mechanism with an activity rule. If the players play undominated strategies then we are able to bound the efficiency loss compared to an optimal mechanism that maximizes the total welfare. We have no assumptions on the underlying distribution from which the players' arrival times and valuations for the object are drawn. Moreover we have no assumption of a common prior on this distribution.

Strategy Proof Risk Assessment
Yaniv Mordecai* and Yigal Gerchak

Risk is a measure of the undesired effect of uncertainty. In complex and unprecedented settings, it is often difficult and, arguably, undesirable to assess risks based on objective (as opposed to subjective) or statistical data, and therefore subjective probability assessments are elicited from experts, professionals and stakeholders, and used in business analysis and decision making processes. While subjective assessments reinforce objective knowledge and information with experience and professional intuition, they are also susceptible to bias and distortion due to methodological, cognitive and strategic reasons. Cognitive Bias, widely discussed in Management Science and quantitative Behavioral Science literature, is psychological, behavioral or cultural, usually unconscious but partly identifiable and reducible. Strategic Bias is the conscious result of personal cost/benefit considerations, and emerges in formally collaborative but, in reality, not quite cooperative assessment processes. Strategic Bias is recognized and discussed in Social Choice and Game Theoretic frameworks and contexts, but has not been seriously explored in Risk Assessment literature, although the phenomenon clearly exists in such assessments, especially when assessors are also stakeholders. Moreover, current collaborative assessment combination methods, such as Averaging and Bayesian Inference, are vulnerable to this conscious bias, a result of Strategic Behavior. Strategically biased assessments may mislead decision-makers and risk-analysts, and lead to wrong and even dangerous decisions. We study the characteristics of Strategic Behavior in Risk Assessment processes, and provide a Strategy-Proof assessment combination mechanism called MEDAS – Median Distribution Assessment Scheme – which neutralizes Strategic Bias. Our mechanism is based on the "Median Voter Scheme", prominent in Social Choice Theory, and extends it from point estimates to general continuous distributions. The mechanism is useful in collaborative assessment of probability and risk. Complex programs, systems, processes and projects, as in R&D (Research & Development), NPI (New Product Introduction), software development, construction and infrastructure, are typically exposed to high risk, emanating from various sources and affecting several objectives, such as lead-time, cost, technical specification, mission success, business continuity and reputation. Risk is not always unavoidable, as risk handling strategies may be available to reduce risk through additional efforts, or to trade-off risk impacts among relevant objectives. It appears that some risk is both necessary and inevitable when undertaking innovative initiatives and projects for future benefits. We formulate the Risk Optimization Model, which seeks to minimize risk through the selection of risk handling strategies. The goal function, reflecting the total risk, is the total weighted Expected Disutility of Objective-Risk in the system. This formulation consists of Utility Theory. We employ a novel selective convolution approach, which generates convolved probability distributions of risk aggregates, according to the binary selection of the risk handling strategies. This process is termed here Stochastic Convolution-Based Optimization. We provide a pseudo-DP (Dynamic Programming) procedure for a full search for small-sized risk optimization problems. In larger problems we would search for relatively good risk balance solutions using local search heuristics.

Auctions with a random number of bidders
Moshe Haviv* and Igal milchtaich

A random number of bidders participate in a multiple item auction, where all items have a common value for all participants. We examine first-price and all-pay actions, with and without a budget constraint. In all models, the symmetric equilibrium bidding strategy is mixed. We compare the equilibrium strategies for the different models, and present explicit solutions in several special cases, in which the bidders' (identical) beliefs about the number of other bidders are described by standard probability distributions.

(W2C) Scheduling 1

A Fully Polynomial Algorithm for Cyclic Scheduling Problems with Interval Data
Eugene Levner* and Vladimir Kats

We study cyclic scheduling problems having a wide scope of applications in various fields, such as planning of robotic cells and periodic manufacturing processes, parallel computations, periodic inspections, database management, digital signal processing, traffic light scheduling, and others. While most of the previous research has been devoted to solving problems with fixed numerical data, we consider a more general case in which the input data (i.e., operation durations) are known to lie within prescribed intervals. Such type of data makes a scheduling model more flexible. These cyclic scheduling problems constitute a subclass of linear programming problems, and, hence, can be solved in weakly polynomial time. In past decades, a question has been debated

whether these problems, with the input data of any sign, can ever be solved in strongly polynomial time. As the above general question remained open, several (strongly-, weakly-, or pseudo-) polynomial algorithms have been developed in the literature for solving various special cases of those scheduling problems and the associated problem of finding the minimum cycle ratio (see, e.g., Lawler (1967); Karp (1978); Gondran and Minoux (1979); Karp and Orlin (1981); Cohen et al. (1985); Roundy (1992); Hartmann and Orlin (1993); Ioachim and Soumis (1995); Hanen and Munier (1995, 2009); Klundert (1996); Lee and Posner (1997); Chen et al. (1998); Levner and Kats (1998, 2008); and Kampmeyer (2006)). Megiddo (1979) proposed a strongly polynomial parametric method for finding a cycle of the minimum length-to-height ratio in a graph in the case if all the cycle heights are positive. This method can be also employed for solving (in strongly polynomial time) the cyclic scheduling problems with positive numerical data. (Surprisingly enough, as far as we know, this fact has hitherto remained unnoticed by the scheduling researchers). However, Megiddo's method does not close the above general question. Indeed, since it can only solve the problems in which the denominator in the rational objective function is always positive, it cannot handle the scheduling problems with interval data, in which case the cycle heights may be either positive or negative. The main contribution of this paper is a strongly polynomial algorithm for solving the cyclic scheduling problems with interval data. To our knowledge, no strongly polynomial algorithm for these scheduling problems has been known before. The new method is based on Megiddo's idea but our technique differs in that it handles a more general case – when the input data are of any sign. The new graph-based algorithm, combines the idea of Megiddo's method (1979) with the parametric critical path algorithm by Levner and Kats (1998). It runs in O(n2m log n) time, where n is the number of nodes and m the number of arcs in the underlying graph.

Modeling Yield vs. Flow-Time in a Production System

Israel Tirkel* and Gad Rabinowitz

Modern industry struggles with the tradeoff between Yield and Flow-Time (FT) due to their significant impact on profit. Short FT reduces costs and increases revenue, and high Yield increases revenue and decreases cost per unit. This work studies the impact of inspections on Yield and FT in a deteriorating production system. It originates in semiconductors manufacturing, but can be applied in other industries as well. Analytical models are validated by simulation. Numerous factors (e.g. maintenance, inspection) impact both Yield and FT, thus call for integrated analysis. The factors considered here focus on (*i*) Inspection frequency - rate at which product items are inspected, and (*ii*) Inspection scheduling - timing and sequencing at which items are inspected. Simplified three-step *Production Cell* is developed to represent a basic segment of a production-line. It consists of two production steps, *Feeder* and *Bleeder,* and one inspection step, *Metro*, each represented by a single queue. The *Feeder's* state randomly deteriorates from in-control (IC) to out-of-control (OOC), and is corrected back to IC only with external intervention. Following a certain Inspection Policy (IP) an item processed in the *Feeder* is sent to the *Metro*where it is correctly detected with Y_{IC} (Yield of item processed by IC *Feeder*) or with Y_{OOC} (Yield of item processed by OOC *Feeder*), clearly $Y_{IC} > Y_{OOC}$. All items then continue to be processed in the *Bleeder*. Fixed Measure Rate (FMR), static IP most commonly used, is applied with various parameters. Variables impacting Yield and FT are identified as (*i*) Measure Cycle (*MC*) - number of items between consecutive inspections, and (*ii*) Feedback Delay (*FD*) - integer number of items entering the *Feeder* during time of inspection. Two cases in multiple scenarios are investigated: 1. *Fixed FD* – where the Feeder is M/M/1, the Metro is $E_{MC}/E_{FD}/\infty$, and the Bleeder is G/M/1 approximated to M/M/1. 2. *Random FD* – where the Feeder is M/M/1, the Metro is E_{MC}/M/1, and the Bleeder G/M/1 approximated as M/M/1. Results illustrate a typical concave curve where growing inspection rate increases both Yield and FT, indicating tradeoff. Yet, at a certain inspection rate Yield reaches a maximum and at higher rates starts to decline while FT continues to increase, which is undesired. The cause of Yield decline is longer delay of corrective feedback to an OOC production tool, due to longer waiting-time for inspection. This inferior behavior may be overlooked, if not properly recognized. Major conclusions are as follows: 1. Preferred Yield-FT combination business-wise is maximized either at peak Yield or just prior to its significant drop, with almost no improvement in FT. 2.Preferred Yield-FT combination can be selected by pre-determining *MC*; this reflects FMR policy in practice. 3. Although not directly pre-determined, reduced *FD* average and variance improves Yield results; *FD* variance decreases with growing *MC*. 4. Model works well in determining preferred *MC* and is favored due to its ease of use and clarity. This work provides a decision support tool for applying in-line inspection, while simultaneously considering Yield and FT. Future studies could further develop this ignored operations management issue of Yield-FT tradeoff.

A Folding Horizon Approach for Setting Gates of Activities in the Stochastic Project Scheduling Problem

Illana Bendavid* and Boaz Golany

This paper addresses the problem of controlling the scheduling of activities in projects with stochastic activity durations. A first approach is to set a gate for each activity i.e., a time before it the activity cannot begin. Since the resources required for each activity are scheduled to arrive according to its gate, we may incur a "holding" cost when an activity is ready to be processed, but the resources required for it were scheduled to arrive at a later time; or we may incur a "shortage" cost when the required resources have arrived on time but the activity cannot start because its predecessors are not yet finished. Our objective is to set the gates so as to minimize the sum of the expected holding and shortage costs. In this method, all the gates are determined in a ``static" way, at time ``zero", before the project starts and before any uncertain parameters (the durations) are realized. In this way, all the risk induced by the uncertainty is assumed by the project manager (PM) only. Introducing flexibility into contracts with the subcontractors through negotiations can help reducing the uncertainty and sharing the risk between the PM and the subcontractors. A second approach is to first solve the problem at time ``zero" in a static way to obtain a basic guideline for the contract and then, each time more information is obtained, the PM can incorporate it to solve the problem for the remaining activities and to adjust his future decisions in a dynamic way, allowing him to reduce uncertainty, thus to reduce his costs. We employ the Cross-Entropy method to solve both problems. We describe the implementation of the method, compare its results to other heuristic methods and provide some insights towards actual applications.

(W3A) Asymptotic Methods in Queues

On the non-degenerate slowdown diffusion regime

Rami Atar*, Itay Gurvich and Nir Solomon

I will discuss diffusion limit results and asymptotic control formulations for a heavy traffic regime, called the nondegenerate slowdown regime, that is distinct from both the conventional and the Halfin-Whitt regimes, in that the service time and the delay are of the same order of magnitude. For a model with multiple streams of arrivals and multiple server pools, routing policies that minimize delay in conventional heavy traffic have been identified by Bell-Williams and by Mandelbaum-Stolyar, by two different approaches. I will exhibit that some aspects of these approaches to study the model carry over from the conventional to the nondegenerate slowdown regime, and some do not.

Control of Fork-Join Networks in Heavy-Traffic

Asaf Zviran* Rami ATar and Avishai Mandelbaum

This work addresses the problem of analysis and control of fork-join networks in the conventional Heavy-Traffic diffusion regime. Standard fork-join networks are feed-forward, which are relatively easy to control. We thus allow probabilistic feedback, which turns the problem into a challenging gone. In our models, activities are associated uniquely with customers. They are hence non-exchangeable in the sense that one can not join activities associated with different customers – this is the case in healthcare (e.g. emergency departments) and multi-project environments (and in contrast to assembly networks). In this framework we introduce a natural concept of optimality, and then solve for the optimal control, asymptotically in heavy-traffic. The central ingredient in the proof is the establishment of asymptotic equivalence between non-exchangeable and exchangeable dynamics.

Pricing for Social Utility Optimization in Cloud Computing Centers

Nahum Shimkin *, Ishai Menache and Asuman Ozduglar

Motivated by the recent advent of cloud computing facilities that offer online computing power on demand, we consider a large service facility that offers simultaneous service to a large number of heterogeneous jobs, each associated with a distinct user. The execution time of each job depends on the amount of resources applied to it. Our main concern here is in maximizing the social utility, which comprises of the users' service utility minus their delay cost. This requires regulating the distribution of available resources between the active users, as well as controlling the arrival rates of the different job types. In this work, we formulate a fluid queueing model that captures the essential ingredients of this problem. We proceed to characterize the social optimum, and propose a pricing mechanism that uniquely induces that social optimum, without requiring the system to be aware of the users' attributes and preferences. The proposed mechanism divides the available resources between the active

users according to their price bids, and is reminiscent of proportionally fair pricing mechanisms of network flow control.

(W3B) Approximation Algorithms

An Extension of the Nemhauser and Trotter Theorem to Generalized Vertex Cover with applications

Dror Rawitz*, Reuven Bar-Yehuda and Danny Hermelin

The Nemhauser and Trotter Theorem provides an algorithm which is frequently used as a subroutine in approximation algorithms for the classical Vertex Cover problem. In this paper we present an extension of this theorem so it fits a more general variant of Vertex Cover, namely the Generalized Vertex Cover problem, where edges are allowed not to be covered at a certain predetermined penalty. We show that many applications of the original Nemhauser-Trotter Theorem can be applied using our extension to Generalized Vertex Cover. These applications include a (2-2/d)-approximation algorithm for graphs of bounded degree d, a PTAS for planar graphs, a (2-\lg\lg n/2\lg n)-approximation algorithm for general graphs, and a 2k kernel for the parameterized Generalized Vertex Cover problem.

Approximating Functions in Logarithmic Space and Time: A ``Plug & Play'' Approach

Nir Halman* and James Orlin

We consider several natural problems related to the design of approximation algorithms and the analysis of their error bounds. We define a set of sufficient conditions on a function f:D \to R^+ and its domain D, so that we can construct good approximations for it in space, time, and number of queries, which are all polylogarithmic in |D| and \max_{x \in D}f(x). Using our ideas we construct a meta algorithm for obtaining Fully Polynomial Approximation Schemes (FPTASs) for stochastic dynamic programming on several families of directed acyclic graphs. Our results are given in a modular way, as a set of ``ready-made'' algorithms and computational rules, so that future (and past) approximation algorithms will be simplified by using them.

Price of Adaptivity in Stochastic Knapsack Problem with Dependence on State of Nature Assumption

Aleksander Vainer* and Asaf Levin

We consider a stochastic variant of the NP-hard 0/1 knapsack problem in which item values are deterministic and item sizes are random variables with known, arbitrary distributions, that depend on random variable \theta denoting the state of the world, and fixing a value of \theta the size variables are independent. So \theta induces a limited dependency among the sizes of distinct items. Items are placed in the knapsack sequentially, and the act of placing an item in the knapsack instantiates its size. Our goal is to compute a solution ``policy'' that maximizes the expected value of items successfully placed in the knapsack, where the final overflowing item contributes no value. We consider both nonadaptive policies (that designate a-priori a fixed sequence of items to insert) and adaptive policies (that can make dynamic choices based on the instantiated sizes of items placed in the knapsack thus far). An important facet of our work lies in characterizing the benefit of adaptivity. For this purpose we advocate the use of a measure called the adaptivity gap: the ratio of the expected value obtained by an optimal adaptive policy to that obtained by an optimal non-adaptive policy. We suggest to study the price of adaptivity of this model as a function of the size of the support of \theta. Assuming that the support of \theta has at most k values, we show a lower bound of \Omega(k) and upper bound of O(k) in our model of the Stochastic Knapsack problem.

(W3C) Scheduling 2

A Bi-criteria Approach to Scheduling a Single Machine with Rejection and Positional Penalties

Nufar Gasper*,Liron Yedidsion and Dvir Shabtay

Single machine scheduling problems have been extensively studied in the literature under the assumption that all jobs have to be processed. However, in many practical cases, one may wish to reject the processing of some or all of the jobs in the shop. The decision to reject jobs may be due to low machine capacity or high scheduling costs and is assumed to result in a rejection cost. In such a

framework, the scheduler has to decide first which jobs will be rejected and which will be accepted. Then he has to schedule the accepted jobs efficiently. Thereby two criterions are used to evaluate the quality of a schedule: a scheduling criterion, F1, which is dependent on the completion time of the accepted jobs, and the total rejection cost, F2. Problems of scheduling with rejection have been previously studied, but usually in a narrow framework .focusing on one scheduling criterion at a time. This paper provides a robust unified bicriteria analysis of a large set of single machine problems sharing a common property: all problems can be represented or reduced to a scheduling problem with a scheduling criterion which includes positional penalties. Among these scheduling criteria are the minimization of the makespan, the sum of completion times, the sum and variation of completion times, and several due date assignment problems. Four different problem variations for dealing with the two criteria are studied. The variation of minimizing F1+F2 is shown to be solvable in polynomial time, while all other three variations are shown to be NP-hard. For those hard problems, we provide a pseudo polynomial time algorithm. An FPTAS for obtaining an approximate efficient schedule is provided as well. In addition, we present some interesting special cases which are solvable in a polynomial time.

Equal Allocation Policy in Openshop batch Scheduling
Baruch Mor[1]*, Gur Mosheiov and Daniel Oron

We check the optimality of the very practical policy of equal allocation of jobs to batches in batch scheduling problems on an *m*-machine open-shop. The objective is minimum makespan. We assume unit processing time jobs, machine-dependent setup times and batch availability. We show that equal allocation is optimal in a 2-machine and a 3-machine openshop, but is not optimal for openshops containing more machines.

Scheduling a Deteriorating Maintenance Activity on a Single Machine
Gur Mosheiov* and Jeffrey B.Sidney

We study a problem of scheduling a maintenance activity on a single machine. Following several recent papers, the maintenance is assumed to be deteriorating, i.e. delaying the maintenance increases the time required to perform it. The following objective functions are considered: makespan, flowtime, maximum lateness, total earliness, tardiness and due-date cost, and number of tardy jobs. We introduce polynomial time solutions for all these problems.

(W3D) Logistics and Transportation

Service Oriented Train Timetabling
Mor Kaspi* and Tal Raviv

The train planning problem can be divided into several sub-problems, mainly Line Planning, Timetabling, Platforming, Rolling Stock Circulation, and Crew Planning. In this study we deal with the two strategic sub-problems: *Line Planning* – deciding which set of lines should be served by the system and at what frequencies; *Timetabling* – deciding upon the schedule of each train in each line subject to track availability and safety constraints. Our objective is to minimize the total time spent by the passengers in the railway system, including waiting time at the origin stations, connections and travel time. We believe that this is the proper measure for quality of service. We formulated an integrated Line Planning and Timetabling Problem (ILPTP) and devise a Cross Entropy Heuristic to solve it. Evaluating a timetable according to the above measure requires finding the shortest journey for each flow of passengers in terms of total time. We construct a graph representation of a timetable such that the shortest journey of each passenger stream is equivalent to a shortest path on it. We then devise a specialized shortest path algorithm for this graph that is used to quickly evaluate the objective function. Our algorithm runs in linear time with respect to the number of arrival/departure events and to the number of passenger stations. This algorithm is significantly faster than general algorithms for the all-pairs shortest paths problem. We then test our algorithm with data obtained from Israel Railway. The timetable created by our algorithm may save about 20% of the total travel time as compared to the one currently in use, allows better utilization of the infrastructure and is more robust. In addition, a variation of the algorithm is used

to solve a bi-objective problem to explore the tradeoff between operational costs and service level. Solutions that dominate the current timetable with respect to both objectives are generated.

Lateral Transshipments with Fixed Transshipment Costs

Reut Bonshtain* and Michal Tzur

Our study deals with inventory systems in which lateral transshipments are allowed. We consider two-retailers in a multi-item periodic-review inventory system and extend existing models by relaxing the assumption of negligible fixed transshipment costs. The goal is to determine the inventory transshipment and replenishment policies that maximize the expected system profit. We first introduce the single-item transshipment problem with fixed transshipment costs and investigate properties of the cost function as well as those of the optimal solution. Then we extend the model to include multi-items with a joint fixed transshipment cost. We analyze the multi-item problem based on the results obtained in the single-item case.

The Uncapacitated Dial-A-Ride Problem on a Tree

Rona Pfeffer* and Shoshana Anily

The Uncapacitated Dial-a-Ride Problem (UDARP) on a tree is defined by a tree, a set of object types (including a null object), and a single vehicle of unlimited capacity, which is initially positioned at the root of a tree. Each vertex of the tree is associated with a pair of object types: the object type supplied and the object type required by the vertex, each of which may be a null object. The total supply (demand) of each object type, except for the null object, is exactly one unit. In addition, each edge of the tree is associated with a non-negative cost. The objective is to design a minimum cost feasible route that starts and ends at the root of the tree, so that the vehicle, while following the route, loads and unloads objects at the vertices in order to satisfy the requirements of all the vertices. The problem is known to be NP-hard. We prove some structural properties that any optimal solution for the UDARP on a tree satisfies. Lower and upper bounds on the optimal cost are provided, as well as necessary and sufficient conditions under which the optimal solution coincides with the lower bound. We present both exact and heuristic algorithms for solving this problem. In particular we provide a mixed-integer-programming formulation of the problem, which is based on the formulation of the Traveling Salesman Problem with precedence constraints. We present a computational study that was performed using CPLEX. In addition, we developed a dynamic programming formulation that computes the optimal solution for small instances. Finally, we propose an effective heuristic, which is shown via a computational study to give near optimal solutions for small instances of the problem.

(T1A) Constraint Satisfaction Problems

Constraint Satisfaction Problems and Operations Research

Michael Veksler

A Constraint Satisfaction Problem (CSP) defines a feasibility problem over unrestricted constraints. The same, systematic, solving algorithm works equally well for many types and mixtures of constraints. Moreover, a constraint is not restricted by any specific semantics and can be a relation represented by a simple table. The field of systematic CSP solving has been evolving since 1960's and modern algorithms have been evolving since 1970's. In this presentation I will outline basic systematic solving algorithms for CSP, as used by state-of-the-art solvers. Some of current challenges and research opportunities will be mentioned such as: optimization, hybrid methods, symmetry, solution/unsatisfiability analysis and modeling.

Constraint-Based Random Stimuli Generation for Hardware Verification

Michal Rimon*, Yehuda Naveh, Itai Jaeger, Yoav Katz, Michael Vinov, Eitan Marcu and Gil Shurek

Functional verification of hardware design is the process of ensuring the conformance of a logic design to its specification. Hardware designers use Hardware Description Languages (HDL) to describe the physical hardware. This HDL code can be simulated by commercial software tools, and can be automatically synthesized into gate-level circuits. In current industrial practice, simulation-based verification techniques play a major role

in the functional verification process. These techniques verify the design's actual behavior (as obtained from simulating the HDL description), to its expected behavior (as derived from its specification). To do this, they generate stimuli with which to drive the simulation model. Hence, the field of simulation-based verification of hardware design is concerned mainly with the generation of stimuli that most effectively exposes design bugs. In this presentation we will describe the architecture of a family of stimuli generation applications developed at IBM which target various aspects of processor and system level verification. Each of these tools decomposes the stimuli generation problem into components which are viewed as constraint satisfaction problems (CSP). We will discuss the modeling of hardware information and domain expert knowledge in the CSP framework, and describe a general-purpose CSP solver that is specialized for the kind of problems which arise from the stimuli generation domain.

Constraint Programming for Cloud Computing
Odellia Boni

Cloud computing is one of the most imminent trends in data centers. The cloud is comprised of virtual machines (computers) that the users are working on and physical machines where the applications are actually running. One of the major problems in this field is the problem of placing the virtual machines on the physical machines (hosts) such that constraints regarding resources, security and efficiency are satisfied, and various objective functions which correspond to the working mode of the cloud are optimized. The problem was modeled as a CSP problem and solved by a systematic search solver. Results show the solution to be scalable as well as flexible.

(T1B) OR Applications 1

Quadratic Model for Budget Allocation
Zilla Sinuani-Stern

We develop a quadratic model for allocating operational budgets in public and nonprofit organizations. The allocations for each organizational unit have lower and upper bounds. The objective function is to minimize the weighted sum of the quadratic deviations of each allocation from its bounds. The weights reflect the priorities, the benefits, or the size of the various units. The optimal allocations are mostly around the midpoint between the bounds. A simple algorithm is presented to derive the solution. The new quadratic model is compared to the familiar linear model for budget allocation which almost always provides extreme allocations on the bounds: for some units on the upper bound while for others on the lower bound. We perform sensitivity analyses, and resolve special cases of the model with closed form solution. The model with its variants was actually used for allocating budget in various university setups, some examples are presented here

Precision agriculture management using segmentation of massive hyper-spectral satellite image
Z. Mhabary*, S.Cohen Y. Cohen, V. Alchanatis and O. Levi

Precision agriculture is concerned with techniques and practices that maximize crop yield with minimum cost through improved use of data, information, and knowledge technologies. Detailed information about land, weather, water, agriculture markets, prior yields, agricultural-chemical options, seeds, etc. is required for planning crop selection and planting. High temporal and spatial resolution information is required to monitor crops, assess risks, and make decisions about appropriate interventions to maintain crop health. Similarly, to maximize yields, decisions about timing harvests require information about current and future conditions. Remotely sensed information about farm assets including information collected from the farm about outcomes of plans, cultivation techniques, and harvests, is integrated within a farm's database for perpetual use. Remote sensing is an essential source of information for precision agriculture management, and specifically multi and hyper spectral sensing provides data this practice. Although the costs of acquiring hyper spectral images is typically high, for specific crops and in specific climates, hyper spectral remote sensing is used more and more for monitoring the development and health of crops. Therefore, segmentation of Hyper-spectral images requires advanced analysis and computational methodologies. We introduce a new state of the art method for segmentation of multi-spectral images. The proposed methodology is based on a multi-scale geometric transformation called the Beamlet Transform. The method is applicable for both mono-spectral and multi-spectral images where each pixel has its corresponding spectral profile vector. The proposed segmentation method is especially effective when the underlying image consist of relatively large segment with smooth

boundaries. In this case it performs exceptional well even in extremely low SNR. The method is unsupervised and assumes no prior knowledge of the image characteristics or features.
Despite of being relatively complex and sophisticated our proposed segmentation algorithm has a surprisingly low computational complexity of $O(N \log N)$ which is achieved by implicit computations through variants of the FFT's. However, given the typical large volumes of satellite imagery data and the need for fast online computation, standard hardware is not sufficient, therefore we combine the super Fast methods in image processing with parallel implementation on a distributed system based on cluster architecture. In order to validate the efficiency of our method we used the known Lark algorithm as a benchmark for segmentation of multi-spectral images and show that our new method out-performs the Lark algorithm.

Evaluating Israeli Police stations' performance using Data Envelopment Analysis (DEA)

Doron Alper* and Zilla Sinuani-Stern (**ORSIS Prize Winner**)

This study estimates the relative productivity of 60 police stations in Israel during 2006-2007, based on two inputs and 16 outputs, using Data Envelopment Analysis (DEA) model. We utilized several DEA versions: constant and variable return to scale, with and without bounds on the virtual variables. Moreover, benchmark analysis was done. The police stations were ranked via Cross Efficiency analysis, and the Maverick index was derived for each station. The validation of DEA was implied when the various models were highly correlated. Regression analysis did not depict external factors to explain the variability of the efficiencies of the police stations.

(T1C) Allocation problems in competitive environments

Existence and computing a unique Nash equilibrium in a stochastic competitive R&D race where winner takes all

Inbal Mund*, Pelin Gulsah Canbolat, Boaz Golany and Uriel G. Rothblum

We model a market with multiple firms deciding on how much to invest in an R&D project. The investment of each firm, together with a firm-specific constant, determines the rate of the exponential project completion time for that firm. The firm that completes the project first, collects all its revenue. The focus of this talk will be on a simple method for computing a unique Nash equilibrium. The sensitivity of the Nash equilibrium to various parameters of the model will be discussed.

A stochastic competitive R&D race where ``winner takes all''

Pelin Canbolat*, Boaz Golany, Inbal Mund and Uriel G. Rothblum

We model a market with multiple firms deciding on how much to invest in an R&D project. The investment of each firm, together with a firm-specific constant, determines the rate of the exponential project completion time for that firm. The firm that completes the project first, collects all its revenue. The focus of this talk will be on comparing the Nash equilibrium (which turns out to be unique and easily computable in this case) and the globally optimal solution(s) maximizing the aggregate utility of all firms. The sensitivity of the market activity to various parameters of the model and the loss due to decentralization in this particular setting will be discussed.

Characterization and computation of Nash equilibria in defensive resource allocation games

Noam Goldberg*, Boaz Golany and Uriel G. Rothblum Rothblum

We consider a two player game in which a defender seeks to allocate a scarce resource to multiple defensive activities, or potential targets which can be protected; the probability of an attack being successful decreases in the amount of resource allocated to the given target. We build on previous work of Golany, Kaplan, Marmur, and Rothblum (2009), proposing a closed form solution for the attacker's strategy, and an algorithm for finding the defender's strategy, in Nash equilibria of a zero-sum resource allocation game. We extend their results to the case where the defender has multiple resources at their disposal and the set of defender strategies corresponds to a more general polyhedral set. In addition, in the case of a single resource being allocated, we consider an

extension of the results of Golany *et al.* in order to compute Nash equilibria of a nonzero-sum resource allocation game.

Determining all Nash Equlibria in a (Bi-Linear) Inspector Game
Yael Deutsch*, Uriel G. Rothblum abd Boaz Golany

An important role that governments play in most countries is to inspect whether businesses, organizations and individuals follow the rules and regulations that are set by the legislative bodies in these countries. In most developed countries, individuals and businesses are free to run their lives or operate their businesses as they wish and accountability is based on the government's ability to inspect the compliance of any individual, business or organization with these regulations. Examples of these settings include: the Internal Revenue Service (IRS) agency selects at random some tax forms filled by individuals and checks for possible fraud, the Security and Exchange Commission (SEC) checks for possible inside information incidents in trade transactions, the International Atomic Energy Agency (IAEA) inspects compliance by member states that have signed the Non Proliferation Treaty (NPT) and more. Attempts to model inspection decisions using a game theory approach date back to the early 1960's. The original model considered a two-person non-constant-sum game between an inspector and an inspectee that has motive to violate some rules without being caught. If the inspectee violates the rule and is caught, it has to pay a penalty. The inspectee can avoid paying the penalty by paying some side-payments. This model has resulted in many extensions, especially during the last two decades, where there has been a growing public concern about the compliance of countries with the rules and regulations of the NPT. Some of these researches prove (or demonstrate) the existence of Nash Equilibria solutions for a game between the Inspector and the inspectees. However, with the exception of some two-person game, none of these researches give and prove closed-form expressions for these equlibria. In our paper we formulate a game theory model which is useful for many of the real-world single inspector multiple inspectees scenarios, of which some we discussed above. Our model has one inspector and multiple inspectees which act independently of each other. There is a unique payoff/penalty for each player for complying or violating the rules and/or executing inspections, noting the inspector has a finite budget for computing the inspections. We find all Nash equilibria solutions for this game and provide explicit closed-form expressions for it. These results can be very useful for inspecting agencies that wish to maximize the effectiveness of their inspections while having a limited budget for executing these inspections. We also discuss and demonstrate the sub-optimality of the Nash Equilibria we found, as they depend on the sign of the payoffs/penalties for the inspector and for the inspectees for each action. There are many options for further research on this topic, for example: changing the model such that some (or all) of the inspectees act together and not independently, extending our single-stage model to a multiple-stage model, which is more accurate for real-world inspection scenarios. Other applications can be easily imagined.

Plenary 2 : Mordecai Avriel

Financial Engineering at Bank Hapoalim

The central unit of financial engineering in Bank Hapoalim is the Department of Analytic Development. I will describe the current and past activities of this department that will, hopefully, shed some light on FE in a financial institution.

(T2A) Mehraz Prizes winner's

Optimal Influenza Vaccination Policy
Dan Yamin* and Arieh Gavious

Influenza (flu) is a contagious respiratory illness caused by influenza virus. It can cause mild to severe illness, and may lead to death. The most efficient way to prevent the disease is through vaccination.Vaccinations are vital for reducing the probability of infection not only for the individual who becomes inoculated, but also for the entire population. Although a fairly efficient vaccine for the disease exists, influenza vaccination coverage in all age groups remains suboptimal. This research analyzes seasonal influenza vaccination efficiency based on self-interests versus social interests. The research has two main objectives. The first is to offer a normative explanation for the small percentage of people taking the vaccine. The second is to offer tools for an optimal vaccination policy based on epidemic theory and economic mechanisms with the goal of benefiting society. We have considered a non-atomic population game to model the decision-making problem of an individual considering whether or not to get the vaccine. The game includes a 'social planner' (e.g., a governmental body) interested in minimizing the overall loss caused to individuals by the vaccination policy. The 'social planner' may encourage individuals to take the vaccine by giving them a financial incentive. The model was formulated first for a homogenous population in which all individuals are identical. Then, we considered a heterogeneous population with two age groups. The probability of becoming infected if a proportion p of the population does become vaccinated is determined according to epidemic theory. We look for symmetric equilibrium where all individuals will become vaccinated with the same probability p when the social planner pays q to those who choose to become vaccinated. Under assumptions of rationality, we find the optimal social symmetric incentive generated by a sub-game perfect Nash. From the model we determined analytically that with no incentives, the proportion of those taking the vaccine is lower than the optimal proportion for society. Therefore, for the welfare of society, an incentive mechanism is necessary. Simulations analysis of the results found that the marginal contribution of the incentives to the proportion of vaccinators is greater when the virus type is considered less contagious and less hazardous. Paradoxically, the results showed that the optimal incentive for the population should be greater for seasonal influenza rather than for pandemic types of influenza such as the swine flu. In the case of a heterogeneous population, in order to maximize the benefit to society, the social planner should favor giving incentives to the non-elderly population over the elderly. Results also showed that preference should be given to inoculating the infant population over the rest of the population. These findings are at odds with the CDC's policy of giving incentives to populations at risk rather than to those who are not at risk.

A Strongly Polynomial Algorithm for Controlled Queues
Alexander Zadorojniy*, Guy Even and Adam Shwartz

We consider the problem of computing optimal policies of finite-state finite-action Markov decision processes (MDPs). A reduction to a continuum of constrained MDPs (CMDPs) is presented such that the optimal policies for these CMDPs constitute a path in a graph defined over the deterministic policies. This path contains, in particular, an optimal policy of the original MDP. We present an algorithm based on this new approach that finds this path, and thus an optimal policy. In the general case, this path might be exponentially long in the number of states and actions. We prove that the length of this path is polynomial if the MDP satisfies a coupling property. Thus we obtain a strongly polynomial algorithm for MDPs that satisfies the coupling property. We prove that discrete time versions of controlled *M/M/*1 queues induce MDPs that satisfy the coupling property. The only previously known polynomial algorithm for controlled *M/M/*1 queues in the expected average cost model is based on linear programming (and is not known to be strongly polynomial). Our algorithm works both for the discounted and expected average cost models, and the running time does not depend on the discount factor.

(T2B) Decision making

Does Competition Improve Performance
Yigal Gerchak* and D. Marc Kilgour

Agents, such as salespersons, are often required to compete, in the sense that their rewards are determined, at least in part, by relative rather than absolute performance. It is well-established that contests motivate rational agents to increase their level of effort. We explore the extent to which contests also change strategy; when agents choose not level of effort but, say, level of difficulty, is rational behavior different when ordinal measures of success become more important than cardinal measures? Basing our conclusions on some simple probabilistic

models where optimal choices can be determined analytically, we show that the structure of the competition is important, and that overall results may be better when rewards are not competitive.

Simultaneously Learning and Optimizing Prices
Bert Zwart* and Arnoud den Boer

We consider a monopolist firm selling a single product. The expected demand is a linear function of the selling price, with unknown parameters. At each time period the firm estimates these parameters using historical sales data; it then chooses the selling price for the next period, and collects new sales data. While setting the selling price, the firm needs to balance between learning the parameter values and maximizing instant revenue; e.g. between exploration and exploitation. We show that a certainty equivalent policy, which puts full emphasis on exploitation and none on learning, is not consistent. We therefore propose another pricing policy, called Controlled Variance Pricing (CVP). The key idea of CVP is to enhance the certainty equivalent policy with a taboo interval around the average previously chosen prices. This guarantees a lower bound on the speed of convergence of the prices. Using existing consistency results in recursive least squares, we show that CVP is consistent and characterize its performance. Numerical evaluations indicate that CVP performs better or comparable to other pricing policies from the literature, not only for linear but also for other demand functions.

(T2C) Stochastic Models 2

Analysis of a TCP System under Polling-Type Reduction-Signal Procedures
Omer Czerniak*, Uri Yechiali and Eitan Altman

The performance of a Transmission Control Protocol (TCP) for a system with N connections sharing a common Active Queue Management (AQM) is analyzed for both Additive-Increase Multiplicative-Decrease (AIMD) and Multiplicative-Increase Multiplicative-Decrease (MIMD) control mechanisms, where reduction signals follow either a cyclic or a probabilistic polling-type procedure. The Laplace-Stieltjes Transforms (LST) of the transmission rate of each connection at a polling instant, as well as at an arbitrary moment, are derived. Explicit results are obtained for the mean transmission rate and (in contrast to most polling models) for its second moment. The analysis of the probabilistic MIMD models uses transformations yielding a system's law of motion equivalent to that of an M/G/1 queue with batch service. The MIMD control mechanism with probabilistic strategy is further analyzed for the case where the transmission rate is bounded above.

Two-Queue Systems where Customers of Each Queue are the Severs of the Other Queue
Efrat Perel* and Uri Yechiali

We consider systems comprised of two interlacing queues where customers of each queue are the servers of the other queue.

Denoting by L_i the number of customers in queue i = 1, 2, we study four models:

(*i*) Queue 1 (Q_1) operates as a multi-server limited-buffer $M(\lambda_1)/M(\mu_1)/L_2/\max\{0, N-L_2\}$ system, while queue 2 (Q_2) operates as a single-server $M(\lambda_2)/M(\mu_2 L_1)/1$ queue. That is, at any moment, the L_2 customers present in Q_2 act as the servers of the limited-buffer Q_1, where service time of each individual customer is exponentially distributed with parameter μ_1. The customers of Q_2 are served by the L_1 customers in Q_1, who join hands together to form a single server with exponentially distributed service time of rate $\mu_2 L_1$. Arrivals to Q_i follow a Poisson process with rate λ_i. (*ii*) Q_1 operates as in model (*i*), but Q_2 operates as a multi-server $M(\lambda_2)/M(\mu_2)/L_1$ system. (*iii*) Q_1 is a limited buffer single-server $M(\lambda_1)/M(\mu_1 L_2)/1/\max\{0, N-L_2\}$ queue, while Q_2 is an $M(\lambda_2)/M(\mu_2 L_1)/1$ queue. (*iv*) Q_1 is as in model (*iii*), but Q_2 is an $M(\lambda_2)/M(\mu_2)/L_1$ system. For each model we derive the (conditional) probability generating function of L_1 (given L_2) and determine the condition for stability, where in models (*i*) and (*ii*) we apply Matrix Geometric analysis. We further calculate the means of L_i, along with their correlation coefficient, and compare between the models.

(T3A) Game Theory 2

Asymmetric Auctions Reconsidered
Yitzhaq Minchuk* and Arieh Gavious

We use perturbation analysis to obtain an explicit solution for expected revenue in asymmetric first-price auctions with two bidders. The bidders' valuations are drawn independently from perturbed uniform distribution functions, thereby introducing a weak asymmetry between bidders. We show that, contrary to common wisdom, with asymmetry, second-price auctions may generate higher revenue for the seller than first-price auctions

Bargaining over multiple issues: to commit or not to commit?
Arieh Gavious* and Ella Segev

We consider a repeated bargaining situation in which a seller sells one identical unit of a good each period to the same buyer. Under incomplete information on the buyer's valuation we compare two situations. In the first situation the seller can commit in advance to a series of prices and in the second he has no such ability and therefore responses to the buyer's behavior - whether he bought previous units and in which prices. We show that the seller's expected payoff is higher in the commitment game than in the no-commitment game while the buyer's ex-ante expected payoff is lower in the commitment game. Moreover, in the commitment game, the seller commits to the same price in all periods. We conclude that the commitment ability is an advantage. Finally, we prove that the social efficiency, the sum of expected payoff over all periods, is higher in the con-commitment game.

Core-stable rings in second price auctions with common values
Ram Orzach* and Françoise Forges

In a common value auction in which the information partitions of the bidders are connected, all rings are core-stable.

(T3B) Continuous Optimization 2

Facility Location: A Robust Optimization Approach
Hussein Naseraldin*, Opher Baron and Joseph Milner

In this research, we apply Robust Optimization (RO) to the problem of locating facilities in a network facing uncertain demand over multiple periods. We consider a multi-period fixed charge network location problem for which we find (1) the number of facilities, their location and capacities, (2) the production in each period, and (3) allocation of demand to facilities. Using the RO approach we formulate the problem to include alternate levels of uncertainty over the periods. We consider two models of demand uncertainty: demand within a bounded and symmetric multi-dimensional box, and demand within a multi-dimensional ellipsoid. We evaluate the potential benefits of applying the RO approach in our setting using an extensive numerical study. We show that the alternate models of uncertainty lead to very different solution network topologies, with the box uncertainty case opening fewer, larger facilities. Through sample path testing we show that the ellipsoidal uncertainty case provides small, but significant improvements over the solution to the problem when demand is deterministic and set at its nominal value. We explore the effects on the solution performance for changes in several environmental parameters.

A parametric method for constrained minimization problem with application to Image Deblurring Problem
Luba Tetruashvili* Aharon Ben-Tal and Amir Beck

We present a method for solving convex constrained minimization problem. In our method the original problem with functional constraints is replaced by the sequence of subproblems in which we minimize an improvement function over a convex set X. We use very simple updating rule for t values in which the next value of t is the sum of the approximated optimal value of the last solved subproblem and the last value of t. We showed that the approximated optimal values of subproblems converge to zero with the linear rate. Thus we can reach neighborhood of the optimal point in Omain steps. We did several numerical experiments as applied to large

image deblurring problems (up to 500000 variables). For solving subproblems we used Yuri Nesterov's in O algorithm

A Fast Proximal Gradient Method for Solving a Class of Nonsmooth Problems with Applications
Amir Beck* and Marc Teboulle

This talk is concerned with the general convex problem of minimizing the sum of a smooth convex function and an extended valued nonsmooth convex function. This model is quite general and encompasses classical convex (constrained) optimization problems. In applications, this model is commonly used in signal recovery problems when one requires to minimize the sum of a smooth least-squares term and a nonsmooth regularizer (such as l_1 norm or total variation function). We begin by exploring the proximal gradient method which is a natural generalization of the gradient method and, much like the latter, exhibits an O(1/k) rate of convergence. We then present a fast proximal gradient method with a guaranteed O(1/k^2) rate of convergence. We will demonstrate the effectiveness of the approach on image processing problems such as total variation-based and wavelet-based deblurring and denoising.

(T3C) OR Applications 2

Service Organization Workforce Optimization using Mixed-Integer0Programming and Simulation
Zohar Feldman*, Segev Wasserkrug, Dagan Gilat

For almost any enterprise, the workforce is one of its most important assets and serves as one of the primary cost factors out of the overall operational costs. There is a need for tools to efficiently manage this workforce, especially for two important needs: Design - long-term decisions like designing the workforce composition, planning the organizational processes and procedures and so on, and Day-to-day Operation - carrying out short-term operation planning, typically for up to several weeks in advance, such as creating a shift plan for all the employees. While there are some analytical methods and tools that attempt to address the above two problems, they are not suited to the complexities of real service operations. For example, many of the existing methods run into difficulty when attempting to address a workforce with a heterogeneous set of skills. In addition, none of the existing methods address environments in which workers have to fulfill complex business processes where each process comprises a set of tasks with routes between the tasks. We present a generic workforce optimization algorithm capable of handling such complexities. Our approach has several important virtues. First, we propose an end-to-end generic algorithm to provide a solution to the scheduling and capacity planning as a self-contained module and solve each of them as a single problem. This is in contrast to contemporary methodologies that decompose the problem into sub-problems, which are solved individually. For example, a common approach is to first calculate the staffing requirements for each time of the day (typically divided into small intervals) to meet desired service levels based on the predicted load. Next, this output is put together with the shift structure data to figure out how many employees should be assigned to each shift. Lastly, a schedule or capacity planning exercise is carried out to take the above into account. Solving each stage independently of other steps may reduce the quality of the solution significantly. Second, as described above, we accommodate complex service models, such as business processes, which are composed of several steps and follow a specific process flow. We try to estimate their performance as accurately as possible, as opposed to simplified approaches that use simplistic stationary queueing models, which do not capture the reality properly. They do not take into account the multi-skill environment, assume a very simple process with one single step, are based on steady-state analysis, assume exponential distributions, and so forth. Hence, results may be highly inaccurate. Third, our model takes into account both business rules stemming from work regulations and requirements, as well as service level objectives; these can take a very wide range of forms and may be expressed over various time horizons (daily, weekly etc.). Our approach is robust in the sense that given the reality of a limited workforce, it can still provide solutions that provide a minimal compromise for the goals defined as less important. We follow the same basic procedure with several modifications to solve both the capacity planning and scheduling problems.

Robust Capacity planning
Michael Masin* and Ira Goldenberg

Semiconductor manufacturing is a highly competitive industry with rapid innovation rate and long and uncertain lead-time. In turn, long lead times and high innovative rate cause very volatile and uncertain demand. The unmet demand is usually lost both through the specific order produced in another facility, but also through future sales. The common policy in semiconductor industry is to build a big building, prepare all needed facility and then to Ramp it up (purchase equipments in sequence). The reasons for purchasing new equipments are replacement of obsolete equipment and introduction of new products / technologies that require innovative, dedicated tools. Currently, a typical process requires about 400-600 different operations with a variety of tens needed tool types. One tool usually cost between one to eight million dollars. The capacity planning in Semiconductor means to determine the timing and amount of equipment's acquisition quantity that purchased from each equipment group. The problem is complicated by constrains on budget, existing tools from each group, available space and highly uncertain demand forecast. The goal is to find a tool-set robust for demand uncertainty and other uncertain parameters, while taking into account high capital investment costs or, in other words, to balance between the risk of low utilization (exceeded capacity) and unmet demand (capacity shortage). In this research, we concentrate on a single location (one facility) problem, with complex tool group / operation relationship, meaning that specific operation performed on more than one tool group and each tool group can perform more than one operation. The model deals with multi-product case, which includes different products that require different process flows (reentrant job shop). We formulate Mix Integer Linear Programming (MILP) model for both the nominal problem and its robust counterpart. The nominal model aggregates and extends all known models for purchasing planning of equipments. To the best of our knowledge, our robust counterpart is the first model for purchasing planning with uncertain demand dealt through robust counterpart. We also develop multi-step Dynamic Robust Counterpart. At each step, a decision regarding tool's acquisition is implemented only when it cannot be postponed. Then the model is updated with the implemented decisions and realized demand and rerun for the remaining periods. Computational experiments, including real size industrial models, provide insights on the problem complexity and the importance of robust solutions.

Real-Time Workforce Scheduling in Multi-Process Environment

Sergey Zeltyn*, Mike Marin, Michael Masin, Segev Wasserkrug and Dagan Gilat

This research is motivated by a practical problem in real-time workforce scheduling. Scheduling is performed in a challenging environment that involves multiple instances of different business processes, general process structure that enables sequential and parallel step processing, stochastic service times and random routing. Such environment is typical for many large service enterprises such as banks and insurance companies. Resources are assumed multi-skilled and can have different service rates. For each process, a non-decreasing penalty function is defined. This function depends on the sojourn time of a process instance and provides relation between operational and business metrics. The objective function, subject to minimization, is the sum of the sojourn time penalties for all instances per time unit. Based on several theoretical insights, borrowed from other application areas, such as call centers and jobshops, we design a decentralized dispatching rule that enables real-time implementation in large-scale business process environments. Given a set of idle resources and a set of waiting steps from all active process instances, a service index for each feasible resource/step combination is computed. Then the maximal service index is determined and the corresponding resource/step assignment is performed. The service index is computed as the product of several components, where each component corresponds to a specific scheduling principle. Our scheduling rule is tested using a broad set of practical examples, including those based on real-life scenarios. It implies significantly lower penalty costs than alternative well-established dispatching policies. In addition, we check that all components of the service index contribute to minimization of the penalty costs.

(T4A) Combinatorial Optimization

The Capacitated Spanning Tree Problem

Nili Beck*, Ester M. Arkin and Refael Hassin

This paper considers a generalization of the capacitated spanning tree problem, in which some of the nodes have capacity and the others have capacity . The results presented in this work: 1. For $k=1$: a. For $K=2$ we present a way to find the optimal solution. B. We present a $K-1$simple approximation algorithm; this

algorithm is suitable for small values of K . C. We also present a 6-approximation algorithm which is suitable for all values of K . 2. For $k = 2$ we present a 10-approximation algorithm, suitable for all values of K . 3. We present a 21-approximation algorithm suitable for all values of (K,k).

4. We consider a generalization of the problem where each node $v \in V$ has its capacity k_v. We present an $2+\alpha$ approximation algorithm for α which bounds the ratio between the maximal and minimal capacities in the graph.

An Improved Heuristic Algorithm for the Set Covering
Tsahi Avrahami*, Amnon Gonen and Eli Shemesh

The set covering problem is well known as an NP-complete problem. A common heuristic family of algorithms that solves the problem is the Greedy type algorithms. In this study we have to cover N sites by allocating minimum number of servers. Each server covers a predefined pattern and by locating it to a site K it covers its neighbors according to its pattern. The algorithm sorts all possible location of servers (sites) by the number of additional sites it covers. One of the minimal, called M, is chosen arbitrarily. Among all server locations that covers site M, it selects the location that covers the maximum number of sites. Now, actually, the algorithm covered the most "problematic" site. All covered sites are erased from lists of sites and the algorithm continues with the rest of sites. The study compares the Improved Heuristic Algorithm (IHA) with the Greedy one. The results showed that in most cases, for symmetric severs, the IHA finds a better solution (more than 10000 problems were tested). For big problems of above 600 sites, the probability to find a better solution is above 80%. Analyzing some cases where the Greedy algorithm found better solutions than the IHA, brought us to the conclusion that the weakness of the IHA is at the selection part of the minimal server location called M. However, we have not yet determined the best solution for these cases.

Probabilistic Testing of non Feasibility of the OS System
Ephraim Korach*, Ron Ben-Yishai, Sagi Hilleli and Michal Stern

The OS problem is: given a hypergraph (S,V) where V is a collection of items and S is a collection of subsets of V. A star clustering tree is a spanning tree of the complete graph G=(V,E) where each subset induces a star. Given a cost function on E *the* OS problem is to find among all feasible clustering star trees the one that have a minimum cost, i.e. the sum of the cost of its edges is minimum. In our research we would like to use the solution of the problem as a tool for solving another problem from the area of economics. As first step it would be helpful to test if a given instance has a feasible solution. We developed a probabilistic lower bound for the existence of a feasible solution, based on several necessary conditions for the existence of a feasible solution.

(T4B) Queuing Theory

Queueing In Multilevel Real-Time Systems
Jozeph Kreimer* and Edward Ianovsky

We consider a multilevel hierarchical Real-Time Systems (RTS) consisting of *s* service/maintenance levels and *r different* channels required to be under nonstop surveillance. There are N(i) identical servers at *i*-th level that provide maintenance for N(i-1) servers of lower (*i*-1)-th level. N(1) servers of first (lowest) level provide service to requests of RT jobs arriving via *r different*channels. There are exactly *r(k)* , jobs in *k*-th channel at *any* instant. Each channel has its own specifications and conditions, etc., and therefore different kinds of equipment and inventory are needed to serve different channels. A server of the *first* level is operative in *k*-th channel for a period of time S(1,k) before requiring R(1,k) hours of maintenance. S(1,k) and R(1,k) are i.i.d. exponentially distributed random variables. Each server of the *first* level coming back from a mission is assigned to the *k*-th channel with probability p(k) (*k*=1,...,*r*). Assignment probabilities can be used as control parameters. The duration R(1,k) of repair is exponentially distributed with parameter and does not depend on the channel. After maintenance, the server will either be on stand-by or serving the channel it was assigned to. A server of *i*-th level, (*i=2,...,s-1*) provides S(i) time units of maintenance (for servers of (*i*-1)-th level) before

requiring R(i) hours of maintenance. S(i) and R(i) are i.i.d. exponentially distributed random.variables. N(s) servers of highest *s*-th level are supposed to be absolutely reliable and practically unbroken. The system works under a maximum load (worst case) of nonstop data arrival to each one of *r* channels. If, during some period of time of length *T*, there is no available server to serve the job in one of the channels, we will say that the part of this job of length *T* is lost.

First come first served matching of multi-type customers and servers, in various queueing systems

Gideon Weiss* and Ivo Adan

We consider a system where jobs of several types are served by servers of several types, and a bipartite graph between server types and job types describes feasible assignments. This is a common situation in manufacturing, call centers with skill based routing, matching of parent-child in adoption or matching in kidney transplants etc. We consider the case of first come first served policy: jobs are assigned to the first available feasible server in order of their arrivals. We are in particular interested in the fraction of customers of each type which are served by a particular type of server, we call these matching rates. We will present results for four different systems:
- A loss system with random assignment and no queueing, with a reversible solution
- A stable manufacturing system, with a product form solution,
- An overloaded system with reneging,
- Infinite FCFS matching of an infinite sequence of customers and an infinite sequence of services.

We will also pose some questions for future research.

The M/G/1+G Queue Revisited

David Perry* Onno Boxma and Wolfgang Stadje

We consider an M/G/1 queue with the following form of customer impatience: an arriving customer balks or reneges when its virtual waiting time, i.e., the amount of work seen upon arrival, is larger than a certain random patience time. We consider the number of customers in the system, the maximum workload during a busy period, and the length of a busy period. We also briefly treat the analogous model in which any customer enters the system and leaves at the end of his patience time or at the end of his virtual sojourn time, whichever occurs first.

(T4C) Supply Chain Management

Robust optimization of multi-period production planning under uncertainty

Michal Rozenblit*, Aharon Ben-Tal and Boaz Golany

We address a single-product multi-period production planning problem in which the inventories are managed periodically over a finite horizon. The demand for the product is uncertain and is only known to reside within a user-defined uncertainty set *U* while drifting around a nominal trajectory. This nominal trajectory is assumed to follow a typical life-cycle pattern. To solve the problem we apply the *Adjustable Robust Optimization* (ARO) methodology which finds a solution that is feasible for all possible realizations of the uncertain parameters with a guaranteed objective value. We compare our solution to those obtained employing the common method that replaces the uncertain parameters by their nominal values, and to those obtained employing a sample-based method. The two latter methods produce solutions that may violate some of the constraints, as the true values of the demand drift from their nominal or sampled ones. In contrast, the ARO solution is immunized against infeasibilities, as long as the parameters drift within *U*. Intuitively, the ARO methodology could be expected to yield "conservative" solutions. However since in practice it is rare that the demand indeed follows the worst possible scenario, it is perhaps better to base the comparison on *average performance* over simulated demand trajectories. As shown by our simulation study, the ARO methodology, although based on suboptimal decision rules, achieved on average, solutions that are surprisingly close to the ultimate upper bound on profit (82% - for high uncertainty and 99.6% - for low uncertainty). Moreover, in the planning stage of a production planning problem (right before the first period) we can use the ARO guaranteed objective value to compare it with revenues that can be obtained by using the resources for alternative investment opportunities. It can also be used

in negotiations, e.g., over the selling price or the shortage penalty (which is sometimes considered as customer compensation). In addition, the ARO solution can be employed together with the nominal solution to estimate a horizon length after which it becomes unprofitable to produce and sell the product. Furthermore, we solve a more realistic problem including fixed order cost and show that a folding horizon approach of RO results in a*higher profit on average* compared to the nominal solution.

Models for Evaluating Profitability of Investing in Improvements

Gabi Pinto and Beni Bar Yosepf

The most common measure of profitability of investment in improving processes and facilities prevailing in Industrial Engineering (IE) in general and in Work-Study (WS) in particular is return-on-investment (ROI). In order to calculate the ROI of an investment, the traditional attitude of IE and WS literature defines two market types, "saturated" and "non-saturated". In a saturated market the general idea is to keep a constant output while reducing the total cost. In a non-saturated market there is more flexibility with a general purpose of reducing the cost per unit, either by reducing cost while keeping a constant output or, preferably, by increasing output while increasing cost at a lower rate. Also, in a non-saturated market the output is measured in a situation of constant product (or service) unit price. Rising markets and different types of environment call for different models of evaluation. The traditional attitude ignores special conditions presented by the new evolving markets, such as a new market, a new product, a learning market and strong influence of output on unit price (not only cost). These conditions call for considering learning-curve of the market and/or the producer. In this paper we present different models of evaluating improvements in non-saturated markets, with new and existing processes. While the traditional ROI model ignores the learning phenomena of a new process, some of the presented models do consider a learning curve of a new process in a non-saturated market. We will present the following models: 1. Selling price for the additional units is lower (e.g., special offer or selling to a foreign country). 2. Price is reduced for the whole output (e.g., as in agricultural products). 3. Price is constant for the original quantity, but is decreasing gradually, for the additional quantity, according to the learning curve (e.g., selling the additional quantity to a learning market). 4. Price is decreasing from the original price according to the learning curve for the whole quantity produced (e.g., as in the semiconductor industry). Most of the models we present go through the same stages; the first stage is to calculate the cost per unit before and after the improvement. The second stage is to calculate the efficiency per time unit, and the final stage is to calculate the ROI.

Staffing and Routing in a Two-call Center System with Transfer Cost: Supply Chain Methodology

Yanir Cohen*, Yale T. Herer and Segev Wasserkrug

In this talk, we address a problem faced by companies which operate call centers in two distinct geographical locations. Each call center is modeled as a Markovian queueing system with a fixed number of servers. Calls arrive to each center according to a Poisson process. In general, the two call-centers are allowed to be unique both with respect to their costs and operational parameters. Moreover, we model the ability to transfer calls among the call centers for a given cost. The service rate may depend on both the call origin and the location where the call is handled. The cost function consists of the following types of costs: staffing cost, which is the cost per unit time of having a server at a particular location, waiting cost, which is the cost per unit time of having a customer wait to be served, and transfer cost, which is the cost of transferring a call from one call center to the other. This cost includes the direct cost of routing the call to the other call center and the indirect costs which are a consequence of the difference in quality between the two locations (e.g. language and/or cultural issues). In this setting, we find staffing levels at each of the call centers and a policy for determining when to transfer calls between the centers in order to minimize the expected total cost. These staffing levels are found by utilizing two different approaches: queueing theory and inventory theory. Our Queueing Theory analysis is carried out under the assumption that each call center is operated under the QED (Quality and Efficiency Driven) regime**.** Under this regime, we consider two policies for transferring calls: static and dynamic. In the static policy, we find when it is preferable to transfer all calls from one call center to the other as opposed to not transferring any calls. For the dynamic policy, we find a call transfer rule for minimizing the expected total cost, basing the transfer decision on the system's state (e.g. the queue lengths). As there is no general analysis for this problem using Queueing theory, we also show how Transshipment methodologies taken from inventory theory may be applied to solve the problem. We do this by making an analogy between calls and physical goods. We also compare the results of the policies provided on the two approaches on a simulation model of the system.