Computational Sustainability: Computational Methods for a Sustainable Environment, Economy, and Society

Carla P. Gomes Institute for Computational Sustainability Cornell University Sponsored by:

This talk is an adaptation of the NSF reverse site visit talk or the Expeditions In Computing Program (June 2008). Thanks to the ICS members who helped shape the vision I ormulated n this talk .



The 1987 UN report, "Our Common Future" (Brundtland Report):

- Raised serious concerns about the State of the Planet.
- Introduced the notion of sustainability and sustainable development:

Sustainable Development: "<u>development that meets the needs</u> of the present without compromising the ability of future generations to meet their needs."

> The UN General Assembly stressed that environmental problems were global in nature and stated the urgency of policies for sustainable development.



UN World Commission on Environment and Development, 1987.





Gro Brundtland Norwegian Prime Ministe Chair of WCED Computational

Follow-Up Reports: **CS** Intergovernmental Panel on Climate Change (IPCC 07) Global Environment Outlook Report (GEO 07)

> "There are no major issues raised in Our Common Future for which the foreseeable trends are favourable."



Temperature change "C 1970-2004

-0.2 0.2 1.0 2.0 3.5

Global Warming

-10







Erosion of Biodiversity

Examples:

•The biomass of fish is estimated to be 1/10 of what it was 50 years ago and is declining.

 At the current rates of human destruction of natural ecosystems, 50% of all species of life on earth will be



INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE (IPCC)





extinct in 100 years.



Main Causes of Damage to Earth:

Poor Management of our Natural Resources

Pollution











Habitat Loss and Fragmentation









Over-Harvesting











I Computational Sustainability II Computational Themes in Our Research III Institute for Computational Sustainability IV Compsust09



Computational Sustainability

II Computational Themes in Our Research III Institute for Computational Sustainability IV Compsust09



 The advancements in communication and computation have dramatically transformed traditional business models.

e.g., electronic markets, just-in-time manufacturing, combinatorial auctions, and customer data mining.

 The impact of information technology has been highly uneven, with little benefit in terms of the environment.



Computational Nature of Decision and Policy Making Problems in Sustainability

Key sustainability issues concerning the definition of policies for sustainable development translate into decision, optimization, statistical and learning problems that fall into the realm of computer science and related fields (information science, operations research, applied mathematics, and statistics).



- Unique in scale, impact, complexity, and richness;
- Often involving combinatorial decisions, in highly dynamic and uncertain environments.

 \rightarrow Offer challenges but also opportunities for the advancement of the state of the art in computing and information science.

 Unfortunately, in general computer scientists are not aware of these challenging problems.



Computer scientists can — and should — play a key role in increasing the efficiency and effectiveness of the way we manage and allocate our natural resources, while enriching and transforming Computer Science.



Computational Sustainability

Computational Sustainability --- interdisciplinary field that aims to apply techniques from computer science, and related fields(information science, operations research, applied mathematics, and statistics) for balancing environmental, economic, and societal needs for sustainable development.



Focus:

Developing computational & mathematical models and methods for decision making concerning the management and allocation of resources in order to help solve some of the most challenging problems related to sustainability



Examples of Computational Sustainability Problems



Examples of Sustainability Themes

I Conservation and Biodiversity

Wildlife Corridors



II Balancing Socio-economic Demands and the Environment

Policies for harvesting renewable resources





III Renewable Energy

Biofuels and other alternative energies







Challenges in Constraint Reasoning and Optimization: Conservation and Biodiversity: Wildlife Corridors

Wildlife Corridors link core biological areas, allowing animal movement between areas.

Typically: low budgets to implement corridors.

Computational problem → Connection Sub-graph Problem



Connection Sub-Graph - NP-Hard

Worst Case Result --- Real-world problems possess hidden structure that can be exploited allowing scaling up of solutions→ Science of Computation.



Connection Sub-graph Problem

Given a graph G with a set of reserves:

Find a sub-graph of G that:

contains the reserves; is connected; with cost below a given budget;

and with maximum utility



"Typical" Case Analysis: Synthetic Instances

Runtime How is hardness affected 100.00 as the budget fraction is varied? optimal solution Runtime (logscale) 10.00 optimal extended soln × Problem evaluated on semi-structured graphs 1.00 *m* x *m* lattice / grid graph with *k* terminals Inspired by the conservation corridors problem 0.10 Place a terminal each on top-left and bottom-right Maximizes grid use Place remaining terminals randomly 0.01 Assign uniform random 0 200 400 600 800 1000 costs and utilities budget slack % (w.r.t. mincost) from {0, 1, ..., 10} **Runtime for Optimal Solution** Utility Gap (Optimally Extended Min cost/ Optimal)



Structure

Real world instance:

Corridor for grizzly bears in the Northern Rockies, connecting:

Yellowstone Salmon-Selway Ecosystem Glacier Park

(12788 nodes)

Scaling up Solutions by Exploiting Structure:

Typical Case Analysis Identification of Tractable Sub-problems Exploiting structure Streamlining for Optimization Static/Dynamic Pruning 5 km grid (12788 land parcels): minimum cost solution

Yellowstone

Glacier Park

5 km grid (12788 land parcels): +1% of min. cost

Our approach allows us to handle large problems and reduced corridor cost dramatically compared to existing approaches [Conrad et al. 2007]

Salmon-Selway

Interdisciplinary Research Project (IRP):

Wildlife Corridors (Conrad, Gomes, van Hoeve, Sabharwal, Sutter)

CompSust09: Poster and Ashish Sabharwal will talk more about this problem



Additional Levels of Complexity: Stochasticity, Uncertainty, Large-Scale Data Modeling

- Highly stochastic environments
- Multiple species (hundreds or thousands), with interactions (e.g. predator/prey).
- Spatially-explicit aspects within-species
- Different models of land conservation

 (e.g., purchase, conservation easements, auctions)
 typically over different time periods
- Dynamical models

Dynamics of Species Movements and migrations; CompSust09: Natural Resource Analysis and Decision Making Williams, Runge, Conroy McDonald-Madden

Species Distributions, Biodiversity & Ecological Models: Elith, Farnsworth,Fink, Hochachka ,Kelling, ns) Langley, Los,Munson, Phillips, Riedewald, Sabaddin, Sheldon

> Ecological Monitoring & Computer Vision Dietterich, Guestrin, Los, Kumar, Krause, Nichols, Pauwels

S ¹⁷

Optimization models for Red-Cockaded Woodpecker management

Dilkina, B., Elmachtoub, A., Finseth, R., Sheldon, D., Conrad, J., Gomes, C., Shmoys, D., Amundsen, O., and Allen, W. Cornell University and The Conservation Fund

Introduction

- Degradation and loss of longleaf pine ecosystem has led to decline of Red-Cockaded Woodpecker (RCW)
- 'Keystone' species primary excavators of nest cavities used by at least 27 vertebrate species
- Historically 1.0 to 1.6 million breeding groups, today only 5,600 existing RCW breeding groups
- Highly specific habitat need mature pine trees infected with Red Heart fungus
- Cooperative breeders territory groups consisting of one breeding pair and up to four 'helpers'
- Conservation and habitat management crucial to continued viability of Red-Cockaded Woodpecker

Research Objectives

The goal of this research is to develop methods to prioritize land acquisition adjacent to current RCW populations to aid in their recovery.

We seek to pose this as a formal optimization problem: where and when should one acquire land parcels and/or translocate birds to maximize the number of RCW breeding groups.

To solve this problem we develop a diffusion model to describe spatial patterns in RCW populations, and pose this as a stochastic network design problem. **Study Area**

Palmetto Peartree Preserve (3P) consists of 10,000 acres of wetland forest in Tyrrell County, North Carolina. As of September 2008, there were a total of 32 active RCW territories within the preserve



Figure 1. 3P RCW territories shown in blue

Patch-based Diffusion Model

 Based on cascade models for spread of influence in social networks; also related to metapopulation models in ecology

Model Description

- Territories *i*, *j* =1, ..., n.
- Occupied or unoccupied at each time step
- May colonize other territories (probability *p_{ij}*), or go extinct (probability β) in each time step
- Unoccupied territories become occupied if colonized by one or more other territories
- All colonization and extinction events independent

Parameters

- Colonization probability decays with distance, and only succeeds if target territory has suitable habitat
 - $p_{ij} = \begin{cases} q_{ij} & \text{if territory } j \text{ is suitable} \\ 0 & \text{otherwise} \end{cases}$

Illustration



Simulation Results 2

• Spatial configuration is very important. Dense and highly connected configurations are most stable.

• The four scenarios below show the effect of territory density on occupancy in the 3P study



Occupancy as Network Connectivity

• We can describe the occupancy patterns of RCWs using a graphical network model



- The circles represent a territory in a specific year
- Horizontal lines between squares inside the circles indicate suitability of that territory in that year
- Horizontal lines between circles indicate nonextinction from one year the next. These are present with probability 1 - β
- Diagonal lines indicate potential colonization events; colonization occurs only if the source territory is occupied. These edges are present with probability *p_{ij}*
- Blue lines indicate actual colonization and nonextinction (e.g. territory i colonizes territory j at t=2)
- **D**The occupied territories at t=3 are only those that can be reached from t=1 by a sequences of
- Wdgsample many scenarios representing different outcomes of colonization and extinction events
- Goal: maximize the number of colonized territories at time T, averaged over all scenarios
- Decision variables: which territories to purchase (i.e., make suitable) and in which time period
- Budget constraint limits the total cost of the territories we can purchase
- Purchase constraints let us buy each territory once
- Flow constraints between territories
- Capacity constraints restrict flow to suitable and colonization edges
- Integrality conditions on decision and flow



Solving Large-Scale Models

- Large mixed-integer programs (MIP) like ours are very difficult to solve
- We have employed the following "LProunding" approach rather than solving the MIP directly:
 - Solve the relaxed LP version
 - Set any integer variables <.1 to 0
 - Set the largest integer variable to 1
 - If new bounds result in infeasibility, set the previous variable to 0
- Repeat until an integer solution is reached
 This approach is generally much faster than solving the original and obtains close to optimal results
- The table below shows the results for testing our 33 territories for 10 years, 5 simulations, random territory costs and a variable budget

D			
^D Budget	IP solution	Lp rounding	%optimal
300	6.6	5.8	87.9%
400	8.4	6.6	78.6%
500	10.2	10	98.0%
600	12	11.4	95.0%
700	13.6	13.2	97.1%

Acknowledgments

The authors gratefully acknowledge the support of the National Science Foundation, award number 0832782. The authors also thank Dr. Jeffrey Walters of the Virginia Polytechnic Institute for granting the use of the RCW DSS.



Challenges in Dynamic Models and Optimization :



We are interested in defining optimal (good) **policy decisions** (e.g. when to open/close a fishery ground over time). **New Class of Hybrid Dynamic Optimization Models** Combinatorial optimization problems with an underlying dynamical model. Challenges in Highly Interconnected Multi-Agent Systems: Renewable Energy

Energy Independence and Security Act





Large-scale investment in new technology provides exciting logistical planning and optimization challenges and opportunities



Resulting optimization models are beyond scope of current "facility-location" expertise in several ways:

- -large-scale input;
 -stochastic nature
- (e.g.,feedstock and demand)
 - →new models to capture uncertainty
 - →new stochastic
 - optimization algorithms;
- -dynamics of evolution of demand and capacity



Current approaches limited in scope and complexity

- E.g. based on *general equilibrium models* (e.g., Nash style)
- Strong convexity assumptions to keep the model simple enough for analytical, closed-form solutions (unrealistic scenarios)

Antonio Bento General Equilibrium Models for Biofuels

→ Limited computational thinking

Transformative research directions

- More realistic computational models in which meaningful solutions can be computed
- Large-scale data, beyond state-of-the-art CS techniques
- Study of dynamics of reaching equilibrium key for adaptive policy making!



How to measure risks/ predict rare events?

Impact of Biofuels: Dynamic Equilibrium Models

Impact of Land-use on Climate

Policies for a carbon cap and trade economy

Power Grid, Transportation & Sustainable Communities







I Computational Sustainability II Computational Themes in Our Research III Institute for Computational Sustainability IV Compsust09



Our approach:

The study computational problems as natural phenomena in which principled experimentation, to uncover hidden structure, is as important as formal analysis

\rightarrow Science of Computation,

Our team has a track record of making compelling scientific discoveries using such an approach.

Small world phenomenon

Pioneered science of networks



Phase Transitions in Computation

Led to interactions between CS, statistical physics, and math





Theme: Science of Computation Discovering patterns, laws, and hidden structure in computational phenomena

Streamlining Constraint Reasoning

Discovery of structural properties across solutions (machine learning);

Divide ("streamline") the search space by imposing such additional properties.





Design of Agronomic Experiments for Studying Fertlizers

 \rightarrow Scaling up of solutions

→Domain Independent Approach: XOR-Streamlining based on random parity constraints; Provable bounds on solution counting

Van Es et al 2005



Key sustainability issues concerning the definition of policies for sustainable development translate into largescale decision/optimization combining a mixture of discrete and continuous effects, in a highly dynamic and uncertain environment

 \rightarrow different levels of complexity

Study computational problems as natural phenomena

 \rightarrow Science of Computation

Many highly interconnected components;

→ From Centralized to Distributed Models

Multiple scales (e.g., temporal, spatial, geographic) → From Statics to Dynamics: Dynamic Models

Large-scale data and uncertainty

→ Machine Learning, Statistical Modeling, Stochastic Modeling Complex decision models

 \rightarrow Constraint Reasoning and Optimization



Complexity levels in Computational Sustainability Problems 27



Transformative Computer Science Research: Driven by Deep Research Challenges posed by Sustainability

Design of policies to effectively manage Earth's natural resources translate into large-scale decision/optimization and learning problems, combining a mixture of discrete and continuous effects, in a highly dynamic and uncertain environment \rightarrow increasing levels of complexity

Study computational problems as natural phenomena → Science of Computation

Many highly interconnected components;

→ From Centralized to Distributed: Computational Resource Economics

Multiple scales (e.g., temporal, spatial, geographic) → From Statics to Dynamics: Dynamic Models

Large-scale data and uncertainty

→ Machine Learning, Statistical Modeling

Complex decision models

→ Constraint Reasoning, Optimization, Stochasticity

Constraint Reasoning & Optimization Gomes PI,W Hopcroft Co-PI Selman^{Co-PI} Shmoys^{Co-PI}

Resource Economics, Environmental Sciences & Engineering Albers^{Co-PI,W}, Amundsen, Barrett, Bento, Conrad^{Co-PI}, DiSalvo, Mahowald^W, Montgomery^{Co-PI,W} Rosenberg,Sofia^W, Walker^{AA}

Renewable

our of the second

Dynamical

Systems

Guckenheimer

Strogatz Zeeman^{PI,W}

Yakubu^{AA}

Environmental & Socioeconomic Needs



Statistics & Machine Learning Dietterich^{PI} Wong^{Co-PI} Chavarria^H

Science of Computation





I Computational Sustainability II Computational Themes in Our Research III Institute for Computational Sustainability IV Compsust09



- Perform and foster research in Computational Sustainability
 - new insights into sustainability questions;
 - new challenges and new methodologies in Computer Science and related fields

(Analogous to Computational Biology)

 Establish a vibrant research community, reaching far beyond the members in the original NSF Expedition.



- Perform and foster research in Computational Sustainability
 - new insights into sustainability questions;
 - new challenges and new methodologies in Computer Science and related fields

(Analogous to Computational Biology)

 Establish a vibrant research community, reaching far beyond the members in the original NSF Expedition.

Multi-institutional, Multidisciplinary Research Team 6 Institutions, 7 colleges, 13 departments





Interdisciplinary Research Projects (IRPs): The Building Blocks of our Expedition

Seedling IRPs

	IRP Name	Faculty Team
1.	Wildlife Corridors for Grizzly Bears	Amundsen, Conrad, Gomes, Selman, Shmoys
2.	Biofuels	Bento, Gomes, Mahowald, Shmoys, Strogatz, Walker, Wong
3.	Bird Conservation	Rosenberg, Conrad, Dietterich, Gomes, Hopcroft, Strogatz, Zeeman
4.	Native Plant Habitat Recovery in Victoria, Australia	Dietterich, Gomes, Selman
5.	Joint Public/Private Management for Biodiversity	Amundsen, Montgomery, Dietterich, Gomes, Hopcroft
6.	Fire Management in Forests	Albers, Conrad, Guckenheimer, Selman
7.	Rotational Management of Fishing Grounds	Conrad, Guckenheimer, Yakubu, Zeeman
8.	Pastoral Systems in East Africa	Barrett, Guo, Gomes, Toth



Nurturing New IRPs

Application/Synthesis Facilitators

			C	computation	onal Them	ies			
		Science of Computation	Dynamics & Optimization	Optimization & Learning	Learning & Dynamics	Agent Integration	Dissemination, Outreach, & Diversity		
reas	Conservation and Biodiversity	4	4		1		4	Conrad, Montgomery, Gomes	L S
lication A	Balancing Socio-Economic Demands and Environment	4	4	4	J	4	1	Albers, Conrad, Guckenheimer	vpplicatio acilitators
App	Renewable Energy	4	4	✓		4	1	Bento, Gomes, Shmoys	A T
		Dietterich, Hopcroft, Selman	Guckenheim er, Shmoys, Zeeman	Dietterich, Gomes, Hopcroft	Guckenheim er, Wong	Bento, Gomes, Shmoys	Hopcroft, Zeeman		
				Synthesis	Facilitato	rs			

Institute Activities

Research			E	Building Research Community				
Constraint Reasoning A Optimization Owner ** Optimize State States **		Coordinating transformative synthesis	C	onference & Workshops				
1.3		collaborations			V	/eb Portal		
Hill Regards Economica			Weekly Seminar					
Dynamical Dynamical Models Dynamical Dynamic		Interdisciplinary Research Projects (IRPs)	External		Hos Sc	t visiting ientists		
Revealed the Dynamics and Training	Constant Constant			Conaboratio	no			
Education				Outreach	K-12 A	ctivities		
	Postdocs						Conservati	on
Doctoral students	Honors	Research seminar	Citiz	zen Science Project			Fund	UII
	projects	series			Science	e Exhibits	6	
Summer REU p minority	orogram targ students	eting	C S	Cornell Center Sustainable Fi	for a uture		Cornell Cooperativ Extensior	/e า
Computational Sustainability courses		OSU Alliance for Computational Sustainability						



I Computational Sustainability II Computational Themes in Our Research III Institute for Computational Sustainability IV Compsust09



Conference CompSust09: Goals

Bring together a community of researchers, educators, policy makers, practitioners, and students interested in applying techniques from computer science and related fields to solve key challenges in sustainability.

About 200 researchers from 80+ different research departments, labs and government institutions.

Establish a two-way street between environment researchers and researchers in computer science and related fields:

- Work on a common language for such a multi-disciplinary community.
- Educate computer scientists about computational aspects of challenges in sustainability;
- Educate of researchers in sustainability about what models and techniques computer science and related fields can offer

Computational Sustainability is a fundamentally new intellectual territory with great potential to advance the different disciplines involved and with unique societal benefits!



Enjoy the Conference!!!

Thank You ©!

Computational facilities for biodiversity research (e-Infrastructures / Examples of computer-assisted photo-

identification of individual animals)

	Tuesday, June 9, 2009	1st Conference on Computational Sustainability June 8–11, 2009	CompSust09	
8:30 am	Registration	Cornell University, Ithaca, NY USA	Toparticipate: email comp_cuttiges.comoil.edu	
9:00 am	Wouter Los. Eric Pa	uwels		

8:00 am Registration 9:00 am **Opening Remarks** Cornell University Provost, Kent Fuchs Computational Sustainability: Computational Methods for a Sustainable Environment, Economy, & Society Carla P. Gomes 9:30 am Ken Williams Natural Resources Analysis and Decision Making 10:10 am Mike Runge Challenges in Framing the Problem: Just What are We Trying to Optimize Anyway? 10:50 am Break 11:05 am Mike Conroy Challenges in Dynamic Optimization for Natural Resource Management 11:45 am Eve McDonald-Madden Spatial and Temporal Issues in Resource Allocation 12:25 pm Lunch 1:50 pm Steve Kelling Biodiversity Research & Conservation in a Digital World 2:30 pm Jane Elith The Art of Modelling Range-Shifting Species 3:10 pm Steve Phillips Maximum Entropy Modeling of Species Distributions 3:50 pm Break 4:05 pm Vipin Kumar Discovery of Patterns in Global Earth Science Data using Data Mining 4:45 pm Jim Nichols On the Ugliness of Ecological Monitoring: Computational Constraints Arising from Ecological Data and Inference Methods 5:25 pm Poster Session I

9:40 am Carlos Guestrin TBA 10:20 am Andreas Krause Sensing Challenges in Environmental Monitoring 10:40 am Break 10:55 am Warren Powell Approximate Dynamic Programming for a Stochatsic, Multiscale Energy Policy Model 11:35 am Brian Williams Risk Sensitive Planning, Deep Sea Exploration and Sustainable Connected Homes 11:55 am Lunch 1:20 pm Ian Dobson TBA 2:00 pm Alan Borning The UrbanSim Project: Using Urban Simulation to Inform Public Decision-making about Land Use and Transportation Choices 2:30 pm James Landay Environmental Sustainability Through Activity-based Computing 3:00 pm Break 3:20 pm Terry Quinn Incorporating Biological and Environmental Realism into Fisheries Stock Assessment Models 3:50 pm Gautam Sethi Have Your Fish and Eat Them Too: Fishery Management Under Multiple Uncertainty 4:20 pm Richard Howitt Modeling Dynamic Network Systems with State-Contingent Penalty Functions 5:00 pm Poster Session II

Wednesday, June 10, 2009



Thursday, June 11, 2009

9:00 am High-Perform versity Studi	Thomas Dietteric mance Computer Vision for Arthropod Biod ies	ר <i>ו</i> -
9:40 am Model-based Map Constru		n e
10:10 am Mathematica munication I Planning	al Programming-based Heuristics: Telecon Network Design Meets Species Distributio	s 7- 17
10:40 am	Break	
11:00 am C o	omputational Sustainability Panel	
12:00 pm	Lunch (on your own)	
1:30 pm	Concurrent Working Groups	
Room A Problems in	Quantifying and Managing Uncertainty	ı
Room B Interactive D	Decision Support Tools	1
3:20 pm	Break	
3:40 pm	Concurrent Working Groups	
Room A Assessing M tainty in Soci	Multiple Sources of Human-Behavioral Unce ial-Ecological Systems	r r-
Room B Species Dist	tribution	3

9:00 am
9:30 am Jon Conrad Maximin Utility with Fractional Consumption, Extraction, and Harvest Rates
10:00 am Antonio Bento <i>TBA</i>
10:20 am Break
10:40 am Claire Montgomery Optimal Forest Fire Fuel Treatment and Timber Harvest in the Face of Endogenous Spatial Risk
11:00 am Pat Langley An Interactive Environment for Constructing Ecological Models
11:20 am Mark Battle Development of an Optimal Method for Calibration of Gas Analyses
11:40 am
12:00 pm Lunch (on your own)
1:30 pm Concurrent Working Groups
Room A Chair: Allen Sustainable Communities & Sustainable Agriculture
Room B Chair: McDonald-Madden & Shmoys Problems in Spatial Modeling and Resource Allocation
3:20 pm Break
3:40 pm Working Group
Room A Chair: Conrad

Dynamic Optimization of Renewable Resources