Human and Modeling Approaches for Humanitarian Transportation Planning

by

Erica L. Gralla

B.S.E., Princeton University (2004)
S.M., Massachusetts Institute of Technology (2006)

Submitted to the Engineering Systems Division
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2012

ARCHIVES
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
OCT 26 2012
LIBRARIES

© Massachusetts Institute of Technology 2012. All rights reserved.

Author ....
Engineering Systems Division
August 31, 2012

Certified by ....
Charles Fine
Chrysler LGO Professor of Management and Engineering Systems
Thesis Supervisor

Certified by ....
Jarrod Goentzel
Director, MIT Humanitarian Response Lab

Certified by ....
John Carroll
Adelman Professor of Management and Engineering Systems

Certified by ....
Jérémie Gallien
Associate Professor of Management Science and Operations
London Business School

Certified by ....
Richard Larson
Mitsui Professor of Engineering Systems

Accepted by ....
Olivier L. de Weck
Chair, Engineering Systems Division Education Committee

Human and Modeling Approaches for Humanitarian Transportation Planning

by

Erica L. Gralla

Submitted to the Engineering Systems Division
on August 31, 2012, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

Recent disasters have highlighted the need for more effective supply chain management during emergency response. Planning and prioritizing the use of trucks and helicopters to transport humanitarian aid to affected communities is a key logistics challenge. This dissertation explores ways to improve humanitarian transportation planning by building on the strengths of both humans and models. The changing, urgent, multi-objective context of humanitarian aid makes it challenging to formulate and deploy useful planning models. Humans are better able to understand the context, but struggle with the complexity of the problem. This research investigates the strengths and weaknesses of human transportation planners in comparison with models, with the goal of supporting both better human decision-making and better models for humanitarian transportation planning.

Chapter 2 investigates how experienced humanitarian logisticians build transportation plans in a simulated emergency response. Based on an ethnographic study of ten logistics response teams, I show how humans come to understand the problem and its objectives through sensemaking, and solve it through a search-like series of decisions guided by goal-oriented decision rules. I find that the definition of objectives is an important strength of the sensemaking process, and that the human reliance on greedy search may be a weakness of human problem-solving.

Chapter 3 defines a performance measure for humanitarian transportation plans, by measuring the importance of the objectives identified in the ethnographic study. I use a conjoint analysis survey of expert humanitarian logisticians to quantify the importance of each objective and develop a utility function to value the performance of aid delivery plans. The results show that the amount of cargo delivered is the most important objective and cost the least; experts prefer to prioritize vulnerable communities and critical commodities, but not to the exclusion of others.

Chapter 4 investigates the performance of human decision-making approaches in comparison to optimization models. The human decision-making processes found in Chapter 2 are modeled as heuristic algorithms and compared to a mixed-integer linear program. Results show that optimization models create better transportation plans, but that human decision processes could be nearly as effective if implemented consistently with the right decision rules.

Thesis Supervisor: Charles Fine
Title: Chrysler LGO Professor of Management and Engineering Systems

Acknowledgments

First and most importantly, I wish to thank my advisors, Charles Fine and Jarrod Goentzel. Charlie has supported, encouraged, and enriched my work in immeasurable ways, from supporting my wide-ranging exploration of a new topic area to pushing me to find exciting and significant insights. Jarrod has been an invaluable source of advice and support. I benefited from his enthusiasm and expertise in humanitarian logistics, our daily discussions on research and the humanitarian sector, our varied forays into "the field," and many other experiences too numerous to record.

I also wish to extend a very grateful thanks to the other members of my committee. John Carroll was always encouraging and supportive, especially in framing and refining my work on human decision-making. Jérémie Gallien helped to shape many aspects of my work, especially the study of humanitarian objectives, and I benefited from his advice on both research and career. Richard Larson provided valuable expertise in disaster response and engineering systems, and I am very grateful for his support since my early days at MIT.

I am extraordinarily grateful to everyone at the World Food Programme (WFP) and other humanitarian organizations who provided valuable data, thoughts, and feedback, gave me the opportunity to participate in the Logistics Response Team training, and welcomed me into the community. I especially want to thank everyone at WFP's Logistics Development Unit and the Logistics Cluster, and all the facilitators and participants at all five LRT trainings I attended. In this thesis, I have tried to do justice to everything that I have learned from them. I also want to acknowledge the assistance, advice, and guidance of several individuals who were particularly helpful in this research, including Bernard Chomilier, Philippe Sallet, Daniel Dupraz, Bekim Mahmuti, Baptiste Burgaud, and Andrew Stanhope. Special thanks also to everyone with whom I discussed my research or who encouraged me along the way (Adam, Amin, Cameron, David, Edita, Ellen, Emmanuel, Giorgi, Hetty, Lado, Madeline, Mallory, Matteo, Pierre, Tony, and many others). Thank you.

At MIT, I was privileged to be part of the Engineering Systems Division and the Center for Transportation and Logistics. The faculty, researchers, staff, and students of ESD and CTL were invaluable to my research, from providing feedback on my ideas to shaping my view of the world. I especially wish to thank Olivier de Weck, Edgar Blanco, Yossi Sheffi, Karen Golden-Biddle (at BU) and Ozlem Ergun (at Georgia Tech) for their various contributions, past and present, to my research and my thinking. In addition, I want to thank the other faculty, researchers, and students, too numerous to name, who have provided advice, guidance, and support over the years I spent at ESD and CTL. I learned as much from hallway conversations as I did from research seminars and formal courses. Thanks also to the ESD and CTL staff, for their support, friendship, and patience. Finally, thanks to all my colleagues who listened to my research presentations, gave me feedback, tested my survey and tools, and bounced around ideas.

I am also very grateful to Lincoln Laboratory, especially to the groups that supported my dissertation work, and the people with whom I worked in Haiti. I particularly want to thank Marc Zissman, Matthew Kercher, and Aradhana Narula-Tam at Lincoln Laboratory for their support, and to acknowledge Louise Ivers and Joe Czarnik for mentoring and inspiration, and, with the rest of the Haiti team, many wonderful conversations. In addition, I want to extend my thanks to Martijn Blansjaar, Charl van der Merwe, and several others at Oxfam GB, from whom I learned much during and after my internship there.

I gratefully acknowledge Sawtooth Software for their generosity in awarding a software grant for this research. My early research was supported by grants from the National Science Foundation and NDSEG, for which I am also very grateful.

I feel very lucky to have passed through the PhD program in great company. Thanks to all my ESD friends and colleagues for their support, in seminars, happy hours, sports, coffee, lunches, and more. To my friends in Boston and elsewhere, thanks for your encouragement and inspiration. Finally, I want to thank my family, especially my husband and my parents, for love and support these many years. Thank you.

Contents

(This page intentionally left blank.)

List of Figures

(This page intentionally left blank.)

List of Tables

(This page intentionally left blank.)

Chapter 1

Introduction

A number of recent disasters have highlighted the importance of logistics in emergency response. Supply chain challenges hampered aid efforts in the South Asian tsunami (CNN, 2005), Hurricane Katrina (Walsh, 2006), and the Haiti earthquake (BBC, 2010). Improvements should be possible, since there are decades of supply chain research to build upon. However, the circumstances of relief work are unique, and many of the models and principles developed for other contexts do not apply directly to the humanitarian setting. Humanitarian logisticians operate under extreme time pressure, in a changing and novel environment, and in many cases cannot access much information about the evolving emergency. Unfortunately, traditional methods of modeling and optimization depend on the availability of data and time. In humanitarian logistics, it seems, the challenging context requires a new breed of tools.

Currently, many aspects of humanitarian logistics are managed by people, using tools like spreadsheets and white boards. There may be very good reasons for relying on this approach, beyond the simple lack of better technological solutions. The challenging context may be better captured by human understanding than by models: people may better see how the environment is changing, perceive the impact of constraints, and guess at missing information. On the other hand, people may make worse decisions without the benefit of decision support tools to handle complex information and search large decision spaces. To better support the humanitarian logistics function, it is important to understand these trade-offs in detail. What are the most challenging aspects of logistics in the humanitarian context, and can they be captured by models? How do people manage these elements of the problem? What are the strengths and weaknesses of these "human approaches" and what we will call the "modeling approach"? In this thesis, I explore these questions, looking to understand how humans solve complex logistics problems in the humanitarian context, assessing the implications of their methods for supply chain performance, and investigating the relative merits of modeling and human approaches to humanitarian logistics.

Humanitarian Transportation Planning

To investigate these general issues, I focus on a specific, analytically tractable problem faced by humanitarian supply chain managers: the planning of aid shipments within an emergency-affected country. In many large-scale, international emergencies, a United Nations-managed coordinating team called the Logistics Cluster provides transport services for other aid organizations. These partner organizations submit cargo movement requests, specifying the origin, destination, type, and amount of cargo to be moved. The Cluster must determine how to use its fleet of trucks and helicopters to deliver the most-needed aid cargo as quickly as possible.

This is a challenging task, which requires logisticians to integrate large amounts of information, create a complex plan involving movement of both vehicles and cargo, and make difficult decisions about prioritizing shipments in light of limited transportation resources. It is also a challenging technical problem for models: optimization algorithms can solve this type of problem, but none are able to solve it rapidly, nor trade off among competing priorities (see, e.g., Crainic, 2000; Wieberneit, 2008; de la Torre et al., 2011). This "humanitarian transportation planning" problem provides a useful setting in which to explore human and modeling approaches, because it provides both analytical tractability and enough critical aspects of the humanitarian context to challenge both models and humans.

The humanitarian transportation planning problem is not only a good case from which to learn about human problem-solving, it is also an extremely important problem in its own right. Transportation decisions determine when and where aid arrives. Poor planning could result in deliveries that do not meet priority needs, backlogs at the arrival port, or inefficient use of limited transportation capacity. Better planning could increase the number of people who receive aid, or ensure that those who need it most receive what they need more quickly. Such improvements could have a major impact on aid delivery worldwide. In 2011 alone, the Logistics Cluster provided services in 12 countries (Logistics Cluster, 2012b). In one operation, after the Haiti earthquake, they transported cargo from 124 different organizations, totaling almost 17,000 metric tons of cargo, including health, water and sanitation, shelter, and food (Logistics Cluster, 2012a).

This research was developed in cooperation with the Logistics Cluster, to develop knowledge and tools to assist in transportation planning. Improving their extensive transportation operations could enhance the ability of many humanitarian aid organizations (those who use their transport services) to reach people affected by emergencies, ensuring that aid is delivered in greater quantities, more quickly, to those who need it most. To further illustrate the challenges of transportation planning and the potential benefits from its improvement, I examine a motivating case: logistics operations in the Pakistan flooding emergency in 2010.

Motivating Case: Pakistan Flooding 2010

Beginning in late July, 2010, and continuing for months, Pakistan experienced some of the worst flooding in the history of the country. Heavy monsoon rains led to flooding in northwest Pakistan, then the floodwaters traveled down the Indus river to the Arabian sea, severely impacting the central and southern parts of the country. Figure 1-1 shows the extent of the flooding in early September. The floods receded relatively quickly in the provinces of Khyber Pakhtunkhwa (KPK) and Balochistan, but lingered for weeks in Punjab and for months in parts of Sindh, where waters still had not receded in some areas in January (Polastro et al., 2011; NDMA, 2011).

The flooding devastated about one-fifth of the country, including 78 districts in 6 provinces, encompassing an area of one hundred thousand square kilometers. The disaster affected 20.2 million people, about 10% of the population of Pakistan (Polastro et al., 2011). Flooding destroyed 1.6 million houses, 23,831 kilometers of roads, and 485 health facilities (NDMA, 2011). More than 12 million people were in need of humanitarian assistance (Polastro et al., 2011). The scale of this emergency, in terms of area and population affected, was larger than the 2004 tsunami and the Haiti earthquake combined.

The main logistics challenges were access to affected areas, storage facilities, and coordination (Logistics Cluster, 2010a). The Logistics Cluster, an inter-agency organization led by the World Food Programme, was quickly launched to provide coordination and fill operational gaps. The Cluster provided transportation services to reach cut-off areas, managed storage facilities throughout the country, and coordinated with other aid organizations and with government and military authorities (Logistics Cluster, 2010a).

Broad-ranging evaluations of the humanitarian response to the 2010-2011 floods in Pakistan reported mixed success (Polastro et al., 2011; NDMA, 2011). Not all needs were met, and moreover the response was not always targeted at those most in need. The response was very fast and effective in KPK province, where many organizations already had operations, but less so in more-affected areas like Punjab and Sindh, prompting the Pakistani government to urge better coordination among aid organizations. Coordination is often a major challenge in an emergency of this scale, and it worked well between militaries and some humanitarian organizations, but there were some difficulties coordinating between many humanitarian organizations and the military (Polastro et al., 2011; NDMA, 2011).

Logistics challenges One of the most important operational challenges was lack of access to affected areas. Many affected areas were cut off when roads were destroyed by the flooding or clogged with people fleeing the destruction (Khan, 2010; British Red Cross, 2010; WFP News, 2010; UN News, 2010). As the United Nations World Food Programme reported, many of the "main roads were under water and the side routes congested with people trying to evacuate" (WFP News, 2010, para. 2). To access these cut-off areas, aid organizations relied on diverse transportation methods, like using donkeys or travelling by foot (Khan, 2010). However, for large-scale deliveries, helicopters were "the only viable means to reach many areas that have been cut off" (UN News, 2010, para. 8), meaning that air transportation was an essential part of the aid supply chain. Aid organizations had to weave together various methods of getting around the access challenges to reach affected areas.

Air transportation was a key bottleneck: it was the lack of sufficient air transport capacity that kept aid from reaching those who needed it. About a month after the start of the flooding, the British Red Cross reported, "There are massive logistical challenges in Pakistan which are holding us back from reaching all the people in need. At the moment there are relief items ready to go but the challenge is getting them to places where roads, rail, and all normal transport have been wiped out" (British Red Cross, 2010, para. 4). Thus, aid cargo was available but could not be delivered because of the difficulty in reaching cut-off areas. As the BBC reported, "Help is coming, but painfully slowly. ... Helicopters are a vital part of the aid operation... though for now there are far too few of them" (BBC

News, 2010, paras. 14, 21). The lack of sufficient helicopter capacity was one of the main reasons for slow delivery of humanitarian relief.

In an attempt to address the air transportation bottleneck, the Cluster managed a huge air operation, using "all available air assets" to reach affected populations, including 9 helicopters from the United Nations Humanitarian Air Service, along with many other aircraft provided by governments and militaries. In six months, the air operation utilized more than 60 aircraft to reach 300 locations, delivering more than 12,200 metric tons of cargo to nearly 1 million beneficiaries (Logistics Cluster, 2011). It was a particularly complex air operation, because as the floodwaters spread into some areas and receded from others, the helicopters had to move within the country to reach newly affected destinations. Figure 1-2 shows the extent of the Cluster-facilitated air operation. Nevertheless, the air transportation capacity was still insufficient to meet the needs (Schlein, 2010; BBC News, 2010).

A lack of sufficient financial resources likely contributed to the logistical bottlenecks, as well as the difficulty of scaling up the wider relief effort. Only a month after the earthquake, WFP was already appealing for funding (UN News, 2010), and the broader appeal of the humanitarian community was only half-funded at the end of 2010 (OCHA, 2010). As a result, the Logistics Cluster director had to spend "lots of time thinking about how to get the most out of finite resources" (WFP News, 2010, para. 7), managing an expensive air operation as efficiently as possible.

Getting the most out of the available aircraft required coordination among at least three sets of stakeholders: the aid organizations with cargo to deliver; the military, government, and UN organizations with available aircraft; and the Pakistani government agencies responsible for directing the response. The Logistics Cluster participated in a Joint Air Coordination Cell, coordinating with all interested stakeholders (Logistics Cluster, 2011) and thereby attempting to efficiently and effectively utilize the air transportation assets available (US Embassy, Islamabad, Pakistan, 2010). Analysis of the response indicated that coordination efforts generally had mixed success, but the Logistics Cluster and WFP were relatively successful in coordinating (NDMA, 2011).

Because of the limited transport capacity available, prioritization of aid deliveries was crucially important. The Joint Air Coordination Cell provided a forum in which to discuss prioritization, which was at least in part implemented by choosing which cargo flew when on the Cluster-managed aircraft. The Cluster received cargo movement requests from other aid organizations, and decided when to deliver each of them (this problem is the focus of this thesis). The Cluster followed Pakistani government directions to prioritize deliveries of kits containing shelter, hygiene, and food items from various agencies, and later delivered shelter cargo to snowed-in areas (Logistics Cluster, 2011). Nevertheless, a later UN analysis lamented a "poor capacity to prioritize" as a challenge in the broader relief effort (OCHA, 2010, "Pakistan," para. 5), and other reports suggested a more general inadequacy of aid targeting (NDMA, 2011; Polastro et al., 2011). Despite steps toward better prioritization through coordination, there is room for improvement.

In summary, the key logistics challenges in responding to the Pakistan floods were a lack of access to flooded areas, insufficient air transportation capacity, insufficient financial resources, coordination difficulties, and prioritization of humanitarian aid deliveries. The Logistics Cluster worked to address many of these challenges on behalf of the humanitarian community or in cooperation with its partner organizations. Nevertheless, many challenges remain.

Importance of transportation planning Better transportation planning could address many of the key logistics challenges identified in this case study. Transportation planning is the process by which the Cluster (or any organization managing a fleet of vehicles) determines where and when each vehicle will move and what cargo it will carry. Currently, the Logistics Cluster does transportation planning by hand, meaning a human planner receives requests for cargo movement and decides where and when vehicles will move, usually with a 3-4 day time horizon. Transportation planning is a challenging problem for people to manage, because it involves searching through a space of many possible vehicle movements and cargo loads, and it is difficult to know the impact of one decision on a later decision or on the final performance metrics for the overall plan. Therefore, there are probably opportunities to improve transportation planning with decision support tools. Better transportation planning would address several important logistics challenges, especially the lack of sufficient transportation capacity and the prioritization of aid deliveries.

First, better transportation planning could identify ways to make limited transportation capacity go farther, by optimizing the use of vehicles to deliver the most possible cargo. In Pakistan, limited transportation capacity was a key bottleneck that slowed the delivery of aid. With better transportation planning, it might have been possible to increase the

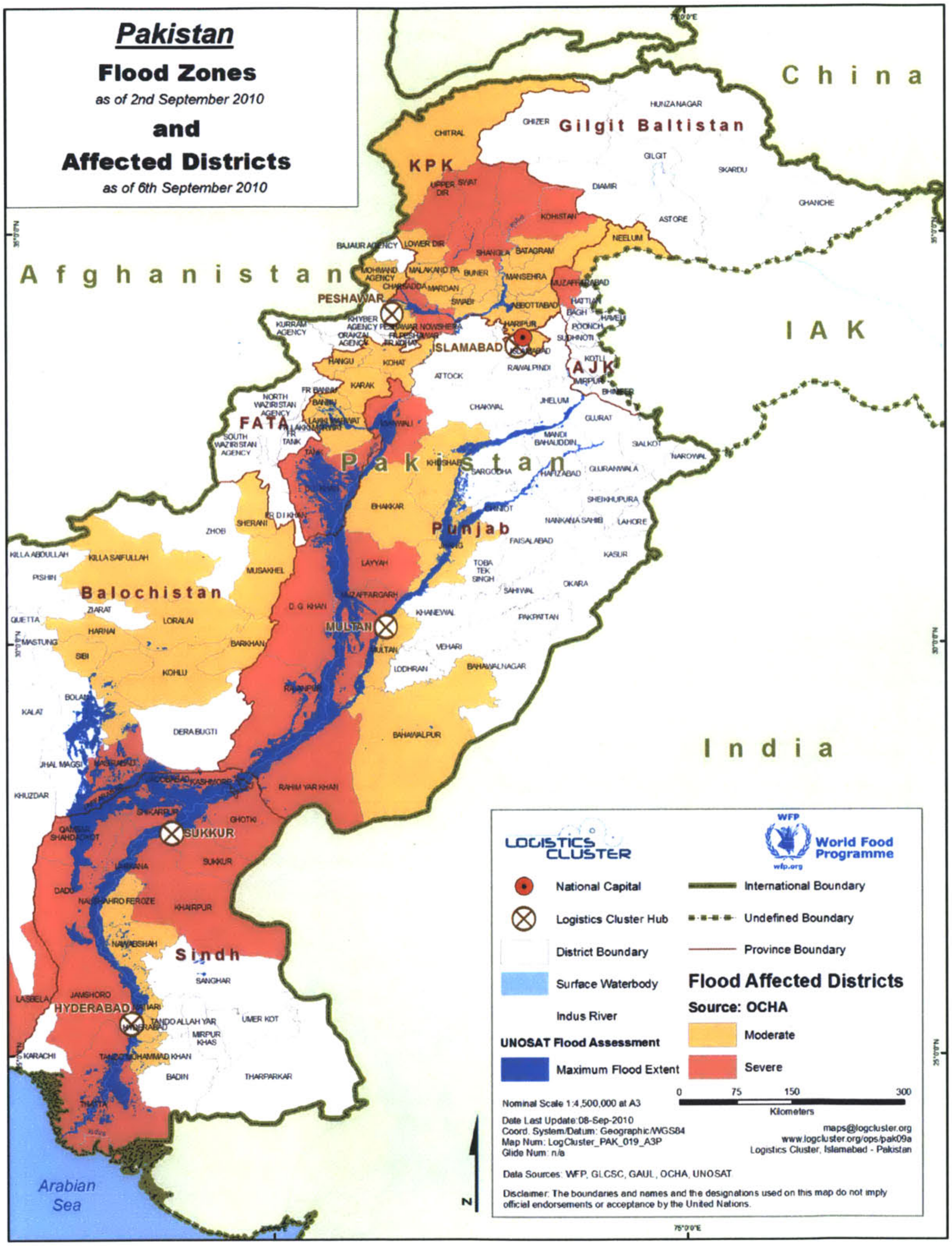


Figure 1-1: Extent of flooding in Pakistan, as of early September 2010. (Logistics Cluster, 2010b)

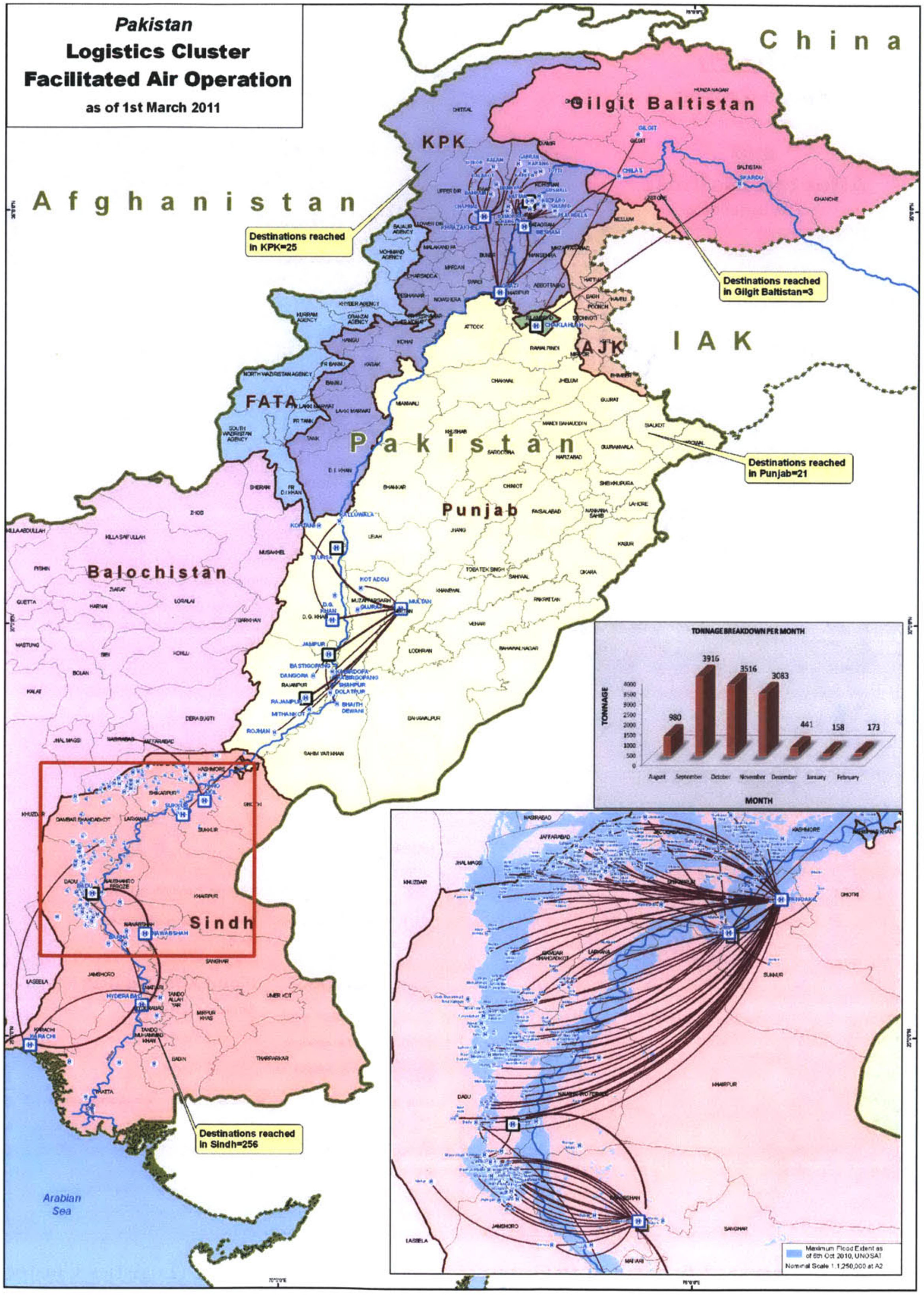


Figure 1-2: Logistics Cluster Air Operations Map (Logistics Cluster, 2011).

capacity of the helicopter fleet without requiring additional financial resources, which were already stretched.

Second, transportation planning is an important way in which prioritization of aid is implemented. Prioritization policies are often handed down from government organizations, as was the case in Pakistan, and it is up to the logisticians to ensure that their plans are in line with government prioritization requests. With careful selection of which cargo to deliver and when, government prioritization policies could be implemented more effectively.

Better transportation planning provides a way to increase transportation capacity and improve prioritization of aid deliveries without drawing on stretched financial resources. In Pakistan, such improvements would have addressed many of the key challenges of the aid effort: the difficulty in reaching cut-off areas because of limited transportation capacity and the lack of effective prioritization of aid deliveries.

Research Questions and Thesis Outline

This thesis is motivated by the need for better humanitarian transportation planning in emergencies like the Pakistan floods. Better planning must take into account the complex challenges of the humanitarian context. In Pakistan, for example, operations had to quickly adapt as floodwaters moved, priorities shifted, affected areas became more or less accessible, and vehicle availability changed. In this dynamic context, good planning approaches may require a combination of human intuition to understand the context and mathematical modeling to handle the complexity. This thesis explores the strengths and weaknesses of human and modeling approaches, to identify ways to improve transportation planning by building on the strengths of each. More specifically, I address the following research questions:

1. How do humanitarian logisticians manage transportation planning, in the emergency response context?

2. How should the performance of humanitarian transportation plans be measured?

3. How do human approaches compare to mathematical modeling approaches? What are the strengths and weaknesses of each?

The thesis is organized into three stand-alone chapters, and each addresses one of the three research questions. With the first research question, I aim to increase our understanding of how humans solve complex operational problems, in realistic and urgent problem settings, by studying the extreme case of humanitarian transportation planning. We know little about how humans solve problems that are both ill-defined and require complex reasoning, challenging humans in both sensemaking and solving (Rudolph and Repenning, 2002). In Chapter 2, I use social science methods to understand how experienced logisticians make sense of the problem and create a transportation plan. I observe logistician teams in a rich and detailed emergency response simulation that replicates many aspects of the humanitarian context. Based on my observations, I identify patterns of decision-making behavior, which show how sensemaking and solving processes are linked in humanitarian transportation planning.

The second research question arose from an interesting finding in the first study. One of the most important elements of understanding the problem was figuring out the right goals or objectives for the transportation plan. Humanitarian logisticians used several criteria to judge their plans (such as the amount of cargo delivered, or the speed of delivery), but it was difficult to determine which were the most important. Weighting these success criteria is important because such weights can be used to develop objective functions for planning tools that account for the goals of humanitarian aid delivery. Chapter 3 aims to quantify the importance of various objectives of transportation planning and develop a performance measure that can evaluate a given plan. I use a conjoint analysis survey of humanitarian logisticians to develop a piecewise linear utility function that describes how experts trade off the multiple objectives of humanitarian aid.

The third research question explores how well people solve transportation planning problems. In Chapter 4, I seek to compare the performance of human problem-solving approaches (discovered in the first part of the research) with that of mathematical models, in order to identify the strengths of each. The simulation field research in Chapter 2 did not enable evaluation of decision-making performance, so in this chapter I develop formal models of the human approaches found in Chapter 2. I compare their performance with that of an optimization model, to evaluate the relative merits of each approach.

Addressing these research questions requires a wide array of methods. I aim to understand both *how* and *how well* humans manage a complex planning problem in a challenging environment. Meeting this goal requires a combination of ethnographic social science methods to address *how* people make plans, conjoint analysis surveys to *measure* their performance criteria, and modeling methods to address *how well* humans plan in comparison with mathematical models. I chose to focus on a single, rich planning problem, and to explore it with a variety of methods. Using multiple methods to understand a single problem provides complementary perspectives, enabling us to identify important elements of human decision-making, then measure and model those elements to draw conclusions about their effectiveness.

In this thesis, I show how experienced logisticians understand the problem of making aid delivery plans, in the messy environment of humanitarian response. I describe how they see the problem, the information they consider in their decision-making, and how their understanding of the problem shapes their decision processes. Then, I build on this knowledge to suggest tools and models to improve aid transportation planning. I identify one key element to which humans pay attention, the goals and objectives of aid delivery plans, and I capture these human goals to create objectives for planning models that fit human goals. I also use the results from the first study, knowledge of how humans see and solve planning problems, to provide the formulation for a set of models. The models provide a simpler environment than the simulation, yet capture certain key elements of the problem considered by human logisticians. In this model world, I explore the effectiveness of human decision-making processes, identifying simple and intuitive planning approaches that are effective and implementable. Thus, the three chapters of this thesis utilize diverse methods but provide complementary insights about human decision-making. Together, they enable the development of better planning approaches for humanitarian aid delivery.

References

BBC (2010). Logistical nightmare hampers aid effort. BBC News, 22 Jan 2010.

BBC News (2010). Pakistan floods: Painfully slow progress of aid effort. BBC News, 23 Aug 2010.

British Red Cross (2010). Pakistan relief effort faces massive challenges. British Red Cross, 26 August 2010.

CNN (2005). Much tsunami clothing aid 'wasted'. CNN, 5 Oct 2005.

Crainic, T. G. (2000). Service network design in freight transportation. *European Journal of Operational Research*, 122:272–288.

de la Torre, L., Dolinskaya, I. S., and Smilowitz, K. (2011). Disaster relief routing: Integrating research and practice. *Socio-Economic Planning Sciences*, (in press).

Holguin-Veras, J., Perez, N., Jaller, M., Destro, L., and Wachtendorf, T. (2010). On the need to reformulate humanitarian logistics modeling: Deprivation costs, and material convergence (draft version). In *Proceedings of the 89th Transportation Research Board Annual Meeting*.

Khan, J. (2010). Landslides complicate pakistan flood relief efforts. Reuters, 9 Aug 2010.

Logistics Cluster (2010a). Pakistan flood response overview as of 20 december 2010. Technical report, Logistics Cluster.

Logistics Cluster (2010b). Pakistan flood zones and affected districts. Technical report, Logistics Cluster Pakistan.

Logistics Cluster (2011). Logistics cluster pakistan air operations overview: Monsoon flood emergency response. Technical report, Logistics Cluster.

Logistics Cluster (2012a). Logistics cluster haiti activities and achievements. Technical report, Logistics Cluster.

Logistics Cluster (2012b). Overview of logistics cluster operations 2010-2011. http://www.logcluster.org/about/logistics-cluster/background-information/overview-of-logistics-cluster-operations-2010-2011.

NDMA (2011). Pakistan floods 2010: Learning from experience. Technical report, Pakistan National Disaster Management Authority.

OCHA (2010). Ocha annual report 2010. Technical report, United Nations Office for the Coordination of Humanitarian Aid.

Polastro, R., Nagrah, A., Steen, N., and Zafar, F. (2011). Inter-agency real time evaluation of the humanitarian response to pakistan's 2010 flood crisis. Technical report, DARA.

Rudolph, J. W. and Repenning, N. P. (2002). Disaster dynamics: Understanding the role of quantity in organizational collapse. *Administrative Science Quarterly*, 47(1):1–30.

Schlein, L. (2010). Helicopters needed to transport food in pakistan. VOA News, 20 Aug 2010.

UN News (2010). Un scales up aid to flood-hit pakistan despite huge logistical challenges. UN News Service, 17 Aug 2010.

US Embassy, Islamabad, Pakistan (2010). Joint air coordination cell expedites international aid to flood victims. US Embassy Islamabad, 28 Sep 2010.

Walsh, B. (2006). Fema says katrina taught it a lesson. Times-Picayune, 13 May 2006.

WFP News (2010). Pakistan: Wfp logistics arm lends a hand to aid community. WFP News, 14 Sep 2010.

Wieberneit, N. (2008). Service network design for freight transportation: a review. *OR Spectrum*, 30:77–112.

(This page intentionally left blank.)

Chapter 2

Humanitarian response transportation planning: a behavioral study of problem-solving and sensemaking

2.1 Introduction

Logistical challenges in several recent disasters, including Hurricane Katrina (Walsh, 2006) and the Haiti earthquake (BBC, 2010), have sparked efforts to improve humanitarian supply chain management. Improvement efforts often attempt to build upon experience in commercial logistics, adapting modeling techniques developed for the private sector for use in humanitarian logistics. However, the humanitarian context is particularly challenging to capture in mathematical models, because it is dynamic, information-poor, and exhibits multiple competing objectives. Human decision-makers may be better able to make sense of the complex context, but may make worse decisions without the benefit of tools to manage information and search for good solutions. Understanding how humanitarian logisticians make transportation planning decisions might both suggest better ways to model the challenging context and also identify weaknesses that could be improved with training or decision support tools. To that end, this chapter investigates human decision-making processes in humanitarian emergency response.

Humanitarian logistics spans many activities, including assessing infrastructure, estimating needs, bringing in warehouses and vehicles, procuring and managing inventory, and transporting aid to those who need it. This study focuses on the latter problem, on planning the delivery of aid to affected communities. After cargo reaches a disaster-affected region, it must be transported over the "last mile" to reach its final destination. Transportation capacity is often insufficient to satisfy the need for cargo movement, so aid may be delayed in reaching those who need it. Improved transportation planning could enable better use of scarce transportation resources, so that more aid could reach beneficiaries faster. In addition, transportation decisions – such as which cargo to load and where to send vehicles – have a major impact on who receives aid and what kind of aid they receive. Better transportation planning, therefore, offers an opportunity to improve the operational prioritization of humanitarian aid, in addition to the amount and speed of its delivery.

Aid delivery planning was chosen as the focus for this study because it is important in emergency response and complex enough to challenge both modeling and human approaches. Related transportation problems have long been studied in the commercial context (e.g. Kim et al., 1999; Grünert and Sebastian, 2000), and recently have been examined in the humanitarian context specifically (for reviews, see de la Torre et al., 2011; Caunhye et al., 2012). While such modeling approaches are very good at finding route plans in a complex network flow problem, they struggle to accommodate some essential elements of the humanitarian context. After an emergency, there may not be enough information to build a useful model, as organizations try to determine the needs of scattered beneficiaries and the state of infrastructure. The goals are not clear – who should be served first and what should be sent? Finally, the situation is dynamic: new information arrives, roads wash out, or priorities change. It might be possible to incorporate some of these elements into mathematical models, but in many cases the "right" formulation is unclear, suggesting the need for a combination of human and modeling approaches.

Human decision-makers may have some advantages over models, especially in understanding the complexities of the humanitarian setting. Experienced aid workers may be able to estimate missing information, make assumptions or form expectations about infrastructure, and interpret the context to determine goals. Research on human problem-solving in other urgent contexts suggests that this sort of "sensemaking" activity is essential, and that human experts are often (but not always) good at it (Klein, 1993; Klein et al., 2007;

Weick, 1988, 1993). On the other hand, humans are typically not as good at managing complex information and searching large decision spaces for the best solutions. In making decisions under uncertainty, humans rely on shortcuts or heuristics that yield suboptimal solutions (Simon, 1956; Tversky and Kahneman, 1974). Humanitarian transportation planning requires both making sense of the context (akin to formulating) and solving a complex technical problem. Therefore, improvements may require a combination of human and modeling approaches. Little is known, however, about how people make sense of these problems and how well they solve them. In this chapter, I explore how humans solve humanitarian transportation planning problems, in order to untangle their sensemaking and solving processes and illuminate the strengths and weaknesses of their problem-solving approaches.

This research has two related goals. The first is a practical goal: identifying ways to improve the practice of humanitarian transportation planning by combining the strengths of human and modeling approaches. To that end, this chapter seeks to precisely describe how humans solve humanitarian transportation planning problems and to identify specific strengths and weaknesses. The second goal goes beyond the specific context of humanitarian transportation planning, seeking to broaden our understanding of how humans solve urgent problems that require serious effort in both sensemaking and solving. Humanitarian transportation planning is an extreme case of such a problem, so its understanding should shed light on the interactions between sensemaking and solving in ill-structured, complex problems.

2.2 Theoretical Framework

Researchers have long been interested in how humans solve problems and make decisions. Decades of laboratory research has shown that people use shortcuts to make decisions. Field research, on the other hand, shows how experts make decisions in certain ill-defined problem settings, suggesting that defining or making sense of the problem is more important than solving it. There is much less research on problems that combine the challenges of defining problems and solving them, but a few studies suggest that sensemaking and solving are intertwined. Understanding problem-solving in realistic settings requires a better understanding of how sensemaking and solving interact. In the following paragraphs, I review relevant findings in these areas, focusing on problem-solving in general and on problems most relevant to humanitarian transportation planning, such as those in the industrial operations domain.

Decades of research on human problem-solving and decision-making shows that people use shortcuts to solve problems and make decisions, in comparison to so-called "rational" mathematical models. Human problem-solving is often described as search through a problem space of possible solutions, a model most famously derived from empirical studies of human problem solvers by Newell and Simon (1972). Given their cognitive limitations, it is thought that humans employ limited search, rather than searching until an optimal solution is achieved (Simon, 1955). In problems of choice under uncertainty, people use heuristic reasoning shortcuts that create biases compared with rational models (Tversky and Kahneman, 1974; Kahneman et al., 1982), though these biases may be considered rational themselves (Gigerenzer et al., 1999).

The same kinds of shortcuts have been observed in studies of people solving industrial operations management problems, in the same domain as the transportation planning problem of interest in this research. This work on behavioral operations has focused on a small set of problems, including scheduling in manufacturing plants (Crawford and Wiers, 2001) and inventory management (Bendoly et al., 2006). Laboratory research confirms that people use the same sorts of non-rational heuristics as they do in other problems of judgment under uncertainty, when solving newsvendor problems (Schweitzer and Cachon, 2000; Ben-Zion et al., 2008; Bolton and Katok, 2008), multi-period inventory problems (Sterman, 1989; Croson and Donohue, 2002), and supply contracting problems (Kalkanci et al., 2011). People rely on anchoring and adjustment heuristics in newsvendor and multi-period inventory problems (Schweitzer and Cachon, 2000; Sterman, 1989), show evidence of statistical biases (Bolton and Katok, 2008; Ben-Zion et al., 2008), and generally rely on simple rules in complex problems (Kalkanci et al., 2011; Moxnes, 1998). These problems are in the same domain but not of the same type as the transportation planning problem of interest in this research. Studies of humans solving traveling salesman problems and vehicle routing problems are more relevant, because they have the same kind of solutions and require search through a large problem space. Humans are surprisingly good at solving traveling salesman problems when they are represented visually: they find solutions less than 15% worse than optimal with performance degrading only linearly with increasing problem size. Various heuristic mechanisms have been proposed to explain this performance, including clustering, avoiding crossed arcs, and utilizing the convex hull (MacGregor and Chu, 2011). In general, then, a wealth of laboratory research suggests that humans use simplified heuristics or shortcuts in solving industrial operations problems, just as they do in making other types of decisions.

The research discussed in the above paragraphs focuses on what we might call the "solving" of problems: going from a formulation, with all relevant information and a description of what is to be done, to a solution to the problem. However, in real-world situations, it is usually necessary to figure out what the problem is, which information is relevant, and what needs to be done. This process of going from an ill-defined problem to an understanding of it could simply be called "formulating", but here we label it "sensemaking" because of its resemblance to sensemaking processes observed in similar contexts.

Sensemaking is the process by which people understand situations. They perceive stimuli or cues, and these cues are placed into a frame that helps to interpret the cues (Klein et al., 2007; Weick, 1995). The frame, in turn, influences what cues are perceived and how they are interpreted. The frame also generates expectations and directs action. Action generates additional cues. Actions are interpreted retrospectively, and this new understanding may revise the frame. Thus, sensemaking is about perceiving and interpreting stimuli *and* creating the actions and stimuli that are interpreted. Many varied aspects of sensemaking have been described by theorists (see, e.g., Weick et al., 2005; Weick, 1995). Most relevant here are the social aspects of sensemaking. While sensemaking may seem to be an individual process, it often occurs through communication in social settings and the "locus" may be in a group rather than an individual (Weick et al., 2005). Sensemaking has long been considered an important component of how people and organizations respond to emergencies (Maitlis and Sonenshein, 2010), with implications that human actions can fix or exacerbate crises (Weick, 1988) and that failed sensemaking can have disastrous consequences (Weick, 1993). More recent studies use the sensemaking perspective to understand how responders effectively manage information and understand the evolving situation (Landgren, 2005; Landgren and Nulden, 2007; Muhren and Van de Walle, 2009). While the literature provides a reasonable understanding of sensemaking and its use in crisis response, there is little attempt to connect sensemaking to decision-making. Indeed, the sensemaking perspective assumes that once a situation is understood, decisions or actions are taken, but the process of going from understanding to decision is not well articulated.

Research on experts making decisions in urgent settings suggests that formulating or sensemaking, rather than solving, drives the decision-making process. Firefighters recognize a situation as typical and match it to an appropriate action, often implementing the first option they consider (Klein, 1998, 1993). In a wide variety of contexts, including firefighting, military operations, design engineering, airline piloting, and speed chess, experts appear to generate only a single option based on their assessment of the situation, evaluate it for appropriateness, and generate a second option only if the first is unsuitable (Lipshitz et al., 2001). Similarly, in crisis response, a study of 15 diverse decision-making processes suggests that options are evaluated successively rather than concurrently (Hale et al., 2006). Doctors train to avoid fixating on a single erroneous diagnosis (Rudolph et al., 2009). This propensity for matching a single action to a situation is not universal; in fact, multiple options may be implemented simultaneously to hedge against uncertainty (Smith and Dowell, 2000). However, it does appear that in a wide variety of contexts, experts faced with urgent decisions assess the situation and match it to an appropriate action. In these cases, the process of "sensemaking" drives the decision-making process, while "solving" consists simply of retrieving a seemingly appropriate action once the situation is understood.

Most of this literature studies problems in which sensemaking is the only challenge. Sensemaking research stops short of studying the decision processes that come after it, and research has focused mainly on environments in which an understanding of the situation is enough to generate an appropriate decision (see Lipshitz et al., 2001). Because sensemaking and solving have been studied in distinct literatures, very few studies have explicitly focused on how these processes work together (Rudolph et al., 2009).

Transportation planning in humanitarian aid requires serious effort in both sensemaking and solving, as does problem-solving in many other contexts. Empirical studies shed some light on how sensemaking and solving interact. Rudolph et al. (2009) advance a conceptual model that describes the relationships between sensemaking and solving in the medical diagnosis problem setting. Actions generate information about the patient's condition, the information is interpreted either in support of a leading diagnosis or an alternative diagnosis (and may induce a change in the leading diagnosis), and new action is generated to further illuminate the plausibility of the diagnoses. Thus, sensemaking guides action intended to gather information to distinguish among two choices (solving), and these choices suggest further action to gather additional information (sensemaking), and so on. This model is specific to choice-oriented problems, and therefore does not apply directly to the search-oriented transportation planning problem. Nevertheless, it suggests that sensemaking and solving are intimately intertwined and must be understood together, by studying a complex problem in a realistic setting.

There are empirical studies of problem-solving in industrial operations, which may be more relevant to transportation planning. Field research in this setting has focused rather narrowly on production planning and scheduling. Observations of planners and schedulers suggest that people predict or monitor potential problems by collecting information (Vernon, 2001; Jackson et al., 2004; McKay et al., 1995; Webster, 2001), and that problems are dealt with using heuristics or rules (McKay et al., 1995; Webster, 2001; McKay and Wiers, 2001; Grant, 1986). These studies do not explicitly consider sensemaking or solving processes, but monitoring suggests an ongoing sensemaking process, while heuristics and rules resemble the solving processes found in behavioral operations laboratory studies. Most of these studies stopped short of generating analytically generalizable theory of human problem-solving, focusing instead on the types of activities performed and how to support them. Nevertheless, sensemaking and solving appear to be intertwined in planning and scheduling, though exactly how they interact is not clear.

To understand how people solve complex problems in ill-defined, realistic settings, it is not enough to focus on the solving or the sensemaking processes separately. This research looks to elaborate the intertwined processes of sensemaking and solving in an urgent, complex, and ill-defined problem. Humans solving pre-formulated (laboratory) operations problems tend to rely on simple heuristics or rules. Humans making decisions in ill-defined, urgent contexts rely on sensemaking to generate appropriate actions. These problems have been studied separately, leaving it unclear how sensemaking and solving interact. Elaborating these processes in a particularly ill-defined, urgent and complex problem-solving setting should shed light on how humans make decisions more generally in this important class of real-world problems.

2.3 Research Design and Methods

This study describes how people solve transportation planning problems in the humanitarian context. As discussed above, little is known about how humans solve real-world problems that require both sensemaking and solving. In such situations, when existing theory is inadequate to describe a phenomenon, qualitative case studies can enable the generation of new theory (Eisenhardt, 1989).

A qualitative study is ideal for several reasons, beyond the lack of existing theory. There are two key elements of the research question that drive this decision. First, the setting is important. We are interested in how planning is achieved under the conditions present in humanitarian emergencies: time pressure, high stakes, information scarcity, rapid uncertain changes, and unclear objectives. Without these conditions, the methods of planning might be different, so it is essential that the research take place in this context. It would be difficult to replicate this set of conditions in a laboratory, but qualitative methods allow us to capture action in context (Pettigrew, 1990; Yin, 2009). Second, we are interested in *how* planning is managed by humanitarian logisticians. We cannot understand how a plan was created by looking only at the plan itself. We require a method that enables us to "see" the behavior and ways of thinking that led to the generation of the plan. Qualitative methods can capture rich data to describe how the action occurs (Langley, 1999; Pettigrew, 1990; Woods, 1993). In short, qualitative methods enable us to explore human approaches to transportation planning, and to build new theories that describe their ways of approaching the problem.

In focusing on *how* logisticians manage transportation planning, this study seeks to build a process theory of the phenomenon. Mohr (1982) distinguishes between process theories and variance theories: variance theories explain phenomena based on relationships among variables (e.g. more of an independent variable leads to more of a dependent variable), whereas process theories explain outcomes based on sequences of events, activities, and choices (Langley, 1999). Process theorizing is appropriate here, because we wish to understand the process by which humanitarian logisticians create transportation plans. The process orientation drives both the kind of data to be collected and the methods of analysis, discussed in more detail below. In short, however, we seek patterns of behavior that represent logisticians' approaches to transportation planning.

Before describing the research design, data collection, and data analysis methods in more detail, we describe the setting in which the research was conducted.

2.3.1 Research Setting

The research setting is a simulated emergency response, conducted for the purpose of training logisticians to work with the Logistics Cluster. The Logistics Cluster is an organization that coordinates logistics efforts among the aid agencies responding to an emergency. It is part of a system of coordination organizations – including a Shelter Cluster, Health Cluster, etc. – that focus on coordinating each component of a humanitarian response. The Cluster system is managed by an inter-organizational committee of humanitarian organizations, and in a given response each of these organizations typically participates in meetings and utilizes services of relevant Clusters. The Logistics Cluster provides coordination, information sharing, and sometimes transportation services to the humanitarian community. The World Food Programme (WFP) is the lead organization, but the Logistics Cluster is usually run by a team of logisticians who are seconded by various aid agencies. The simulated emergency response, the setting for this research, is run by WFP to train logisticians from other agencies to work with the Logistics Cluster. As such, it brings together teams of experienced logisticians to create a logistics-focused response to a major emergency, and it is repeated periodically, with virtually identical scenarios, to train multiple teams.

Each training is a week-long simulated emergency, run 24 hours a day for 6 days. The participants live and work in prefabricated housing and tents, with satellite-based internet and limited electronic resources; these are the same facilities that WFP deploys to real emergencies. The trainees, or "participants", are divided into 2 teams of approximately 10 people each. They are tasked as Logistics Response Teams, the group that arrives immediately after a disaster to set up a Logistics Cluster. The two teams are independently responding to the same emergency (and are instructed to ignore the existence of the other team). A team of facilitators are also present to play the roles of other actors in the emergency and to manage the logistics of the training. The scenario is defined ahead of time, and is the same for each team and each training (minor variations will be described in the next section). Each facilitator has talking points and information to discuss in each meeting, email and document "injections" are already prepared to send to the teams, and an extensive script describes how the scenario unfolds over the course of the week.

The teams are given only news releases before their arrival, saying that a large earthquake has overwhelmed a fictional developing country called Snowland (this and other names have been disguised throughout the thesis). They spend most of the first two days in meetings with representatives of other humanitarian organizations (played by facilitators), in which they are given more, often conflicting, information about the situation they face: how many are affected, which areas are worst-hit, what the population needs, what other organizations are doing, what roads and air facilities are available, what transportation and storage assets can be found, the customs situation, etc. By the third day, they must create a logistics plan to address the situation. This "operational plan" should include the roads and hubs they plan to use, temporary warehousing and vehicles they need to bring in, and the services they will provide to the humanitarian community. On the fourth day, they are asked to use their operational plan to execute (virtually) a logistics operation. They are given a large set of cargo movement requests from other organizations, and asked to create a 7-day plan for using their fleet of trucks and helicopters to deliver these requests. On the following day, the fifth day of the simulation, they must present their plan to donors and to the humanitarian community. Throughout the simulation, they are given changing and conflicting information about the situation, and they face challenges in coordinating their plans with the rest of the humanitarian community. They are put under extreme pressure by those they work with (played by the facilitators), and rarely get more than three or four hours of sleep each night. Thus, the context is a fair representation of the urgent, information-poor environment of humanitarian logistics. Data are collected throughout the simulation, but this study focuses on the data from the fourth and fifth day of the simulation, in which (day four) teams are asked to create a 7-day transportation plan and (day five) present their accomplishments to donors and the humanitarian community.

I chose to study the simulation rather than real emergencies for three reasons. First, because the simulation is repeated with the same scenario for multiple teams, it enables a powerful multiple case research design. Second, it is intrusive and ethically difficult to (merely) observe in a real emergency. All available personnel should be working to assist those in need, and an observer would not only fail to help but also get in the way. The simulation provides the opportunity to observe the teams with minimal disruption to their work. Third, the simulation is realistic enough to spark the same ways of thinking that logisticians would employ in a real emergency. The participants, all of whom have worked in real emergencies, support the idea that the simulation is realistic. According to WFP (2009, para. 3), "participants reported that the training had replicated the feel of a real emergency". One participant wrote that, although it was a "simulation exercise, it feels like we've been through the wringer and back" (Freeman, 2010, para. 6). Thus, the simulation comes close to replicating the challenging humanitarian context. The teams sent to a real emergency typically have fewer than 10 members, and they usually plan 1-4 days ahead, though they aim to plan farther ahead when possible. The simulation may be more difficult than transportation planning in reality (requiring coordination among more planners to create a longer plan), but captures the same kinds of problems and challenges faced by real Logistics Response Teams.

While the simulation is very realistic in many ways, it does not probe all the dynamics of real emergencies. Two important dynamics are not emphasized within the transportation planning portion of the simulation: uncertainty and changes. In the early days of response to most emergencies, there is a lot of uncertainty about the needs, state of infrastructure, and other relevant details. As planners learn more about the situation, and as they gather feedback over time, their understanding of this information will change. In the context of the transportation planning problem, this could mean changes in the cargo movement requests, in the priority of locations or types of aid items, or in the availability of vehicles or roads. The simulation only included one day of transportation planning, and in this limited timeframe only one small change was made to the problem: additional cargo movement requests were made in the middle of the day. Anecdotally, based on observations of this minor change in the simulation and on conversations with Logistics Cluster personnel about past operations, planners seem to deal with changing information by re-planning to take into account the new situation. However, the simulation did not prompt enough discussion of change and uncertainty to draw further conclusions. On the other hand, despite the limited changes to the transportation planning scenario, teams were certainly considering the situation uncertain and likely to change, based on their experiences in the first three days of the simulated response. In those three days, they had received changing and conflicting information, and needed to update their plans around it. Thus, while the transportation planning problem on the fourth day included few changes, teams were already primed to consider the dynamic situation as they made transportation planning decisions.

2.3.2 Research Design

The repetition of the training described above enables a powerful multiple case study research design. Each case study is analogous to an entire experiment, rather than a sample in an experiment (Yin, 2009). Theory arising from one case is confirmed or disconfirmed by the other cases. Because the theory ultimately has been "tested" a number of times, more confidence can be placed in it.

This study is designed to develop a process theory of how humanitarian logisticians solve the transportation planning problem. Each case therefore represents an instance of such a process: a single team of participants navigating days four and five of the logistics training described above. The cases are chosen to be "literal replications" of one another (Yin, 2009), meaning that they are chosen to be similar and the same kinds of results are expected in each case (as opposed to "theoretical replications", chosen to vary along some dimension, in which the same theory is expected to explain different results).

There is no law dictating the number of case studies required to generate a theory. In general, in theory-building research, one concludes there is "enough" data when a point of theoretical saturation is reached, meaning the incremental improvement in theory is small because any new data confirms existing suppositions (Glaser and Strauss, 1967). In case study research, the same principle applies (Eisenhardt, 1989). It is difficult to know, ahead of time, the number of cases required to reach theoretical saturation. Eisenhardt (1989) recommends, as a guideline, selecting between 4 and 10 cases: fewer than 4 cases is less convincingly grounded in the data, and the data from more than 10 cases becomes difficult to handle.

I studied five trainings, each with two (independent) teams. The unit of analysis is the Logistics Response Team described above: a group of approximately ten experienced logisticians who are sent to manage logistics in the simulated emergency. It makes sense to define the team as the unit of analysis because such teams are, in reality, the main decision-makers in emergency response (though they are usually smaller teams). Moreover, observations of the interactions between the team members will shed light on their problem-solving process without the need for more disruptive techniques like thinking aloud (Woods, 1993).

With each team as the unit of analysis, there are two options for defining the multiple case study design. Each training could be considered a case, with two embedded units of analysis. Alternatively, each team's experience could be considered a single case. The former design would be more appropriate if the trainings differed significantly; in that case, it would make sense to consider each training separately and compare how the two teams within that case responded to the particular set of circumstances in that training. However, in this case, each training is very similar to the others. While the participants and facilitators are different, the scenario is roughly the same. An extensive script governs the information provided to the participants at any given meeting or interaction. The script is adapted to the needs and decisions of the teams, but the set of possibilities is restricted by the training scenario. Moreover, every communication with the teams is approved by the training coordinator, who designed the scenario and supervises all the trainings. As a result, each of the ten teams observed (in five trainings) had roughly the same experience. Therefore, it makes more sense to consider each team a separate case. Thus, there are ten cases, each consisting of one team solving a transportation planning problem in the same emergency scenario.

The ten cases do nevertheless vary in some ways. Most importantly, in the fourth and fifth trainings (the last four cases), I introduced a tool for transportation planning. Previously, the teams had been given no support beyond the data for the problem. In cooperation with the training organizers, I developed a simple Excel-based tool, which provided a template for writing down the movements of trucks and cargo, and additionally verified the feasibility of the plans and calculated some basic performance metrics about the plan. This tool was provided to both teams in the final two trainings, but it was only used by two of the four teams. The key difference in these four cases is that the teams were provided with (1) a structured template for recording the plan, and (2) the ability to check its feasibility and its performance (how much was delivered, of what cargo). I expected that the addition of such a tool would reduce the amount of time teams spent defining data formats, and would lead them to think more carefully about the performance of a plan, thus illuminating the aspects of problem-solving in which I was most interested. The second, more minor, source of variation within the cases was that the specific problem data was different in each training. The number of cargo movement requests was generally reduced, and the number of vehicles increased, to make it easier to deliver a significant portion of the requested cargo (mainly to increase the morale of the team). The problem remained overwhelmingly difficult to solve in the time provided, so these changes should not have made much difference to the teams' problem-solving behavior.

Even though I observed teams, the team dynamics were not the focus of this study. The team setting was a convenient way to ensure that participants verbalize and explain their thought processes as they work (which will be discussed in more detail later). Of course, with teams of ten people under pressure, team dynamics were important, but they appeared to impact problem-solving work only when problems were very severe, perhaps in one or two of the ten teams observed. Such conflicts were generally either resolved, occasionally with help from facilitators, or other members of the team picked up more of the problem-solving work. I focused on capturing the problem-solving processes rather than the interactions between team members.

It would be useful to link problem-solving processes with performance on the transportation planning process. Unfortunately, the research setting did not enable linking process and performance, for two reasons. First, it was practically impossible to measure the success of a given team's transportation plan, because there was no requirement to "turn in" a completed plan, and collected data was insufficient to judge performance. Second, it would be difficult to determine which problem-solving processes contributed to the team's performance. There were many other factors that contributed to success in planning, such as leadership and teamwork. In addition, most teams used a mix of problem-solving processes, so it would be difficult to link a team's performance to the performance of a given type of process. Because it was impossible to analyze the performance of problem-solving processes in this research setting, I chose to analyze performance using models of human problem-solving processes, as described in Chapter 4 of this thesis.

As in experimental research, threats to validity are considered within the research design. Internal validity requires that the proposed explanation (our theory) actually explains the effect, rather than some other plausible explanation. In experiments, there are a number of clear threats to internal validity (Campbell and Stanley, 1966), and these are generally addressed through careful research design. In case studies, as opposed to experiments, the threats are less clearly delineated, and we concentrate instead on the key underlying point: the ruling out of rival explanations. In case study research, internal validity is addressed by taking care to consider and rule out rival hypotheses during data analysis (Yin, 2009).

External validity deals with questions of generalizability: to what situations beyond the case study itself are the results generalizable? Generalizability of any experimental or case study result requires an assumption that the same kinds of laws seen in one setting – the experiment or case study – work similarly in another setting (Campbell and Stanley, 1966). In quantitative research, samples are chosen to ensure "statistical" generalizability to a specific population, but the theory resulting from such experiments is often considered to apply more broadly, to other similar populations. This is termed "analytic generalization" by Yin (2009). He argues that case studies enable analytic generalization rather than statistical generalization: the theory arising from specific case studies can be applied across other similar settings.

Addressing external validity, then, requires selection of cases that will enable the generation of theory that applies beyond the specific instance studied. In multiple case studies, the replication logic – showing that the same theory applies in multiple independent cases – confirms that the theory is at least generalizable to all the cases included in the study, and suggests that it is probably generalizable to similar cases beyond those considered in the research. In this study, I argue that the theory arising from ten diverse teams solving a transportation planning problem in one (simulated) humanitarian scenario (designed, for training purposes, to be representative of other major emergencies) generalizes to humanitarian logisticians solving similar problems in similar emergency settings, and may apply to experienced planners solving similar problems in other urgent, information-poor environments.

2.3.3 Data Collection

The primary method of data collection was field observation, which could also be termed participant-observation, or ethnography. Documents and emails from the simulation were also collected.

Selection of Data Collection Methods

The nature of the phenomenon of interest dictated the choice of data collection method. Problem-solving processes are at least partly realized in the hidden cognitive processes of each participant. The data collection method had to enable some understanding of these cognitive processes. Two clear options were interviews and field observation.

Interviews offer the advantage of more direct access to participants' ways of thinking:

participants can be asked to describe how they made choices and decisions. However, interviews would have required participants to recall their thought processes retrospectively. Retrospective recollections can be unreliable, because subjects may over-simplify, rationalize decisions, or simply not remember what they were thinking at the time (Golden, 1992; Miller et al., 1997). While such concerns do not preclude the use of retrospective interview data, they seemed likely to be especially severe in the fast-moving context of emergency response. Due to stress and lack of sleep, participants might be especially susceptible to the problem of inaccurate recall. In addition, this study focuses on intuitive shortcuts and heuristics; participants might be especially tempted to rationalize their use of these shortcuts. For these reasons, field observation was chosen as the primary method of data collection.

Field observation enabled data collection in real time, with confidence that problems of recall and rationalization would not obscure participants' thought processes. However, the problem remained that much of the problem-solving work is hidden in the minds of participants. Various techniques have been developed to "externalize" such internal processes (see Woods, 1993, for a review), including asking participants to verbalize their thought processes (either afterwards or in real time), withholding information to see when and how participants use data, and making the task cooperative so that participants discuss their ideas with each other (e.g. Miyake, 1986). These techniques vary in the extent to which they disrupt the normal thought processes of subjects and twist the study scenario for the purposes of gathering data. As argued earlier, it was essential for this study that the problem-solving context remain as authentic as possible; therefore, we leaned toward those techniques with minimal disruption. The training scenario offered a natural opportunity to use the cooperative technique: teams of participants solve the transportation planning problem cooperatively, so they verbalize their thought processes to a certain extent. Through observations of their behavior as they interacted with each other, with facilitators, with maps, and with computers and other aspects of the environment, it is possible to make inferences about their problem-solving processes.

The selection of field observation as the primary data collection method and the logistics training as the research setting prohibited interviews as a further check on our inferences about participants' thought processes. The immersive environment of the training made it impossible to conduct interviews during the simulation. The trade-off was considered reasonable, given concerns about interviews and the advantages of the research setting.

Instead of interviews, I attempted to accomplish the same purpose – asking participants to describe their thought processes – within the training scenario. As part of their meetings on the fifth day of the simulation, participants were asked to describe their "methodology" for transportation planning, meaning the way they went about making the transportation plan. Their descriptions provided a second data point to verify and triangulate at least some of the inferences about problem-solving thought processes.

Documents and emails produced during the simulation were also collected. These documents served three related purposes. First, they provided another data point with which to verify evidence from other sources (Yin, 2009). Second, they provided evidence of the teams' starting points for problem solving, as well as the results of the problem-solving and decision-making processes. Third, and more broadly, they filled in the story: they show what information the teams had, believed, and sought; what decisions they made and how they were justified; and more subtle elements like what they considered important.

Observation

I spent the full week of the simulation on-site, as a training facilitator. I played the role of an aid worker for Oxfam, another aid organization. In that role, I had two meetings with each of the teams during the first two days, and I participated in the nightly coordination meetings hosted by the teams, in which I was acting as their customer, asking what services they would provide and how quickly they could help me solve my logistical problems. I also had a secondary role in the development of the simulation. The organizers asked me to help them with the transportation planning portion of the training (the fourth day). I helped to introduce the problem, adapt the exercise materials to each training, and ultimately developed tools and techniques to improve the training. Most of these tools and techniques were minor and were 'inserted' after the conclusion of the simulation: I debriefed participants and developed materials for them to take home with them. Therefore they had no effect on the course of the simulation itself. In the last two trainings that I observed, however, I introduced an Excel-based tool to capture plans; this was described earlier. The teams chose to utilize this tool in only one of the two trainings. During this training, I spent a little more time providing technical support to the teams, showing them how to use the tool and solving the one or two problems that arose. Through both of these roles – as Oxfam and managing the transportation planning exercise – I was both participant and observer.

However, it was generally clear to the teams when I was playing each role, because during observation I wore a red ribbon which clearly marked me as observer.

The teams are told at the start of the training that anyone wearing a red ribbon is "invisible", and should be ignored. It was common practice for one or two facilitators to visit the teams' offices, wearing a red ribbon, to observe their progress or their meetings. I wore a red ribbon and sat in as inconspicuous a place as possible. Meetings and discussions could occur anywhere: in the prefabricated offices, in the large meeting tent, outside around a table, in the kitchen area, during a smoking break, or even in the back of a Land Rover when the wind was very strong. I tried to find a corner in which I would be less likely to be noticed, preferably seated (though more often standing), and I took notes in a small notebook. At first, the teams would be somewhat distracted by my presence and my note-taking, but the distraction appeared to lessen over the course of the week.

My goal in taking notes or "jottings" during observation was twofold: to capture enough details to jog my memory later (Emerson et al., 1995), and to capture as much as possible in the language of the participants, verbatim where possible (Spradley, 1980). I also used drawings and photographs to capture maps, diagrams drawn on flip charts, and other graphical information. (It was not permitted to record the sessions. Moreover, recording would not be very useful because the discussion happened in many different places, the nonverbal action was very important, and it would be hard to follow on audio or even video recordings.) I knew that I would have no time to write extensive field notes during the simulation, because I wanted to spend as much time as possible observing the action. Therefore, after each session of observation (when I went back for lunch, for example), I looked through my jottings and added enough detail (by hand, in the margins) to clarify what I had written and seen. Immediately after the end of the simulation, I wrote more extensive field notes, using these extended jottings as a framework and filling in the details from memory (Emerson et al., 1995).

In my notes, I tried to capture discussion in the language of the participants, and to note other elements which might normally be tuned out or missed (Spradley, 1980). I also tried to include concrete details rather than generalized descriptions (Emerson et al., 1995). I focused on collecting data related to my broad research question, looking for behavior and actions that might shed light on how the teams made decisions and solved problems.

In the time outside of my once daily meetings in the Oxfam role, I was able to observe the teams as they worked. The amount of time varied depending on the nature and duration of their activities for the day, but on average I spent 8-10 hours with the teams each day. On the day of the transportation planning exercise (the fourth day), which is the focus of this study, I stayed with the teams from 7am until the dinner break around 8pm, sometimes with an hour away for lunch. After the dinner and debrief (from the media interview exercise), the teams return to work on the transportation planning exercise. In most trainings, I was able to return around midnight to check on the progress of the teams.

The teams' days were filled mainly with scheduled meetings (with facilitators) and unscheduled work time. During the unscheduled work time, team members were usually engaged in one of several activities: large team meetings, working with one or two others, working alone in front of a computer, meeting with facilitators, eating, or taking a coffee or smoking break. I could observe meetings and listen to the dialogue, I could see and record maps and diagrams drawn, and I could peek at laptop screens to see what people were working on. I kept my data collection unstructured, adapting it to the type of work that was happening and what seemed the best way to record it.

As a single observer, I was only able to observe one meeting or discussion at a time. During the days, there were often several meetings happening at one time. Most official meetings – those with facilitators – were scheduled ahead of time and I could choose to observe those most relevant to my study. The unofficial meetings – the discussions within the teams – were less predictable, but I learned that the most interesting discussions tended to happen when the entire team was together, after most of the scheduled meetings. I tried to be present at these times. I moved between the two teams often, but I tried to capture the whole of a discussion before moving to see the other team. I chose where to collect data based on the activities I thought would be most illuminating for this study. After the first training (first two cases), I had a much better idea when interesting activities would happen, and I used this knowledge to prioritize my observations.

However, it is important to stress that the data are necessarily incomplete. While I tried to observe most of the significant action, I was unable to see everything happening in both teams at all times. My data show many things that happened, but they do not show everything that happened. Nevertheless, much can be learned about teams' problem-solving processes from the data I was able to collect.

Documents

As mentioned earlier, I also collected documents produced by the teams during the simulation. A large number of documents are utilized within the simulation. The teams are provided with information about the scenario in documents like assessment reports, operational plans, maps, tables, and spreadsheets, usually given by one of the facilitators in their role as representative of another aid organization. These documents are part of the "script" from the training, and are not produced by the teams themselves. Email is a primary form of communication between teams and facilitators. Finally, as part of the scenario, the teams are required to produce a number of documents, including situation reports, operational plans, operating procedures, a transportation plan, and presentations for meetings.

I collected all emails and documents that were sent between teams and facilitators, and many (though not all) internal team documents. Different types of documents are used in different ways. Emails and other documents that are part of the training's "script" are used mainly to fill in the story and show what information the teams had access to at any given time. Those produced by the teams are used as an additional source of data, to confirm the results of decisions, and to triangulate evidence from other sources.

Other sources

Two other types of data deserve mention. First, as noted above, interviews were not conducted. Instead, participants were asked to describe their own thought processes during the simulation. As part of the meetings with donors and the humanitarian community, participants were asked to describe their "methodology" for transportation planning: how they created the plan and why they made the decisions they made. This is not a different source of data, because it is based on observation and documents, but it is treated as a source for triangulating inferences about participants' thought processes.

A second source of information were the informal conversations I had in my role as a facilitator of the training. There were many situations (such as over meals, during car rides, or waiting for meetings) in which I chatted with other facilitators or with the participants. I did not use these conversations as formal data, but they contributed to my own knowledge and understanding. In particular, I learned much about the participants' and facilitators' backgrounds and careers, and their opinions about the training's realism and usefulness.

I also told many people about their research, and gained insight from their reactions and ideas.

2.3.4 Data Analysis

The goal of data analysis is to build a process theory describing how humanitarian logisticians manage transportation planning. Building theory is an inductive process, in which the theory emerges from the data. The challenge is in "capturing the complexities of the real world, and then making sense of it"; in other words, taking the variety of data available and building from it a "structured understanding" of reality (Pettigrew, 1990). Most qualitative data analysis involves the development of conceptual categories or codes to organize the data and begin developing theory. Methods stress constant comparison between data and emerging theory (Glaser and Strauss, 1967; Corbin and Strauss, 2008), and sorting and re-arranging data to create and strengthen insights (Miles and Huberman, 1984; Eisenhardt, 1989; Langley, 1999). Process-oriented methods focus attention on what happens over time, so analysis often involves the development of narratives and orderings of events or incidents (Langley, 1999). Any of these methods can be applied to each case within a multiple-case research design, and supplemented by a set of techniques designed to search for cross-case patterns (Eisenhardt, 1989). My process of data analysis was guided by these methods.

A second tradition of data analysis was developed within the problem-solving and decision-making communities to deal with verbal protocols – subjects thinking aloud as they solve a problem – and related forms of data. These techniques focus on coding the verbal protocol data to describe the information-processing activities happening as a subject accomplishes a task (Woods, 1993; Carroll and Johnson, 1990; Newell and Simon, 1972). Problem behavior graphs can be created to show how subjects moved through a problem space (Newell and Simon, 1972). On the surface, these methods appear different from the more open, data-grounded approach described above, but they are closely related. Both involve coding to develop conceptual categories that describe the behavior of the subject, but one method assumes behavior fits an information-processing theoretical frame while the other makes no such assumption. In this thesis, I draw on both sets of ideas.

I concentrated first on understanding each case in detail (Eisenhardt, 1989). To develop an initial set of conceptual categories or codes, I began with the case I thought would provide the richest set of data. In a process of open coding (without a pre-existing theoretical frame), I went through my field notes closely, noting in the margin what I thought was happening in the data. By comparing these open codes with each other and with the data, I developed an initial set of codes that captured how teams worked through the problem and what elements of the problem they focused on. I continued to refine the codes through constant comparison between data and concepts, and by using the codes to define event sequences and chart team behavior. I added a second case to enrich the set of codes, following the same process (beginning with open coding). For each case, I wrote a narrative (Pettigrew, 1990) that described the team's behavior and progress toward solution, and I created graphics and tables that described that behavior in other ways (e.g. categorized by type of activity and time). With each of the remaining eight cases, I continued to refine the current set of codes, and I wrote a narrative and created charts to describe each team's behavior. In the process, I developed ideas and wrote memos to refine my understanding of each code and the relationships between them. What emerged from this process was a set of codes describing the teams' ways of thinking and ways of organizing work and data, and a second set of codes describing the specific elements of the technical problem as they saw it.

The second phase of data analysis involved searching for patterns across cases (Eisenhardt, 1989). The aims were to refine the theory and to confirm that the theory from one case (patterns of problem-solving behavior) were replicated in other cases (Yin, 2009). The analysis was anchored around actions or procedures enacted by the teams. I used the codes that emerged from the within-case analysis to divide the data and search for patterns within sets of codes, noting similarities and differences in the actions of each team (Eisenhardt, 1989). For example, I looked at how all ten teams made decisions about allocating cargo to vehicles and how they organized the set of cargo movement requests. I also tried to connect each of these single-code-centric analyses to each other, looking for patterns of behavior on a larger scale. For example, I looked at what activities preceded or were related to allocating cargo to vehicle (such as choosing a destination for that vehicle). In this manner, I identified three key processes enacted by the teams: understanding the problem, allocating resources to locations, and dispatching cargo and vehicles.

I went back to the data to compare how each team enacted these macro-processes, using process diagrams and charts to describe each team's behavior, and looking for similarities and differences across the teams. Two important processes emerged as the focus during this work. One was the process of understanding the problem, which appeared to be a process of sensemaking (as described in Section 2.2). The other was the process of "solving" the problem: making the decisions that constitute the transportation plan, including both dispatching cargo and vehicles and strategic decisions governing the allocation of resources. To investigate the sensemaking process, I used the codes relevant to understanding the problem to create "sensemaking maps" for each team, summarizing each coded incident in the order in which they occurred, and I analyzed the patterns by which teams moved through sensemaking incidents. To investigate the dispatching process, I returned to the raw data with the decision-making codes in mind, and drew a flow chart diagram to describe each team's decision-making process. Next, I compared the charts to one another, looking for similarities and differences, until two archetypal processes emerged. I returned to the data to check whether the patterns I found were confirmed by the data in each case. The patterns found from these two processes, sensemaking and solving, are described in the following sections.

2.4 Introduction to findings

I began by asking, broadly, how humanitarian logisticians handle transportation planning. As data collection and analysis proceeded, two kinds of work emerged as important: sensemaking and solving. Therefore, they became the focal points for the continuing research. Sensemaking is a process of understanding and framing the problem, while solving involves making decisions that fix elements of the solution. Two main kinds of decisions are made in the process of solving: strategic and dispatch decisions. Strategic decisions have a broad scope, such as determining the major routes along which cargo will flow, or allocating vehicles to bases. Dispatch decisions are specific allocations of cargo to vehicles and vehicles to movements. These three kinds of work – sensemaking, strategic decision-making, and dispatching – are intertwined as the teams progress through the problem. Sections 2.5 and 2.6 describe how the teams performed each of these activities, and how the activities relate to one another. First, however, the problem faced by the teams is described in more detail.

Recall that the setting for this research was an in-depth simulated emergency. In the first three days after the earthquake, teams of experienced logisticians had to gather information, assess logistics requirements, and plan their operation, including locating warehouses and bringing in vehicle fleets. On the fourth day, they must put this plan into action by planning a large number of requested cargo movements. At the same time, they are being interviewed, one by one, by a reporter asking tough questions about why aid is not reaching people in need. The team leaders are being called away to important meetings with angry government officials. They are also preparing for two important meetings the following day: a meeting with donors from whom they hope to secure funding for the operation, and a meeting with their Logistics Cluster partners who are expecting to hear that their cargo has been moved (or, if it has not been moved, why not). In short, there are a number of distractions from their transportation planning task. More importantly, the true goals of their task are not clear, as they may be worried about delivering cargo efficiently, moving cargo for important partners, or showing donors that their work is essential. Furthermore, by this point in the week, the teams are fully immersed in the simulation: they have learned that they cannot rely on all the information they receive, have gotten very little sleep, and are under very real pressure to provide results that will help the people of the fictional country in which they are operating.

The transportation planning task is introduced on the morning of the fourth day. Teams are told that they have "jumped in time": ten days have elapsed, cargo has started arriving, and vehicles have been brought into the country. The teams are given a list of about 100 cargo movement requests (CMRs) that they have received from other agencies. They have also secured a fleet of large 40-ton trucks, a fleet of small 10-ton trucks, and a fleet of helicopters. The 40-ton trucks are available immediately, and the 10-ton trucks and helicopters will arrive throughout the upcoming week. They are given information about these vehicles, including their cargo capacities, the speed at which they travel, and the roads and routes on which they can move. The transportation network is shown in Figure 2-1. The most affected regions are in the southern part of the country, around Mammoth and Vail. Most of the cargo originates in Snow, and some comes from Jay as well. Many destinations, especially those near Vail, are only accessible by helicopters, and Vail itself can only be reached by 10-ton truck or helicopter.

Teams are asked to plan the movements of their trucks and cargo for a period of 7 days. Tomorrow, they will be required to present their plan to both donors and the humanitarian community; the former will wish to see progress in delivery of aid, and the latter will be eager to know how quickly their goods will be delivered. This is the transportation planning problem that forms the basis for this study: it focuses on how teams create the

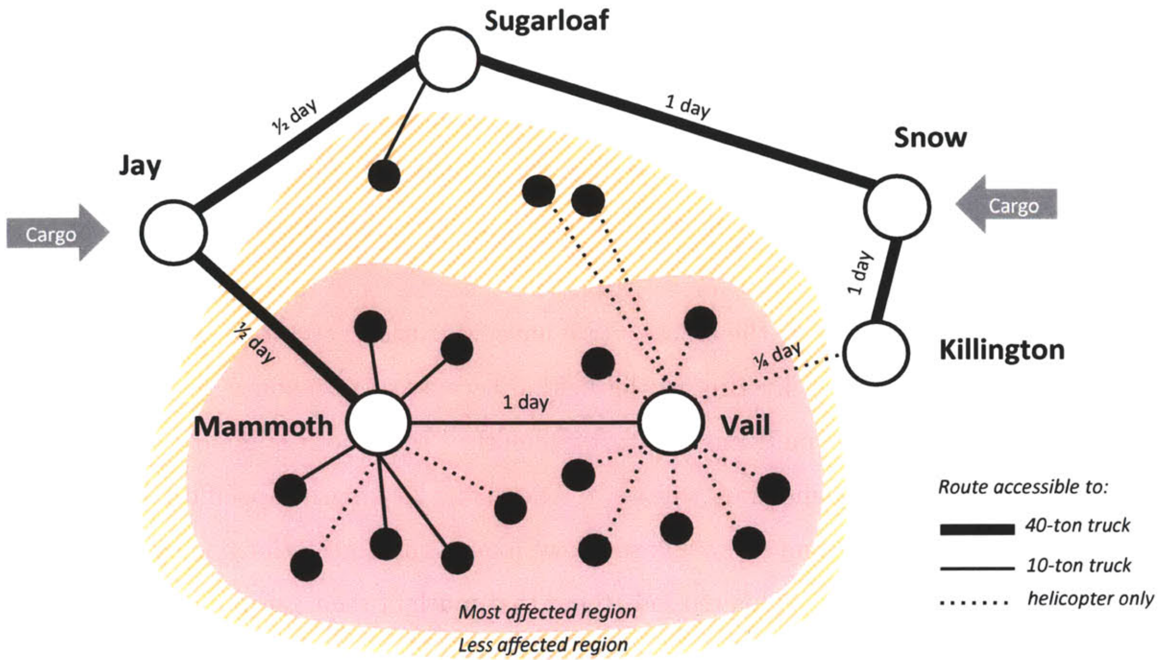


Figure 2-1: Transportation network in Snowland

transportation plan describing the movements of trucks and cargo for the next seven days.

2.5 Findings: Sensemaking

One of the first things teams must do when confronted with the transportation planning problem is to understand it, to figure out what they must do. The following quote from my field notes on Team G illustrates the activities of one team, observed as they were starting to work together on the problem. (Quotation marks indicate direct quotes from a team member; other text is directly quoted from my field notes, with the exception that team names have been replaced with "the team" and the names of locations have been disguised. Throughout the thesis, specific teams are occasionally referenced by a letter, such as Team G. Each team was assigned a letter, in a random order, so that teams cannot be individually identified.)

> The team was discussing how to start, and how to get a handle on the information. Everyone was making suggestions at once. "Let's see who started in Vail and who started in Mammoth." ... "We should first focus maybe" "Let's see what are the CMRs for each place." ... Someone suggested that they

> "filter per destination." Someone else said to filter "first per destination then per priority." Several people have computers out and the word "filter" refers to the Excel function. As they talk they are filtering and also printing some of the filtered lists. "The priority should be from the priority list. First step is filtering by places and we know... then we select the cargo according to the capacity we have." ...
>
> They start looking at the trucks. "We have 20 trucks... 800 tons available." ... "So let's start by delivering to the field... because of warehouse space." ... Someone says they should move it to Mammoth: "Mammoth is good because that's where the beneficiaries will be." They start looking at specific CMRs. Looking at one, someone says, "Ok so Stowe [one of the destinations] is by air... so what is [by] 40-ton." This reminds them that much of their cargo must go by air. "Oh yeah so we cannot make deliveries to the field except by helicopter." They look at another destination. "Heavenly total is... we have 155 tons for Heavenly."

This quote illustrates how one team, in a few minutes, started trying to understand the problem. First, they looked at the list of CMRs they were given. They tried to figure out where (at what towns) the goods were waiting to be picked up, where they needed to be delivered, and what their priorities were (based on the type of item - shelter, health, water and santitation, food, or other). This is an example of exploring the data they are given in order to understand what they must accomplish. Next, they looked at the vehicles available, and started thinking about how and where to move cargo first - focusing on delivering to areas with many beneficiaries, or moving cargo out of congested warehouse space. This is an example of stating goals or justifications for decisions. Finally, they began to think through deliveries, by looking at a specific CMR, and realizing that it has to go by helicopter rather than 40-ton truck. By the end of these few minutes, the team has some idea of how much cargo they must move, and they are starting to understand how it will flow through the network, but they do not have much idea of how to create their transport plan.

This quote suggests that sensemaking involves progression through related and unrelated ideas, looking at available information, and thinking about what to do next, without any clear organization to the process. Through analysis of many sensemaking episodes like this one, across all 10 teams, I found some patterns in the elements of the problem that were "made sense of" by the teams, and in the mechanisms by which teams came to consider them. The following paragraphs describe these common patterns, then advance a model of the process of sensemaking and its interaction with solving.

2.5.1 Elements and outcomes of sensemaking: understanding

Sensemaking, as its name suggests, is a process of "making sense" of the problem, trying to understand what must be done and how to go about doing it. Across all 10 teams it is clear that there are certain elements of the problem that must be "made sense of." Through the data analysis process, certain concepts (captured as codes) appeared again and again across teams as part of their process of understanding what they had to do. Table 2.1 shows these elements, along with the number of teams that were observed considering each one and sample quotes to illustrate each element. (In this and all following tables, note that I did not observe all the activities of every team. Therefore, my lack of observation of some activity by some team does not mean that it did not occur.)

The elements shown in Table 2.1 represent the "outcomes" of sensemaking, suggesting that it is a process of understanding the problem. Four main kinds of understanding seem to be gained through sensemaking: the physical constraints of the problem, the challenges which must be addressed, the goals or objectives to work towards, and an ongoing situational knowledge of the current state and what tasks must be accomplished. These four kinds of understanding are described in the paragraphs below, and noted in Table 2.1.

First and most clearly, teams become familiar with the *physical constraints* of the problem, by seeing how cargo flows through the network, what must be moved from where to where, how much transportation capacity is available, and other more detailed constraints. One of the most important realizations is that cargo flows through the network in multi-modal paths, which means that much of the cargo originating in Snow must be transported first by 40-ton truck to Mammoth, then switch to 10-ton truck to reach Vail, and finally to helicopter to reach its final destination. The problem appears much easier before teams realize they must plan transfers between modes. Another important element of physically understanding the problem is organizing the cargo movement requests. For example, teams that sum up cargo to move all over the network have a much different understanding of the "demand" for transport than teams that simply add up cargo waiting at each origin, without considering where it will go next. (As I will discuss later, these two different "pic-

Sensemaking outcomes: what teams "made sense of"	No. teams observed	Type of Understanding	Sample Quote(s)
Cargo			
Picture (i.e. organize) cargo by origin, destination, item type, mode, route	10	Physical constraints; Situational	"Split cargo into those [categories] - shelter, [health, water and sanitation, and food] on the screen," "then split them where they are [now]... then where they are going."
Sum cargo from, to, on route, by mode	8	Physical constraints; Situational	"Heavenly total is... we have 155 tons for Heavenly."
Make filtered/sorted lists to dispatch from (each origin)	6	Physical constraints; Situational	"I want [the list of cargo] separated by location - one sheet for each location - so we can put it up [on the map] so we can see." ... "Five locations, sorted by priority and then where it's going."
Deal with individual CMRs	4	Physical constraints; Situational	"High-energy biscuit is considered a priority, right? I will put not as food I will put as health."
Network			
Describe transport corridors	8	Physical constraints	"[The corridor goes from] Snow to Vail. The first leg is to Mammoth, then move it to 10-ton."
Describe hubs	5	Physical constraints	"Ok from Snow to Heavenly you have to go through Mammoth," meaning that Mammoth is a kind of hub for this shipment.
Note time (on major routes, to reach, ...)	8	Physical constraints	"We start on day 1. We have all our trucks in Snow." "One day they can go to Jay." "No it will take 1.5 days to go to Jay."
Note accessibility of destinations/roads by modes	6	Physical constraints	"Write down from location to destination the type of truck we can use."
See that cargo follows multi-mode paths (through hubs)	7	Physical constraints	"We cannot send Snow to Vail with 40-tonners so you send to Mammoth."
Assign routes/hubs/modes to CMRs	7	Physical constraints	"[Getting to] Okemo means that you [first] have to drop off at the hub Sugarloaf. So can you put the hub in a column [of the spreadsheet]?" ... "Someone needs to do that on all of the CMRs."
Assign routes/hubs to destinations	4	Physical constraints	"What destinations is Mammoth [hub] going to serve?"
Consider/decide on Killington airbridge	5	Physical constraints	"If the helicopters are free, it's better to send the cargo to Killington and make the [airbridge] rotation."
Tracking and breaking down			
Track movements (to merge, to report)	8	Situational	"Let's agree on the [tracking] format that's compatible so we can merge [the road and air teams' plans] later."
Break down by mode	7	Physical constraints;	"Can you separate surface transport and air transport?"
Break down by location (or consider it)	3	Physical constraints;	"We have responsibility for a hub - this is my hub, this is what's coming and going."
Realize planning must happen in order, by day, or link up carefully	6	Physical constraints	"So you [the helicopter team] plan for the first and second days and then we [the road team] will tell you what you have [received by road]."
Transport demand and capacity			
Find demand for transport	4	Physical constraints	"What is the demand for transport - determined by goods to ship at each hub [and transit times]?"
Find transport capacity (per destination, per origin, per day)	7	Physical constraints	"I think it's a multiplication of two numbers. You have the number of rotations and the capacity and like this you can have the tonnage [capacity] per day."
Other (minor) operational constraints			
Consider warehousing (can ignore it)	9	Physical constraints	"At the end of every day we will calculate how much we have in the warehouse at each location."
Consider restrictions on helicopters: fly from bases, capacity based on rotation length, multi-stop routes	7	Physical constraints	"above 75 [km] [helicopters] can do half the rotations"
Consider loading time (ignore it)	1	Physical constraints	
Consider splitting CMRs, or using either weight or volume to plan	6	Physical constraints	"You want to dispatch your shelter, WASH, and your NFIs by m3, and your food by MT." "You may not be able to move the full CMR. The CMR may have [to be split into] 2-3 legs."
Challenges			
See that it's difficult (due to time, capacity) to deliver to Rocky/Vail	6	Challenge	"Today we can't move anything from Mammoth to Vail because we don't have 10T trucks." "We either accept that we're going to take a long time to deliver or... we make a request for more helicopter capacity."
See that helicopters are idle first few days in Vail because no cargo is there for them to transport	6	Challenge	"We have to move the materials to Vail in order to use the helicopters."
Worry about congestion in hubs	1	Challenge	
See that there is not enough capacity (to transport all, of some mode, ...)	6	Challenge	"The trucking capacity certainly over-meets the needs... but we don't have the air capacity to meet the needs."
Goals			
Use vehicles well, avoid idle vehicles	7	Goal	"We have one helicopter today - let's look at making the use of that helicopter for the highest priority cargo to make best use of it."
Feed cargo to onward transport	6	Goal	"We have to make sure we always have enough stuff in Mammoth and Vail to move stuff with the helicopters.. and secondary transport."
Prioritize (by item type, destination, organization, ...)	10	Goal	"So whenever we have a shipment we consider whether the item is a priority... and the geographic priority."
Send mix of cargo types	4	Goal	"the other way [P1 first] I was very worried that we would send a lot of shelter and there would be no room for anything else."
Send what is there / move lots	2	Goal	"Our priority is to move as much cargo as we can move with those trucks."
Cost less important (than speed)	3	Goal	"Cost should be second priority. First we try to deliver the cargo."
Situational			
Note there is (some, lots, none) - or check what there is - going (from, to, on route)	4	Situational	"If you look at Snow to Vail, you have a lot..." "bulky things from Jay."
Note there are vehicles available to task	6	Situational	"Remember you've got one helicopter today"
Tasks			
Move lots (from, to)	7	Situational	"Today we have to focus on moving as much as possible from Snow to these locations"
Send this (specific) cargo	3	Situational	"Those blankets - let's try and move them today so we can show them we're working on it."
Return to pick up cargo left behind	2	Situational	"We'll have to leave a whole bunch of stuff in Snow and come back and get it."
Position vehicles at a base or on a route	8	Situational	"But we have to bring the Jay trucks to Sugarloaf and then down [to reach Okemo]"
Given this situation, what next?	5	Situational	"Ok where is it going?"
Use vehicle(s)	3	Situational	"The 40-tonners have to go somewhere... the 10-tonners have to go somewhere"

Table 2.1: Elements of sensemaking

tures" of the cargo movement requests lead to different problem-solving strategies.) Thus, sensemaking can lead to different understandings of the physical constraints, which provide different perspectives and influence their general understanding of the problem.

A second kind of understanding gained through sensemaking is a sense of the *challenges* to be overcome. One major challenge is that the transportation capacity is insufficient to transport all the cargo. A second is that it is difficult to reach the most affected areas around Vail, because cargo must be transported all the way around the mountains. As a result, a third challenge arises, that the helicopters in Vail are initially idle because cargo cannot reach them for several days. Logisticians consider this extremely undesirable because it is a waste of very expensive transportation capacity. These challenges are distinct from constraints, though they may be related to them. Vehicle capacity constraints dictate how much cargo can be moved, but when the teams interpret this to mean they can only move some of the cargo, they are recognizing a challenge. This challenge, in turn, may influence them to think carefully about prioritization. While it is possible to create a transportation plan without recognizing challenges, they can strongly influence teams' planning.

A third kind of understanding gained through sensemaking relates to the *goals* or objectives of the problem. At the start of the transportation planning exercise, teams have only a vague sense of what is important (transporting cargo to people who need it, defending plans to stakeholders, and pitching them to donors), but as they work in the exercise they formulate several possible goals. All teams discuss different ways to prioritize cargo: based on the type of item, its destination, the organization shipping it, or other schemes. Many teams also consider simply sending out as much cargo as possible, or addressing some of the identified challenges, like keeping the helicopters busy. Some teams discussed whether they should consider minimizing the cost of the operation, though most decided it was the last priority. Identifying these many possible objectives is a crucial outcome of sensemaking. However, identifying goals is only a first step: teams must then decide which of these objectives they will use as they solve the problem. Goal-setting therefore includes both a sensemaking outcome – understanding a goal – and a solving outcome – deciding on its importance. In the discussions below, goal-setting is considered a "solving" activity, but it actually resides in both categories.

A fourth kind of understanding gained through sensemaking is an ongoing sense of what *tasks* must be accomplished in the current situation (meaning given the current state of the plan). For example, consider the situation before the team has planned any movements. Many teams noticed (through organizing the CMRs) that there was a lot of cargo to move from Snow, and that there were many vehicles in Snow. Often, their next step was to move the cargo out of Snow on the available vehicles. This situational kind of understanding is distinct from the two described above, in that it changes as the team plans movements of cargo and vehicles. Later, after teams had planned some of the cargo movements, they realized that in this new situation, they had to return to pick up cargo they had left behind. This situational understanding includes seeing, given the currently planned movements, what the situation is (there is a lot of cargo to move, there are vehicles available) and noting tasks to be accomplished (pick up cargo left behind, use the vehicles). This kind of understanding is one of the important links between sensemaking and solving, since it often leads to the planning of specific cargo movements (to be discussed in more detail later).

Thus, sensemaking leads to an understanding of the physical constraints of the problem, the challenges to be overcome, goals or objectives to aim for, and what to do or what can be done in a given situation. This understanding emerges gradually over time, as sensemaking and solving progress.

2.5.2 Sensemaking activities: exploring, understanding, and solving

Describing the elements of sensemaking shows us the results of the process, but does not illuminate the activities that constitute it. How does sensemaking proceed? To answer this question, I looked at the process by which each team came to understand the problem they faced. I used the conceptual codes that emerged in data analysis to highlight the incidents relevant to sensemaking, and I created a "sensemaking map" for each team that showed all the incidents, in the order in which they occurred. Figure 2-2 is an example of one of these maps, from Case F.

Examining the sensemaking maps across all cases suggested that there were three general categories of incidents. One set of incidents showed teams gaining new understanding of the problem; these corresponded with the outcomes of sensemaking described in Section 2.5.1. A second set of incidents described exploratory actions, aimed at probing the problem space. The third set of incidents described actions that embodied progress in solving, such as setting a goal or determining a major route. The sensemaking map in Figure 2-2 shows incidents in three columns, corresponding to these three categories of incidents.

In the paragraphs below, the exploring and solving categories are described in more detail (the understanding category, in the center column, was described in Section 2.5.1.) As each category is discussed, it will become clear that incidents in one category often led to incidents in another category. These interconnections will be the focus of the final section.

Exploring

The first set of incidents described actions in which, deliberately or not, teams explored the problem. The most simple actions in this category involved *noting* things, such as the vehicles available or the accessibility of roads. For example, someone noted the number and types of vehicles by saying, "We have some 40-tonners first... 10-tonners and you have helicopters. Three types of moving assets - resources." This team member was simply verbalizing information given to them that morning. A related activity involved *noticing* the situation, as opposed to simply noting information about the problem. For example, while looking at the cargo waiting at each origin, someone noticed they had "bulky things from Jay." On another team, someone looked at the cargo and realized, "there are [passenger] vehicles in the cargo." Often, these noting and noticing incidents led to further action, such as deciding not to prioritize vehicles as cargo, or trying to use available vehicles, so even these simple activities were important components of sensemaking.

Beyond these simple activities of noting and noticing are somewhat more complex kinds of exploratory actions. One involves the organization and manipulation of data, which I have labeled *picturing* to emphasize the various perspectives that can be created. Teams were given a long list of cargo movement requests (CMRs), and they struggled to sort, filter, sum, and diagram it to understand what they must move where, by what mode of travel. For example, someone said they need to "assess what we have to get to where - some data sorting that has to be done." One team tried to accomplish this by listing the cargo in each location: "We're gonna list down what [cargo] we have at each location. Then we look at the [vehicles] we have at each location." They later tried to sort cargo by its mode of travel – "Put a column for 40-ton [truck], put a column for 10-ton, put a column for helicopter" – and then by the next major hub it was going to: "We've divided it into everything going to Mammoth, everything going to Jay, and everything going to Sugarloaf." All the teams created some version of these organized lists of cargo, using them to "calculate the tonnage that you have [to move] per day," for "filling the CMRs per asset [onto vehicles]," and

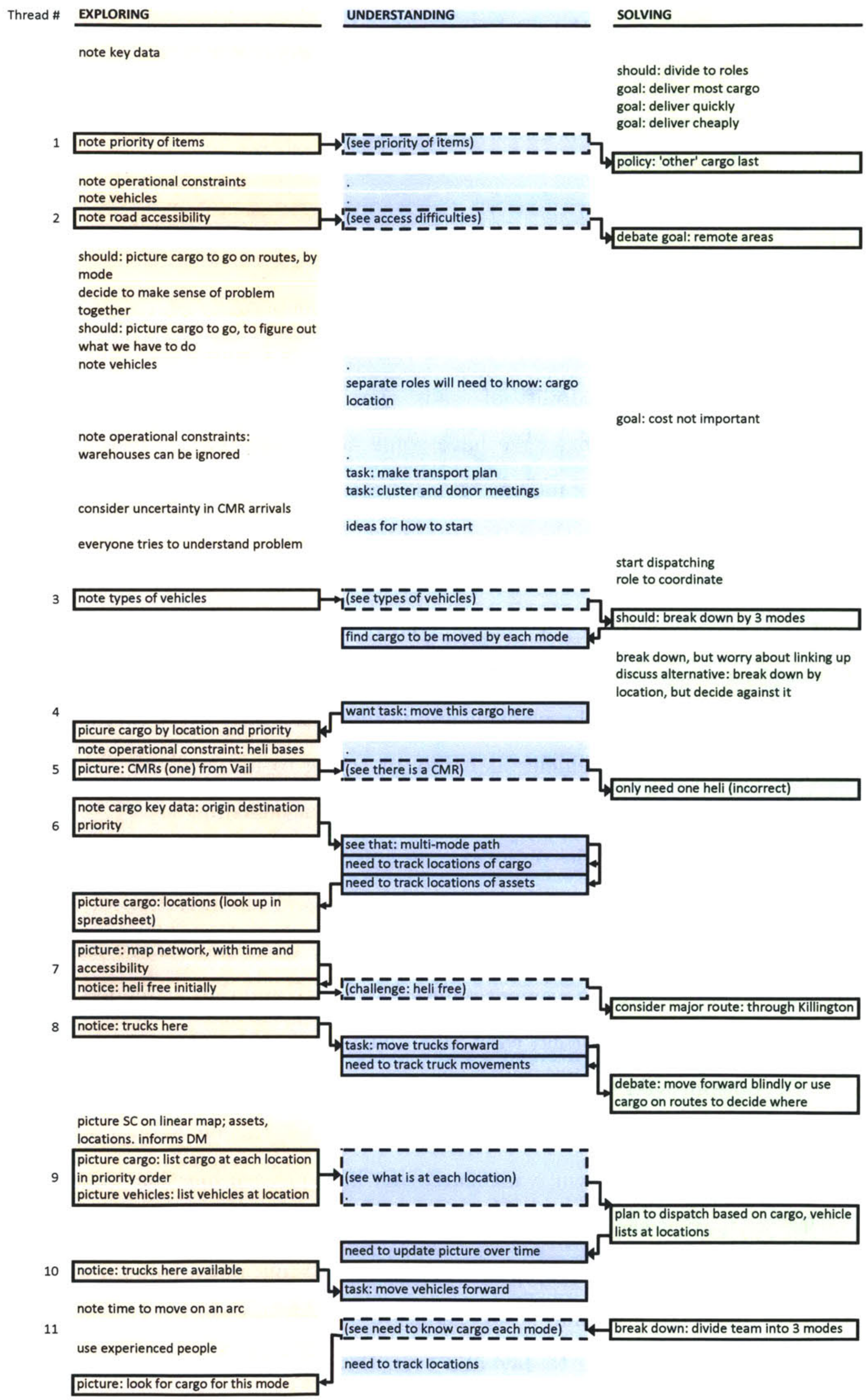


Figure 2-2: Sensemaking map from Case F (first half). Sequential table of sensemaking incidents. Outlined boxes with arrows indicate threads of related sensemaking activities; dashed outlines show implicit rather than explicit understanding activities.

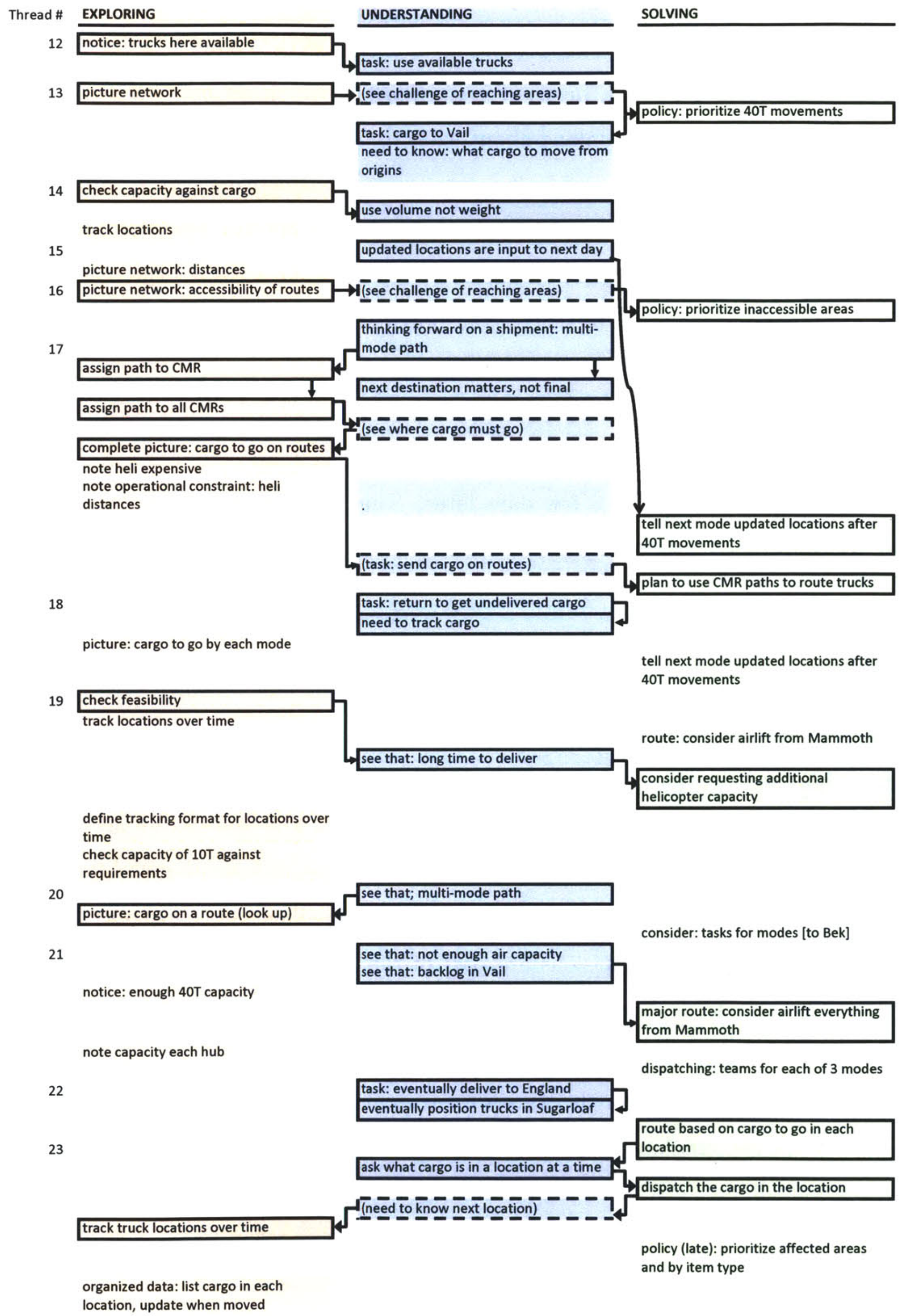


Figure 2-2: Sensemaking map from Case F (second half). Sequential table of sensemaking incidents. Outlined boxes with arrows indicate threads of related sensemaking activities; dashed outlines show implicit rather than explicit understanding activities.

generally to "know where all the [CMRs] are and where you have to move them." Such pictures were central to the solving process as well as to sensemaking.

A third kind of exploration involved *projecting* forward through the next steps in the problem. This is not simply thinking about what to do, but actually following some entity, such as cargo or a vehicle, forward as it moves through the network. For example, looking at how to move a CMR from its origin to its destination, someone said, "Ok from Snow to Heavenly you have to go through Mammoth." Projecting the steps led the team to realize they had to stop in Mammoth to switch transport modes. Another team projected forward through steps to understand the time it took to reach destinations: "We start on day 1. We have all our trucks in Snow." "One day they can go to Jay." "No it will take 1.5 days to go to Jay." By projecting, the teams were able to understand the implications of a given decision now on the situation a few days later. Thus they started to probe the problem space and learn about the problem.

Solving

The activities in the "solving" column of the sensemaking maps involve decisions that fix some aspect of the transportation plan. For example, teams might set a goal of prioritizing the most affected locations, or decide to position helicopters in a particular base, or send a truck to a new destination. The first of these examples is a goal-setting activity. Recall that goal-setting includes aspects of both sensemaking (coming up with a possible goal) and solving (deciding to work towards it), but in the sensemaking maps goal-setting is in the "solving" column. Solving decisions are part of the sensemaking maps because, in many cases, sensemaking leads directly to a solving decision, or solving activities lead directly to sensemaking. In these cases, solving activities were included in the sensemaking map because they were the result of or the trigger for sensemaking activities. The presence of these decisions in the sensemaking maps highlights the strong interactions between sensemaking activities and solving activities. These interactions are the focus of the next section.

It is worth mentioning here that in this case, "solving" refers to the development of a plan, rather than the implementation of it. A solving decision is one that fixes, at least temporarily, some aspect of the plan: deciding to send a particular truck to a particular location at a particular time, for example. In this paper, the actual sending of that truck, in fulfillment of the plan, is not considered. The teams solving the problem use language that suggests implementation as well as planning, but in all cases decisions refer to plans.

2.5.3 The process and mechanisms of sensemaking

The preceding paragraphs described the kinds of activities involved in sensemaking. This section steps back to consider the sensemaking process, looking to understand how the activities of exploring, understanding, and solving are interrelated by examining how teams proceeded through them. The sensemaking map in Figure 2-2 will serve as an example, but similar analyses were conducted for all ten teams. (While there were differences in the specific processes followed by each team, there was evidence of the same kinds of interactions between activities.)

In the sensemaking map, individual incidents are shown in colored boxes, and many of them are connected to other incidents. I will refer to these connected sets of incidents as "threads". As is clear from the long quote at the beginning of this section, sensemaking is a somewhat disorganized or random process: ideas often lead to other ideas, but just as often they lead to nothing (observed) or arise from nothing (observed). My field notes reflect this lack of organization, occasionally showing direct connection between ideas, sometimes showing one idea that led to another a few minutes later, and often showing ideas that did not seem to lead to anything at all. In creating the sensemaking maps, I noted all incidents that seemed part of sensemaking, even if they did not clearly lead to anything. I included arrows between incidents only when the field notes suggested that one incident led to another, whether directly (clear, immediate progression from one idea to another) or indirectly (one idea seemed to follow from another, even though separated somewhat in the field notes). I label these connected incidents "threads." Note also that some boxes are outlined in dashed lines. In these cases, the incident is *implicit* in the flow of work, but not explicitly stated by any participant nor recorded in my field notes.

The first sensemaking thread in Figure 2-2 begins with an exploring activity: noting the priority of items. Most of the CMRs are labeled as shelter, health, water and sanitation, or food items, and given a priority number (1-4) accordingly. In the first incident in this thread, one team member, looking at the list of CMRs, said, "On CMRs clear on priority - if [there's] anything outside of that let's put those at the top." In fact, there are several CMRs that are labeled "other" and have no priority number attached to them. The team tried to decide what to do with those, saying, "we've got to make a value judgment," and decided they should be given last priority: "the four categories first, then the other." This is the last incident in the first sensemaking thread. Thus, this thread shows how, from the initial activity of noting the priority of CMRs, the team surfaced a challenge, or a dilemma in this case, in which they needed to decide how to prioritize the "other" CMRs. In response to the dilemma, they made a solving decision, giving a particular priority level to a set of cargo movement requests. Thus, the thread progressed from exploring, to (implicitly) understanding a challenge they must solve, to solving.

Different sensemaking threads progress through activities in different orders. Thread 4 in Figure 2-2 begins when a team member wants to know, "what do we have to do to get to the point where we can say this cargo goes there?" This suggests that the team, or at least this team member, has formulated a task of figuring out what goes where. In response, the team begins picturing (re-organizing) the cargo to answer this question. In this case, understanding a task the team must accomplish leads them to an exploring activity. Other threads show solving activities that lead to understandings (e.g. 23), or even understandings leading to other understandings (e.g. 18) or exploring leading to other exploring (e.g. 7). Of course, solving activities often lead to other solving activities, but these progressions are less relevant to sensemaking so they are not included in these maps. Generally, this map suggests that Team F, at least, made sense of the problem by moving between exploring, understanding, and solving activities.

Teams moved between these three types of activities unconsciously, without making distinctions between them. There is no clear pattern to the order in which teams moved between activities. However, on close examination, the understanding activity is often central. Most threads involve an understanding activity, usually linked to exploring or solving (or both). Those threads that do not explicitly involve an understanding activity often do so implicitly. Take, for example, thread 3. It begins when one member of Team F says, "We have some 40-tonners first... 10-tonners and you have helicopters. Three types of moving assets - resources." I classified this as an exploring activity, noting the vehicles available, but the second sentence shows that the team is also understanding the implications: there are three different types of vehicles. The team next discusses breaking down the problem by putting teams on these vehicles. Even though the team's conversation only indicated exploring and solving activities, there is an implicit understanding activity between them.

One implication of this centrality of understanding is that a new or different understanding of the problem leads the team to new or different actions. For example, consider thread 17. After realizing (a second time) that cargo followed a multi-modal path, they began assigning such paths to each CMR, which, in turn, made it possible to find the amount of cargo requiring transport on any given route. Understanding a physical constraint led them to picture the cargo in a different way. Understanding can also lead to solving actions. In thread 19, understanding the challenge of delayed deliveries led the team to a solving action: requesting additional helicopter capacity.

If new understanding leads to new actions, what leads to new understanding? New understanding can result from exploring or solving activities, or even from other understanding activities. In thread 10, exploring the problem by noting the availability of trucks led the team to formulate a new task to use them; this is a simple example of exploring leading to new understanding. In thread 23, the team's attempt to route trucks based on waiting cargo led them to formulate a task of figuring out what cargo is in each location at any given time; solving led to new understanding. In thread 17, understanding that cargo follows multi-modal paths led the team to realize that they must plan movements to hubs rather than final destinations; a new understanding led to another new understanding.

Ultimately, the evidence suggests that sensemaking is a cycle between understanding and actions. Exploring and solving activities are actions that lead to new understanding, which leads the team to new or different actions, which lead back to understanding. As this cycle continues, the team's understanding of the problem changes and increases, and they choose different actions accordingly. For example, in thread 3, just after Team F considered breaking into teams for each mode (a solving action), they realized that they needed to know which cargo had to be moved by each mode, giving them a new task to add to their understanding of the problem. Previously, they had done some general sorting of cargo based on locations and modes, but now they planned to re-sort the cargo (back to exploring) to answer this more specific question. This sort of reframing is a common result of new understanding gained by the teams, leading to new actions, as over time the team's understanding of the problem evolves.

At this point, it is possible to describe a more general model of the process, depicted in Figure 2-3. Exploring and solving actions lead to understanding, and progress in understanding in turn leads to more, sometimes different or more targeted, exploring and solving

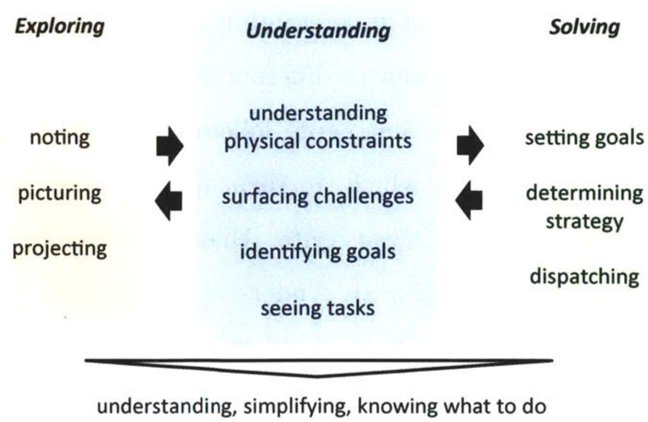


Figure 2-3: Model of sensemaking in humanitarian transportation planning

actions. In this way, teams make sense of the problem and begin to solve it at the same time. As teams move between exploring, understanding, and solving activities, their understanding of the problem increases and changes. They begin to understand the problem's structure and constraints, they identify the challenges which they must work to address, and they notice the situation and formulate a set of tasks that guide their immediate actions.

2.5.4 Insights and implications of the sensemaking model

The word "sensemaking" has been used to describe the phenomenon studied here: how teams came to understand the problem they were trying to solve. However, it has not yet been demonstrated whether and how the process observed here relates to the concept of "sensemaking" described in the literature. As discussed in Section 2.2, Weick (1995), Klein et al. (2007) and others describe sensemaking as updating a frame by perceiving and generating cues; the frame directs what cues are perceived and how they are interpreted. The frame also influences what actions are taken, which generate additional cues. In the humanitarian transportation planners' process, the frame is equivalent to the "understanding" generated through the sensemaking process (the central box in Figure 2-3). Exploring and solving actions led to increased understanding, or an updated frame, which in turn directed new action. The humanitarian logisticians' process fits neatly into existing sensemaking theory.

These findings elaborate upon sensemaking theory by showing the mechanisms of sensemaking in humanitarian transportation planning. The frame consists of an understanding of physical constraints, a sense of the key challenges and goals to be addressed, and a set of tasks to accomplish. The actions that generate cues to update the frame are exploratory, involving noting aspects of the problem, picturing or re-organizing problem data, and thinking forward step-by-step through possible actions. Solving actions can also trigger updates to the frame. In the cases studied here, solving actions led to sensemaking when teams were unsure what to do next or how to accomplish the next step. In humanitarian transportation planning, then, exploratory and solving actions update the frame by probing physical constraints, surfacing challenges, and bringing to light new situations in which new tasks can be formulated.

This study's elaboration of sensemaking's outcomes and mechanisms has additional implications for understanding the interactions between sensemaking and solving. As discussed in Section 2.2, sensemaking and solving are typically studied separately. Few studies have focused on problems in which both are required, and there is not much theory that describes their interaction (Rudolph et al., 2009). My findings suggest that sensemaking and solving are intertwined and inter-dependent, that sensemaking can be seen as a process of formulating or simplifying a problem, that the formulation resulting from sensemaking is oriented around tasks, and that goals emerge through a combination of sensemaking and solving. Each of these implications is discussed in more detail below.

Sensemaking and solving are intertwined and inter-dependent

The model in Figure 2-3 emphasizes the intertwined nature of sensemaking and solving, and illuminates the mechanisms by which they are connected. As described earlier, sensemaking often leads to solving actions. The surfacing of a challenge may spark a solving action to address it; understanding constraints makes clear how solving can proceed; and formulating tasks directs immediate solving actions to accomplish the task. In the other direction, solving actions also lead to sensemaking, especially when teams are stuck, in that they wonder what to do next or need to find something out.

In all cases, teams moved fluidly and often between sensemaking and solving: the processes were temporally intertwined. Evidence suggests they are also more fundamentally intertwined and inter-dependent, in that sensemaking requires solving, and solving requires sensemaking. It is clear that solving requires sensemaking: teams had no idea what to do or how to do it when they were given the problem data, but they figured it out as they gradually made sense of the problem. Sensemaking also depends on solving, though perhaps in a practical rather than a pure sense. It would, in theory, be possible to completely make sense of the problem before beginning to solve it, but the teams' experience suggests that it would be extremely difficult to do so in practice. Teams had difficulty understanding the implications of early decisions on later situations; it was only through making some decisions and progressing forward in the solution steps that they were able to surface challenges, understand new situations, and formulate new tasks. Thus, sensemaking and solving are not only temporally but also fundamentally intertwined: solving requires sensemaking to provide an understanding of what to do and how to do it, and sensemaking requires solving to surface challenges, situations, and tasks that are evident only in later planning stages.

Rudolph et al. (2009), in their study of clinicians diagnosing and treating an illness, come to the same general conclusion: sensemaking and solving are "inputs to each other". This study shows the same dynamic in a very different problem context, in which solving is a series of moves rather than a choice between options, and sensemaking is understanding the problem rather than cultivating multiple options. This study expands upon the theory explaining interactions between sensemaking and solving by providing more general descriptions of the components of the frame and how it both directs and is updated by solving. This expanded theory may be able to explain the observed behavior of planners and schedulers in manufacturing facilities. Empirical studies suggest that schedulers monitor problems by collecting information (Vernon, 2001; Jackson et al., 2004; McKay et al., 1995; Webster, 2001), and deal with them using heuristics or rules (McKay et al., 1995; Webster, 2001; McKay and Wiers, 2001; Grant, 1986). The latter may be a solving process, surfacing constraints and challenges, while the former may be a sensemaking process, which updates the planner's understanding of challenges and needed tasks. The findings in this study suggest more generally applicable mechanisms by which sensemaking and solving interact in complex real-world problems.

Sensemaking is a process of understanding, formulating, and simplifying

In this study, sensemaking was a process of understanding the problem: what to do and how to do it. In effect, it was an ongoing definition and redefinition of a problem formulation, or what Newell and Simon (1972) define as a problem space. The formulation is critical, because people or optimization models search for a solution within the space defined by the formulation. Formulating and solving are typically portrayed as a sequence: first a formulation or problem space is conceived, then solving proceeds within that space (e.g. Volkema, 1983). This study, in contrast, shows the formulation emerging over time as the problem is solved.

The nature of the formulation can influence the success and speed with which a problem can be solved. Simpler, more restrictive formulations can lead to faster and easier searches, but may also exclude the best solutions. Such a dynamic may be at work in contexts, like firefighting, where sensemaking consists of recognizing a problem (a simplistic formulation), which leads to an immediate action (a short search) (Klein, 1993). In contrast, in operations research, much effort is devoted to finding formulations that both speed search and retain the optimal solutions (e.g. Armacost et al., 2002). These observations suggest one way to interpret the sensemaking process observed in this study: as an ongoing simplification of the problem formulation. The most un-simplified formulation would consist only of an understanding of allowed moves – the physical constraints. In this study, the formulation includes not only the physical constraints, but also challenges and situational tasks. What is the function of these components? Tasks may function as subproblems that are simpler to solve than the complete problem. Challenges may function as subgoals, directing or constraining the search to moves which address those challenges. These two components of the formulation suggest that sensemaking is a continuous process of not only understanding but also simplifying and re-formulating the problem as solving progresses.

Sensemaking leads to a task-oriented formulation

If sensemaking is a process of formulating, it is particularly interesting that tasks are a key part of that formulation. The other two elements – constraints and challenges – are analogous to the constraints and objectives that make up the typical problem formulation, but tasks are an unexpected component. Moreover, tasks are an important component, perhaps central to progress in solving. In a given situation, teams identified a set of tasks, solved them, then identified new tasks in the updated situation. Search was directed not only by objectives but also by the set of tasks contained in the sensemaking frame, or formulation, of each team.

Newell and Simon (1972), in their study of human problem solving, found evidence of search directed by simplified "planning spaces" and sub-goals. Even in abstract problems, like the symbolic and game-playing problems they studied, humans came up with simpler formulations to help them solve more complex problems. The same kind of simplification has been observed in ill-defined real-world problems as well. Klein (1998) suggests experienced decision-makers recognize solve-able sub-problems, and Moldoveanu (2009) suggests that managers confronting strategic problems tend to choose simpler problem formulations as approximations to the real problems. It makes sense that there is a tendency to simplify in formulation, since it is well-established that humans tend to simplify in solving (Simon, 1956; Tversky and Kahneman, 1974). This study suggests that these simplifications in formulation take the form of smaller tasks to be accomplished within the larger problem, and moreover, that they emerge from an ongoing process of sensemaking intertwined with solving.

Goals emerge through sensemaking and solving

Goal-setting is worth special mention here because of its importance in problem-solving and the relative lack of research devoted to the process of setting goals. Goals have been studied for their role in strategic decision-making (Eisenhardt and Zbaracki, 1992) and in motivation (Latham and Locke, 1991), but this research emphasizes the importance of goals without making clear the process by which they are set. In naturalistic decision-making, a field devoted real-world decision-making processes, ill-defined goals were seen as one of the defining characteristics of "naturalistic" problems (Lipshitz, 1993), but research has since focused on how decision-makers use forward-directed reasoning (matching situations to actions) rather than identifying and working toward goals (Lipshitz et al., 2001). My findings suggest that, at least in this problem, both strategies were important: forward reasoning strategies like exploring and solving informed sensemaking, but goals were also identified and, as will be shown later, had a major impact on solving.

Moreover, this study of humanitarian transportation planners provides a glimpse into the process of setting goals. In this multi-objective context, potential goals emerged through sensemaking and solving. The sensemaking maps, such as that in Figure 2-2, suggest two main pathways for the identification of potential goals. First, when teams saw something challenging or undesirable, they thought of a goal to address it. For example, they saw the undesirable situation of idle helicopters for a few days, and came up with a goal of minimizing vehicle idleness. A second pathway was through what we might call a dilemma: teams confronted a decision and needed a goal in order to decide. One team said, "if you have five items here, which is the most important?" Another team discussed various possible prioritization schemes, then said, "We've got to make a value judgment," and pick a goal. These two pathways – addressing challenges and dilemmas – suggest mechanisms by which problem-solving goals emerge through sensemaking and its interaction with solving.

2.6 Findings: Solving

For this problem, solving means fixing aspects of the transportation plan. For example, teams are solving when they decide that a particular truck will make a particular movement, or that some number of trucks will be based in a location. The following quote from my field notes provides an example of how solving proceeded in Team E.

> I went back to the team. They were describing what they're working on. "What we are doing is we are allocating different trucks to go in different directions." Continuing, "We have 20 trucks arriving tomorrow morning." We are sending "2 of them up north to Sugarloaf - because we have some cargo to distribute - and 18 down south" to the areas around Mammoth and Vail. They tally up the capacity of these trucks. "18 [trucks] times 25 [m^3] so we have a capacity of [450] cubic meters" (because each 10-ton truck can carry 25 m^3).
>
> Someone is looking at specific CMRs. "We agreed we take the whole of this shelter number 1 and shelter number 3." "Ah, it's all Federation [a humanitarian organization] - someone will cry." They are talking among themselves as they allocate CMRs to trucks. "Should I put IOM [cargo] now?" Someone says, "I have done 11 trucks. All of these CMRs is in Sugarloaf." I think he means he's moved them to Sugarloaf. ...
>
> Someone suggests, "We should be moving more from Mammoth to Vail." "So we need to keep the helicopter busy. Right now the helicopter is idle." He is suggesting they move cargo to Vail for the helicopter to deliver.
>
> Back on the 40-ton truck movement, someone gives the status. "Remember the 40-ton trucks are gonna wake up tomorrow here [in Sugarloaf] and bring their shipments down here [Jay] and here [Mammoth]."

In this quotation from my field notes, there are many kinds of solving decisions being made (and some evidence of sensemaking, which was discussed in the previous section).

In the first paragraph, the team is allocating trucks among two different areas, sending two to the north and 18 to the south. In the second paragraph, they are loading specific CMRs onto specific trucks. In this case, they have loaded two sets of shelter cargo, both from one organization, and they are deciding what CMR to load next. Then, once a truck movement is planned, they note its updated location. In the third paragraph, someone raises a potential problem – that the helicopters are idle – and suggests they start trying to send more cargo towards the helicopters. In the fourth paragraph, they note that they need to plan next steps for the vehicles they've just loaded and moved.

There are two distinct categories of decisions evident in this quote and throughout my field notes. Assignments of cargo to trucks and trucks to movements were called "dispatch" decisions by many teams. Team E was making dispatch decisions when they were loading shelter cargo onto trucks and moving them to Sugarloaf. However, their decision to send 18 trucks south and 2 trucks north was not a dispatch decision; it had a broader scope, in that it was an allocation of resources rather than a specific movement of cargo and vehicles. This is an example of a strategic decision. A different kind of strategic decision is evident in the third paragraph, when someone suggests a goal of feeding cargo to the helicopters.

The teams did not, in general, make a strong distinction between strategic and dispatch decisions. Indeed, it is evident from the quote above that these two kinds of decisions were made whenever they came up, rather than in a specific order. Nevertheless, the distinction is useful because the two kinds of decisions are made in different ways, as the following analysis will show.

Strategic decisions included assignment of vehicles to bases or routes, the definition of the main transportation corridors, and goals and policies to govern dispatching decisions. Dispatch decisions included allocations of cargo to vehicles and vehicles to movements. Teams must make hundreds of dispatch decisions in order to create a 7-day transportation plan. Strategic decisions, on the other hand, are not strictly necessary; when they are made, however, they often govern the way in which dispatch decisions are made. The sections below elaborate upon each of these two types of decisions, and show how they are made and how they relate to each other.

Strategic Decisions	No. Teams Observed	Sample Quotes
Assignment of vehicles to bases or corridors		
Assign 10-ton trucks to:		"You see what the volume is in various locations and then you can see what kind of fleet you need there."
Deliver from Sugarloaf to surrounding villages	3	"We will eventually have to send the 10-ton trucks to Sugarloaf."
Deliver from Mammoth to surrounding villages	8	"Is Mammoth gonna be our main access point?" "Yes, I think we need a lot of trucks [there]."
Transship from Mammoth to Vail	6	"Just instinct we do 18 here [in Mammoth]. We do 6 here [Mammoth to Vail]..."
Assign helicopters to:		"You might have a helicopter that will be one day in Vail, one day in Mammoth..."
Deliver from Vail to surrounding villages	7	"Here [Vail] we will use the helicopters, and here we [Mammoth] we will use the trucks."
Deliver from Mammoth to surroudning villages	8	"It means that one helicopter will have to stay in Mammoth. It doesn't make any sense to move it."
Transship from Killington to Vail	2	"One helicopter will do rotations between Vail and Killington because the road is not accessible."
Transportation corridors and service along them		
Consider airbridge from Killington to Vail	5	"Why don't we send from Killington to Vail [by helicopter]."
Consider shuttle-style service from Snow	3	"Two ways: you do one truck all the way to do all the movements or you use your logistics hubs [and shuttle between them]."
Policies		
Prioritize by item type - lexicographic	10	"Every time we have a choice, we respect the priority [given by the government]"
Prioritize by item type - weighted mix	5	"Obviously, you can't send all shelter out, you gotta have a mix."
Prioritize by destination - lexicographic	5	"Geographical consideration into where the cargo goes... e.g. Rocky [province] is higher priority."
Prioritize by destination - weighted mix	1	"We can't completely leave out New [the low-priority location]."
Send something for each organization	6	"We tried to send something for everyone."
Send what is there, what will deliver most	3	"Do we fit what we have [available] or go for the most in need?"
Serve closer destinations first	2	"Do whatever's closest first."
Feed cargo to onward transport	6	"No no we have to move materials to Vail to get them to the helicopters."
Avoid idle, empty vehicles	4	"I have an MI-8 [helicopter] I have to use today my MI-8."

Table 2.2: Strategic decisions

2.6.1 Strategic decisions

Three types of strategic decisions emerged from the data analysis: assignment of vehicles to bases or transport corridors, definition of the main transport corridors and service along them, and policies to govern cargo loading and vehicle movement. This section describes the nature of each category of decision, how each of these strategic decisions were made, and how they function in the larger problem-solving efforts of the team. Table 2.2 summarizes the types of strategic decisions and shows how many teams were observed discussing each one.

The data analysis process suggested the importance of these three types of decisions, because they were part of the set of codes describing important elements of the problem. To refine and test the three categories of strategic decisions, I compared and contrasted all the data incidents that included each decision. This process resulted in descriptions of each type of decision. Next, I looked to understand how each decision was made by examining the common precedents or justifications for decisions in each of the data incidents of each type. In this manner, I found what led to each type of decision.

Assignment of vehicles to bases or corridors

One of the ways in which teams organized the use of their vehicles was by assigning them to a base or a corridor. An example was provided in the quote from Team E above, when the team discussed their fleet of 20 10-ton trucks, saying they would send "2 of them up north to Sugarloaf - because we have some cargo to distribute - and 18 down south". Across all 10 teams, only three different assignments are given to the 10-ton trucks in this problem: deliveries from Sugarloaf to destination villages (the "north"), deliveries from Mammoth to destination villages, and transshipment from Mammoth to Vail, the helicopter base (the latter two assignments make up the "south" referred to by Team E). There are also three possible assignments for the helicopter fleet: deliveries from Vail to destination villages, deliveries from Mammoth to destination villages, and flights between Vail and the hub Killington (an "airbridge" between the easily accessible Killington and the difficult-to-access Vail). Table 2.2 summarizes these six possible assignments and shows the number of teams observed considering each.

Note that these assignments are either to bases or to corridors. Vehicles may be based at a hub to make deliveries to destinations surrounding that hub, or they may be assigned to a specific corridor connecting two hubs. Teams did not make a distinction between these two kinds of assignments. Some teams varied the assignments over time. For example, in the first two days of operation, there was very little cargo in Vail for the helicopters to deliver. Many teams either based the helicopters elsewhere for the first two days, or used them to provide an airbridge to get cargo into Vail more quickly. After the first two days, they often changed the assignments so that more helicopters were based in Vail to deliver newly arrived cargo. Given the limited number of 10-ton trucks and helicopters available, allocating vehicles among the various possible assignments had a major impact on cargo delivery capacity.

How were these assignment decisions made? There were two components to the assignment decisions. Teams had to come up with the potential assignment – the idea to base helicopters at Mammoth, for example – and they had to decide how many vehicles to assign to each. The idea of the assignment often came from a "task" in the sensemaking frame, or from examining a "picture" of cargo that must be moved (see Section 2.5 for a discussion of tasks and picturing). For example, once a team noticed that there was a lot to move from Mammoth by helicopter (a task derived from picturing), they often decided to base a few helicopters there (the assignment). In the example from Team E, the team justifies allocation of two trucks to the north because they "have some cargo to distribute". In both of these examples, seeing a task to deliver cargo led teams to assign some vehicles to it.

Another common precedent for assignments was a "challenge", such as assigning helicopters to the Vail-Killington airbridge to avoid having idle helicopters in Vail. Thus, the idea for assignments came from tasks and challenges, two elements of the sensemaking frame.

The number of vehicles allocated to each assignment could sometimes be "instinct", as one team put it, or sometimes was based on additional analysis. One team compared helicopter capacity to the demand for transport in order to figure out how many vehicles were needed at each assignment: "In 6 days we can only manage 30 rotations and we have [cargo to fill] 26 in Mammoth already... before considering what is coming from Snow... it means that one helicopter will have to stay in Mammoth. It doesn't make any sense to move it." Note that this is an example of moving from solving activities back to sensemaking. When confronted with the question of how to allocate vehicles among assignments, the team conducted an "exploring" activity to figure out the demand for cargo. This type of calculation could be considered a sensemaking or a solving activity, because it contributes to both the team's understanding of the problem and its decision about vehicle assignments. In this and other examples, the exploring activity was directed to understand the demand for transport on each assignment – in this case, the total number of helicopter rotations required. In other cases, teams based assignments on the total amount of cargo waiting at a node, or the total amount of cargo requiring transport along a corridor. The number of vehicles on each assignment may emerge from instinct or may be the result of additional exploring, generally to find out the demand for transport on the assignment.

To summarize, assignments of vehicles to bases or corridors is sparked by two elements of the sensemaking frame, tasks and challenges, that suggest the need for some vehicles to be assigned to accomplish the task or address the challenge. Sometimes, after teams see the need for an assignment, they conduct additional analysis or sensemaking to determine the number of trucks to assign based on the demand for transport on the assignment.

Transportation corridors and service along them

Transportation corridors are the major routes along which cargo and vehicles move, which may include multiple modes of travel. While most teams did not explicitly define corridors, most of them did so implicitly, as part of the sensemaking process (see Table 2.1). Transportation corridors are included in this section on solving because in some cases teams made decisions about which transportation corridors to use, or how to provide service along them.

The key strategic decision in this category dealt with whether to include a transport corridor from Snow to Vail using helicopters to provide an "airbridge" from Killington to Vail. As one team put it, "If the helicopters are free, it's better to send the cargo to Killington and make the rotation [to Vail]." When this idea was brought up, discussion usually followed. One team's discussion started, "But you have doubled the cost," followed by, "Sometimes time is more important than cost." Many teams decided it was worth the cost of using helicopters to reach a city, Vail, that could also be reached, but much more slowly, by truck. Other teams decided it was not: another team reported that they "decided to make our whole pipeline [go by truck], even stuff going to Vail... instead of going through Killington."

A second strategic decision in this category dealt with the nature of service along corridors. The main example of this was a "shuttle" service for the 40-ton trucks, considered by a few teams. They considered sending only a few 40-ton trucks away per day, using this "shuttle service because otherwise it takes forever to get back to Snow," as one team put it. They anticipated the need to have capacity to send additional trucks out tomorrow, so they did not want to send all trucks out today.

Both of these strategic decisions were inspired by challenges. In the first case, teams saw the challenge of transporting cargo to the helicopter base Vail, and considered addressing it by creating an "airbridge" transport corridor. In the second case, teams anticipated a challenge of additional cargo requiring transport tomorrow, and considered addressing it by adopting a shuttle service policy. As in the assignments decisions above, strategic decisions were inspired by elements of the sensemaking frame.

Policies

I use the word "policy" to describe a phenomenon that was very common in my observations of the teams: they articulated guidelines or decision rules to help them reach goals. The complete set of observed policies is given in Table 2.2. Most policies dictated how some kind of choice was to be made. For example, in choosing cargo, a policy might dictate choosing shelter cargo before other types; in choosing which village to fly to, a policy might dictate flying to closer villages before those farther away.

Many of the policies dealt with prioritization of deliveries. Prioritization was important because there was not enough transportation capacity to make all the requested deliveries. All teams decided to prioritize deliveries based on the type of item, probably because this policy was mandated by the government in the scenario. The government indicated that shelter was most important, followed by health, water and sanitation (WASH), and food. Many teams laid out what I call a lexicographic policy: "We try to load as much as possible priority 1, then if there is some space" fill it with cargo of other types. A few teams instead decided on what I call a "weighted mix" policy: "Obviously, you can't send all shelter out, you gotta have a mix," something like 50% shelter, and smaller percentages of other types of cargo. While all teams considered the type of item as a prioritization criterion, many teams also considered the destination of the shipment, looking to serve more affected geographical areas first. Usually, the two criteria were considered in combination: as one team put it, "whenever we have a shipment, we consider whether the item is a priority... and the geographic priority." A third criterion was the humanitarian organization requesting the delivery. In the quote from Team E on page 73, teams worried about delivering too much cargo from a single organization. Many teams, like Team K, "tried to move something for everyone."

Another set of prioritization policies are based on more operational concerns. In some cases, teams worried that prioritizing based on item type or destination might limit the total amount of cargo that could be delivered. Instead, someone said, "Our priority is to move as much cargo as we can move with those trucks." Such a policy sometimes translated into sending what was immediately available. A related policy was to serve closer destinations first, enabling more cargo deliveries to be completed in a shorter period of time. Another policy was based on the recognition that cargo had to be transported to forward bases before it could be delivered by helicopters or small trucks. Many teams tried to prioritize cargo that required onward transport; for example: "We need to fill up Mammoth slowly slowly [*sic*] so that when we have the trucks [for onward transport] we can move." Finally, most teams had an explicit or implicit policy to avoid having idle and empty vehicles, meaning that they tried to make use of vehicles when they were free (for example, someone said, "Yeah but today we already have the trucks let's use them") and tried to fill them to capacity whenever they were moved (for example, they wanted "not to send them [the trucks] empty to Sugarloaf").

Policies are particularly interesting because they represent the implementation of goals.

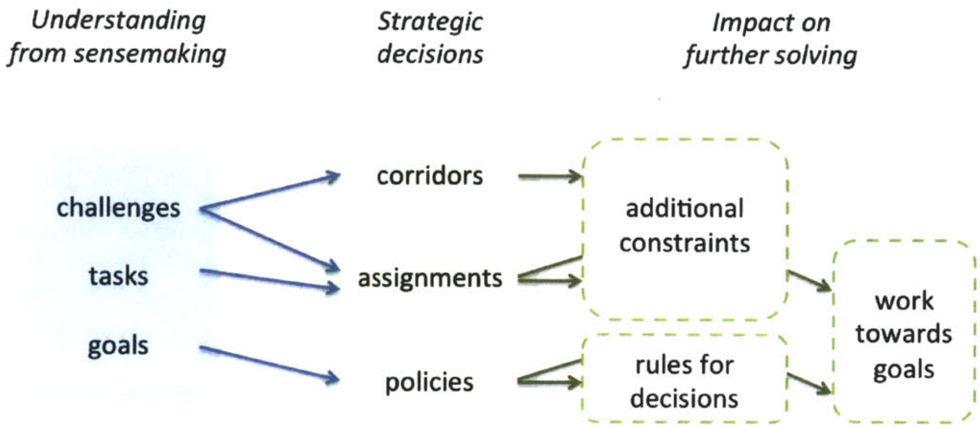


Figure 2-4: Drivers and impact of strategic decision-making

Recall from Section 2.5.4 that goals emerged as part of the sensemaking process. When I looked at the data to understand how policies were formulated, it was very clear that they were formulated based on goals. Consider the policy of feeding cargo to onward transport. On one team, someone said, "Remember that the helicopters in country you pay whether you utilize them or not," and someone else added, "That's why we have to use them at maximum." The team was articulating a goal to use the helicopters effectively. Next, someone asked, "How much cargo do you need in order to keep the choppers busy?" The team formulated a policy of feeding cargo to the helicopters in order to achieve the goal of using the helicopters "at maximum." The same type of process was evident for all the other policies. A goal of serving inaccessible and highly affected areas was translated into a policy of prioritizing by destination; a goal to deliver as much cargo as possible was implemented using a policy of sending what was available or serving closer destinations first. Operational goals, developed through sensemaking, were translated into implementable policies.

Drivers and impact of strategic decision-making

Looking at all three types of strategic decisions, it appears that sensemaking drives strategic decision-making. As shown in Figure 2-4, strategic decisions are made in response to challenges, tasks, and goals, all of which are elements of understanding gained through sensemaking (recall Figure 2-3). Strategic decisions are solving rather than sensemaking activities because they represent choices, not understanding. Teams choose whether or not to include the airbridge as a transport corridor; they choose to allocate vehicles to assignments; and they define and choose one policy or another. Strategic decisions represent one of the direct links between sensemaking and solving.

In fact, most strategic decisions are not strictly necessary in order to solve the problem, since a feasible plan can be created without a prioritization policy, vehicle assignments, or specified service along transport corridors. What, then, is the function of strategic decisions? As shown in Figure 2-4, they have various but related functions.

Assignments and corridors function as additional constraints, in that they restrict the formulation or problem space beyond that defined through sensemaking. Deciding that a helicopter will be based in Mammoth means that "the helicopter stays in Mammoth," as someone said while dispatching helicopters. Policies, on the other hand, restrict not the formulation but the decisions that can be made while searching it, by providing guiding rules or heuristics for making choices, say, among cargo to be loaded. By restricting the problem space and adding rules for searching within it, strategic decisions simplify the solving process.

Some strategic decisions are also a way of working toward goals. It is clear that policies represent a way of implementing goals, because they are rules for making choices that satisfy goals. Assignments can also work toward goals, in that allocating a vehicle to an assignment may push a goal forward. For example, one team discusses prioritizing deliveries to important destinations: "There's a lot of things going to Sugarloaf but we could forget about Sugarloaf if we want to," meaning they can assign all vehicles to other areas and thereby ignore low-priority areas around Sugarloaf. On the other hand, assigning many vehicles to the corridor from Mammoth to Vail might ensure that enough cargo could reach the helicopters in Vail, thus reaching the goal of keeping helicopters busy. Thus, not only policies but also assignments enable teams to work towards goals.

In sum, strategic decisions are driven by challenges, tasks, and goals, all of which emerge from the sensemaking process. Strategic decisions enable simplification of the problem by restricting the problem space and adding rules for searching within it, and they do so in a way that pushes solutions toward goals.

2.6.2 Dispatch decisions

Dispatch decisions are the specific allocation of cargo to vehicles and vehicles to movements. For example, a team might assign a CMR to a truck and assign that truck to leave Snow on the morning of the first day heading towards Sugarloaf. The quote on page 73 provides some examples of dispatching, as someone says, "We agreed to take the whole of this shelter [cargo]," then asks, "Should I put IOM [cargo] now?", and finally provides an update, "I have done 11 trucks. All of these CMRs is in [have been moved to] Sugarloaf." He also notes the next steps, saying, "Remember the 40-ton trucks are gonna wake up tomorrow here [in Sugarloaf] and bring their shipments down here and here." This and other similar quotes exemplify the types of decisions made: selecting cargo to load on trucks, deciding where and when to move trucks, and tracking updated locations as the plan is created.

Note that teams spoke about their plans as if they were being implemented, saying the trucks *are in* Sugarloaf rather than that they *will be* in Sugarloaf on day 2. I have adopted this convention throughout this chapter, to match their language.

In data analysis, it was difficult to isolate incidents of data that dealt with particular dispatch decisions (such as selecting cargo or choosing a destination) because dispatching is an ongoing process. Instead, I tried to understand how dispatch decisions were connected to each other, and more importantly, the process by which they were made. To do so, I analyzed my field notes on two levels. First, I looked at all the data instances, across all teams, that seemed to include dispatch decision-making, and for each one I noted what kinds of decisions were being made and what seemed to precede and follow each decision. I looked across all these mini-process diagrams and consolidated similar decisions, then tried to connect them to each other. For example, I often saw cargo selection decisions preceded by looking at a list of cargo waiting at a node, and followed by dispatching a vehicle on a movement. I also saw dispatching a vehicle on a movement followed by updating locations. This suggested a longer string of connections between these decision types. This resulted in a handful of hypothesized dispatch decision-making processes.

Next, I took a step back to look at the decision-making processes of each team individually. I kept in mind the key results from the first analysis, namely the types of dispatch decisions and their hypothesized connections to one another. I re-read the field notes for each team from start to finish and diagrammed their dispatch decision-making processes, loosely relying on the decision types I found in the first analysis step, but focusing primarily on capturing faithfully the activities of each team.

The final step was to combine the decision-centric and the team-centric analyses. I looked for patterns of decision-making that appeared in both, and came up with two archetypal processes that captured the main flow of decisions in all teams, though each

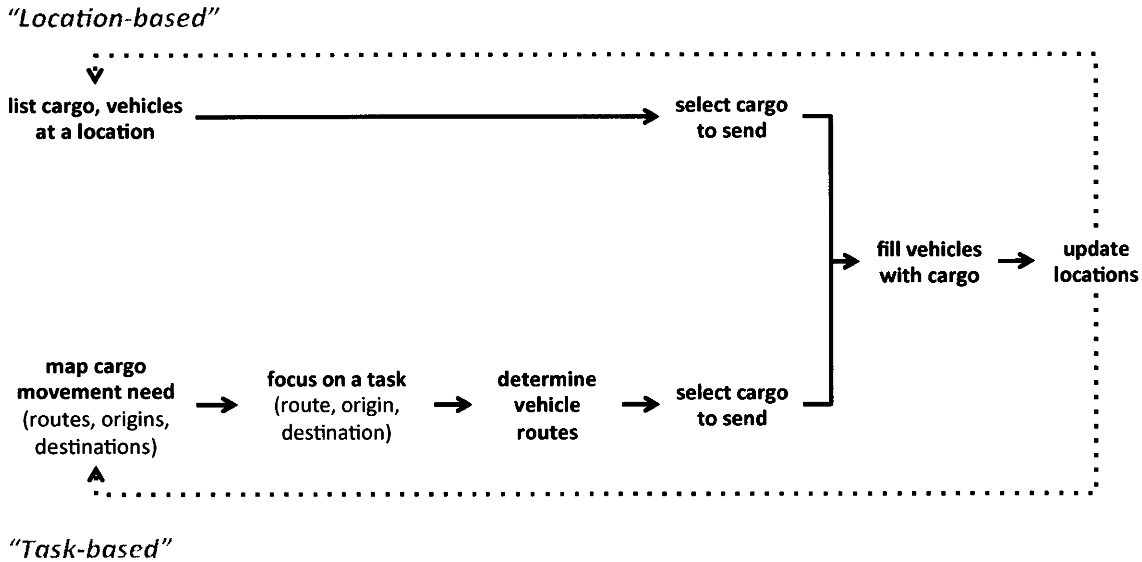


Figure 2-5: Archetypal dispatching processes: location-based and task-based

team showed individual variations on it. I went back to the data one last time to make sure the archetypes fit the team data. This data analysis process resulted in the findings described in the following sections.

Archetypal dispatching processes: location-based and task-based

The two archetypal dispatching processes are shown in Figure 2-5. Each consists of a series of activities, some of which are decisions; a few activities are shared by both processes. They are labeled "location-based" and "task-based" because of the central unit around which plans are constructed. They are described in detail below, followed by a discussion of whether and how the teams followed these two processes. Table 2.3 shows how many teams followed each archetype and variations thereof.

Location-based archetype The location-based process plans around a single location at a time. It is driven by a list or some other representation of the *cargo and vehicles available at a location.* The field notes provide many examples of this activity, akin to the "picturing" described as part of the sensemaking process. For example, one team wanted a "snapshot of where we stand at the moment. In Snow, we have X..." Another team started dispatching by saying, "Let's see what we have here." In another example, one team listed, "at each hub, this is what we have, in order of priority." Next, someone said, they planned

Dispatching Process Variations	No. Teams Observed	Sample Quotes
Heuristic		
Location-based	6	
List cargo waiting at location		"At each hub, this is what we have, in order of priority."
Select cargo going forward		"If there are 5 items here, what is the most important item?"
Task-based	6	
Map cargo movement need		"Biggest concentrations of cargo are in Snow and Mammoth."
Focus on a task		"Today we have to focus on moving as much as possible from Snow to these locations."
Route vehicles		"Let's have this truck [go] to Mammoth then it will make the rotation [back]."
Select cargo		"I'm taking what is in Mammoth that I can take to Vail."
Activities in both heuristics	10	
Fill vehicles and dispatch		"So we have to fill it with other materials that from Mammoth go to Vail."
Update locations		"All of these CMRs is in Sugarloaf... remember the 40-ton trucks are gonna wake up tomorrow
Breakdown		
Air and ground	4	"Can you separate surface and air transport?"
3 vehicle types	3	"Three types of moving assets"
Location	1	"We have responsibility for a hub - this is my hub, this is what's coming and going."
Updating		
Day by day	5	"Go day by day. Then we will know what is arriving in each place each day..."
Update on arrival	1	"On the third day you will have more stuff to move... so you plan for the first and second days and then we will tell you what you have."
Managing CMRs		
Do not split CMRs	2	"It'll be easier to allocate a number of trucks to agencies even if there is some space left."

Table 2.3: Dispatch process activities and variations

to "move through as much as possible."

This latter quote shows how the list was used in the second activity in this process: *selecting cargo to load.* When both vehicles and cargo were listed at a location, cargo was selected from the list to load onto the vehicles. In another example, one person said, "How should we lift it [the cargo]? Maximum priority. We don't have choices we just go." Given the list of cargo, this team will simply go through it in priority order and load it onto the available trucks. Implicit in this step, though not always mentioned by the teams, is the selection of a destination for the vehicle. There is no sense in loading cargo that needs to go in two different directions, so by choosing cargo to load, a destination – or at least the next movement – is implicitly chosen. In this problem, in the first few planning days, much of the cargo went in the same direction, so the choice of direction was not important; but it became increasingly important as cargo moved closer to its final destination.

The next two activities are shared by both archetypes. The cargo selected in the last step may or may not fill the vehicle being loaded. If it does not, teams often try to *fill the vehicle with cargo.* For example, someone trying to fill a truck whose destination was already set said, "So we have to fill it with other materials that from Mammoth go to Vail." Teams were reluctant to leave empty space: one person lamented, "it could have taken something else...", and another called attention to the problem, "Ah, we have an issue. It's not full." Often, vehicles could be filled to capacity, though not necessarily with high-priority cargo, but in other cases they had to be sent partially filled if there was not a full load of cargo going to the same destination.

The final step in the process, *updating locations*, called for noting or tracking the movements that had been planned so far (this step is also common to both archetypes). In the example from Team E, the planner said, "all of these CMRs is in Sugarloaf," and, "Remember the 40-ton trucks are gonna wake up tomorrow here [in Sugarloaf]." This is an example of the crucial step of updating the locations of both cargo and vehicles. Most of the time this updating activity fed back into the first activity in the process, the list of cargo and vehicles at a location. For example, one team said, "We don't know how to follow the total volume in each location," and worked on a tracking format that would enable them to update their location-based lists. Another team was more explicit, noting, "On day 2 we got the same data, same equation: location, volume, priority, resources [vehicles]." The location-based process, then, is a loop. After updating locations, i.e. creating a new list of cargo and vehicles at each location, the process starts again, looking at either a different location or a different time (there is little evidence of how they chose the next location on which to focus, but they generally stepped forward incrementally in time).

One of the key features of this location-based process is the local perspective: only a single location is considered at a time. Planners work from a list of cargo and vehicles available in one city, or one node in the network, and send cargo out from it. This focus on moving cargo *from* the location is another key feature. A discussion from one team provides a good example; they were focusing on the location Snow, in which there are a large number of 40-ton trucks and a lot of cargo. One person noted they had "20 trucks in Snow... nearly all trucks in Snow need to move today." Someone asked, "Ok where is it going?", and another answered, "Doesn't matter where it's going - today, they've got to go." In this discussion, by focusing on Snow, they saw they could send a lot of cargo in the general direction of the affected areas, without worrying about its ultimate destination. To emphasize the point, someone later said that dispatch decisions are "driven by where the 40-tonners [trucks] are today, 10-tonners are tomorrow." This exemplifies the local point of view, in that cargo is moved forward from the locations of vehicles. The implications of the local perspective and the emphasis on pushing cargo forward will be explored later.

Task-based archetype The task-based process centers around "tasks" rather than locations, and around a small set of tasks in particular. These tasks involve moving cargo either on a path, from an origin, or to a destination. For example, Team G noted demand on a path, saying, "Jay [to] Mammoth we have 350 [tons to move]." Another team focused on the need for transport to a destination, noting, "Heavenly total is... we have 155 tons for Heavenly." Yet another example focuses on moving cargo from a place, noting, "There's only one CMR out of Vail." Teams described these tasks, then *focused on a task* around which to make dispatching decisions. For example, one team said, "Today we have to focus on moving as much as possible from Snow to these locations," choosing to focus on the task of moving cargo out of Snow.

The task-based process is driven by what I call a *map of cargo movement need*, meaning an understanding of the demand for transport within the network. As one team described it, they wanted to "define the demand side of this - what do we need to do from where to where?" The maps result from picturing activities, as described in Section 2.5.2. One team described the implications of their 'map', saying, "We have [loads] in Snow... only one [CMR] in Vail," then noted that the "biggest concentrations of cargo are in Snow and Mammoth." The map of cargo movement need thus gave teams an indication of the important tasks to accomplish (as described in Section 2.5.1 on sensemaking).

Once a focus task was selected, teams started *planning vehicle routes* to accomplish it. For example, after Team G saw the demand on a path, "Jay [to] Mammoth we have 350 [tons to move]," they started thinking about how to plan vehicles to accomplish it: "Let's have this truck to Mammoth then it will make the rotation." With vehicles routed, the next step was to *select cargo* to load onto the vehicles. For example, someone looked for cargo going on a vehicle's route, asking, "Can you tell us everything that needs to be moved from Mammoth to Vail?" Selecting cargo happened in much the same way as described above for the location-based algorithm, except that here the vehicle's route had been explicitly chosen in advance, to satisfy some task.

The remaining tasks, *filling vehicles with cargo* and *updating locations*, are the same as those described above as part of the location-based process. Like the location-based process, the task-based dispatching process is also a loop. Updating locations leads to an updated map of cargo movement need and the selection of a new task to focus on. Again, there is little evidence of how teams looped, how often they chose new tasks, or how many tasks were considered at once.

One of the key features of this archetype is the centrality of tasks. Plans are driven by the team's understanding of what must be done, derived through sensemaking; in fact, the first two steps of the task-based process mirror parts of the sensemaking process. The tasks that drive dispatching decisions are a subset of those described as resulting from sensemaking, including only those that require moving cargo on a path, from an origin, or to a destination. The focus on tasks makes this process flexible, in that in one iteration through the process, cargo might be moved out of a congested origin, and in the next, it might be moved along a high-demand path.

Another key feature of the task-based process is that it provides a somewhat global perspective, especially in comparison to the location-based process described earlier. The location-based process can be considered a more myopic special case of the task-based process, in which all tasks are deliveries from a location. In the task-based process, instead of focusing on a single location at a time, teams consider a global map of the need for cargo movement, then pick some task within it on which to focus their planning.

Team processes The location-based and task-based processes are archetypes, in that they describe idealized versions of two distinct patterns of behavior evident in my observations of human problem-solving. Each team had its own variations on these archetypal processes, many of which are described in the next section. Many teams utilized aspects of both archetypes, sometimes trying one and then switching to another, other times using one for planning helicopters and the other for planning trucks. In a few cases, my observations do not provide enough evidence to characterize the team's behavior as following one archetype or another. In general, however, these archetypes provide a useful characterization of key patterns of behavior followed by many of the teams. Table 2.4 provides a brief description of the dispatching behavior of each of the teams.

Variations on dispatching processes

Across all ten teams, there were several common variations on the two archetypal dispatching processes; this section describes these variations.

	Sensemaking	Dispatching
Team A	This team understood the way cargo moves through the network along multi-mode routes. When they realized the helicopters would be idle, they decided to feed cargo to them and to use the Killington airbridge initially. Their initial dispatching was guided by this feeding task, a policy of weighted mix prioritization, and a picture of cargo at origins. As they dispatched, they moved helicopters in an ad-hoc manner among tasks of Killington runs, Rocky deliveries, and Cal deliveries, attempting to use helicopter capacity well. The teams worked separately, with little updating between them until late in the day.	There is not enough data to determine definitively whether they used a location-based or task-based algorithm. There is evidence of thinking in both these ways, so they probably used both strategies. They pictured the cargo by origin and worked from these pictures, tracking CMR locations in the same spreadsheet. They discussed movements, then selected cargo, based on item type and feeding helicopters, then noted when they were free, stepping forward in time. They worked in separate teams for air and ground, with little communication between them.
Team B	This team pictured and organized the data by location, saw that vehicles were available, and started dispatching, in air and ground teams. As they dispatched, they learned more about the network and bottlenecks, and refined their plan for dispatching and the positioning of their vehicles. In the afternoon, they had a major discussion on prioritization and defined clear policies.	Ground dispatching was not observed directly, because it started late, but it appeared they used a combination of location-based and task-based, where the task-based is generally either individual CMRs or from origins. In addition, they assigned tasks to their 10T trucks. Air dispatching appeared entirely location-based, except for decisions about where to base the helicopters. They considered, but did not apparently implement, an idea to draft a plan then fix it later if the prioritization was not adequate.
Team C	This team began by playing with the data, then saw how cargo moves through the network along multi-modal routes. Then they were able to develop a sense of how much cargo had to be moved from origins and along routes. They realized how important the helicopter hub Vail was, and that it could not be reached for several days. As a result, they decided on an alternative route that was faster but more expensive, and they defined prioritization policies. Finally, after developing this sense of the problem, they began dispatching, and as they worked, developed tracking formats.	There is not enough data to describe a dispatch process, but elements of it are clear. They listed CMRs based on their origins and their routes, and they assigned each a number/fraction of trucks. They probably worked around tasks, or at least routed trucks before selecting cargo. They did not appear to break down the problem by air and road.
Team D	This team explored the network, but I saw less clear organization of data until they started dispatching. Once they started, they tracked the process on a wall chart, deciding to move whole CMRs and assign them whole numbers of trucks. As they worked, they figured out how to break down the problem by air and ground, discussed how to link it up again, stated a weighted mix prioritization policy, and later realized they needed to feed cargo to the helicopters.	This team relied on location-based dispatch. They filtered a list of cargo waiting to be dispatched, selected (whole) CMRs to send forward, then determined the number of trucks needed to transport the CMRs, assigned them, and updated locations. Then they looked at the next CMR, location, or day.
Team E	This team pictured the network and allocated vehicles to tasks they saw within it; they refined this allocation as they worked. They pictured the cargo by listing items in each location and hung them on the map, then used them to dispatch. They prioritized by item type, and considered fairness to organizations, along with feeding the helicopters.	This team used the location-based dispatching strategy. They kept lists of cargo in each hub, in item type priority order, and hung them on the map. Then they took one down and dispatched from it, then moved on to the next location, or next day. They updated locations and noted next time it needed to be moved. This was guided by an earlier allocation of vehicles based on the global need for cargo transport.
Team F	This team understood the network relatively quickly, and jumped into dispatching based on lists of cargo and vehicles in each location. They broke down into teams for each of the 3 modes and developed ways of communicating and tracking as they dispatched, eventually passing updated cargo locations among the teams by means of spreadsheets for each location. As they dispatched, they noticed challenges and formulated tasks: for example, limitations on capacity, idle vehicles, and the need to deliver to other locations. They appeared not to worry extensively about prioritization until very late.	They used the location-based dispatching strategy. They listed the cargo in each location, along with the vehicles, then loaded cargo going forward to its next destination. When it arrived, they updated the location information and handed it off to the team doing the next mode.
Team G	They pictured the CMRs and thought through the network, which led to formulation of goals and policies. The key picture was the cargo to go on each route, from which they started dispatching by routing trucks on needed routes. They appeared to prioritize based on item type, organization, and reaching beneficiaries.	This team appeared to use primarily task-based dispatching and the variant in which the task is moving forward from a location. They had the ability to check what cargo was waiting to go on routes or from locations, and this drove their routing of trucks. Interestingly, after a draft plan was finished in the tool, they went back and looked at their final plan, and modified it to even service out among organizations and use all trucks.
Team H	This team started dispatching from a cargo picture and a set of goals and policies formulated early. They ended up breaking down the problem by assigning location managers and truck and cargo trackers. They later defined a more precise weighted mix prioritization policy.	This team used a variant of the location-based dispatch algorithm in which the team was broken down by location. Each location manager dispatched from a list of cargo and vehicles in each location, loaded cargo according to policies, and gave the information to the transporter, who handed it off to the next location manager. They moved forward by day.
Team J	This team came to a basic understanding of the network, and started dispatching quite early. Dispatching was guided by the need to use their capacity well by filling vehicles and using them when they were free. As they worked, they formulated additional tasks, like feeding cargo to onward transport, and they developed tracking methods that highlighted the status of vehicles and CMRs. They reported having prioritized by area and item type.	The data are not completely clear but all incidents of observed work suggest task-based dispatching, along with the variant where some tasks involve moving cargo from a location. They used task-based dispatching processes, which appeared to be driven by deliveries to places as well as from.
Team K	This team spent a lot of time in front of the map that showed the cargo to go on each route, attempting to route trucks to take it. They also stated policies of prioritization, and positioned vehicles, as they discussed the challenges by thinking forward in front of this map.	They used a global map of cargo to go on routes, routed trucks to accommodate the need, then assigned cargo and noted when trucks were free. They considered prioritization policies. My data show their early work, so they certainly started this way, but may not have finished this way.

Table 2.4: Overview of teams' sensemaking and dispatching processes

Breaking down the problem One of the most important variations on the dispatching process was in how each team broke down the work and assigned a role to sub-teams or individuals. Each (full) team had between 9 and 11 people, and only a day to complete a transportation plan, so they were keen to divide responsibilities. Assigning roles had at least two advantages. First, it simplified what each person had to manage, as someone noted, saying, “probably don't need to go in this detail at this stage because it's probably going to be the task of [someone].” Second, it enabled the team to take advantage of specialized skills, as someone noted when suggesting that someone manage helicopters: “He's done aviation before... helicopters get a bit tricky because you have to split their loads.”

Initially, the best way to break down the problem was not clear, but as teams made sense of the problem, they came up with different possibilities. One person suggested, “Can you separate surface and air transport?”, meaning separate teams would plan the movements of trucks and helicopters. A similar alternative recognized that there were “three types of moving assets”, 40-ton trucks, 10-ton trucks, and helicopters, and put separate teams on each. Almost all teams eventually settled on this set of roles, breaking down the problem by mode, but before doing so, they voiced concerns. One team was afraid that “the helicopters are going to see that it doesn't work then we have to go back.” They were worried, and rightfully so, that it would be difficult to make sure that cargo for the helicopters was delivered on time by the ground planning team; more generally, the plans had to link up. Nevertheless, most teams divided the problem in this way, and devised ways of tracking and communicating to support it (described in the next section).

A few teams considered an alternative set of roles: dividing responsibility by location. For example, one suggested "we have responsibility for a hub - this is my hub, this is what's coming and going." In fact, this breakdown may represent many real supply chains, in which a planner at each warehouse is responsible for receiving and sending cargo. However, only one team was clearly seen to adopt this division of the problem.

Updating, tracking, and picturing Once the problem is broken down for multiple planners, there is a clear need for updating each of the planning teams on what the other team has moved. As one person said, “We will have a need to pass information,” and another person said, “they will need a clear idea where what is.”

To facilitate updating, every team settled on the same basic way of moving through the problem: forward in time, one day at a time. For example, one person suggested to "go day by day. Then we will know what is arriving in each place each day and what trucks are available in each place each day and then you can dispatch." Within this forward-solving framework, many teams updated each other at the end of each day. For example, someone asked, "Now it's day 2. Where are the trucks?" Others recognized that they did not need to update as frequently, because the helicopter team would not receive any new cargo until several days later. As one road planner told a helicopter planner, "On the third day you will have more stuff to move... so you plan for the first and second days and then we will tell you what you have."

Having described when (in the planning horizon) updates are made, the next question is how they are accomplished. Some teams simply asked or told their teammates what they had moved. For example, someone asked for "whatever you've got in Jay on the 16th... from this he will be able to dispatch when they are free." However, most teams eventually developed a common tracking format, and used it to pass updates around; for example, one person suggested, "Let's agree on the format that's compatible so we can merge [the separate plans] later." Some teams passed around Excel documents containing updated locations, while others kept a master tracking sheet. Tracking systems varied widely across teams. Some teams tracked everything within the list of CMRs they were initially given. Others developed their own format for a master sheet describing all movements. Still other teams developed a location-based tracking system, in which lists of cargo and vehicles in each location were maintained for each day, and updated when cargo or vehicles moved. Recall that four teams were provided with a simple tracking tool that I developed; only two teams chose to use this to track movements.

The picturing activities that drive the dispatching process (listing or mapping cargo to be moved) also varied widely across teams. The type of picturing used by each team depended on their tracking and updating formats and also on whether a team was using a location-based or task-based dispatching process. For example, one team using a location-based process printed out lists of cargo waiting at each location and hung them on the map. They took down a location's list and acted as planners for that location, deciding what cargo to send forward. Then they updated the master sheet with the new locations and printed out a new set of lists to hang on the map. They described their process as using a "piece of paper on each location with what is there," "because then you can sit and do your own decisions." Task-based processes generally depended on keeping a master list of locations and filtering and querying that list to understand the next important tasks. While picturing formats varied widely across teams, the goals and results of the activity were essential and consistent ingredients in the dispatching process.

Managing cargo A minor but potentially important variant on dispatching processes was in the rules, implicit or otherwise, for splitting CMRs. Many CMRs were too large to fit on a single truck, while many were small enough to fit in trucks already carrying other loads. Many teams simply attempted to fill trucks with cargo, without worrying about whether the CMR fit in its entirety, thus splitting many CMRs. Some teams, on the other hand, allocated a number of trucks to a given CMR rather than allocating CMRs to trucks. In this manner, CMRs were never split. For example, while dispatching, someone said, "CMR 3 will take 6 trucks. So share in into 6 10-ton trucks." This strategy could result in wasted space, but one person said, "It'll be easier to allocate a number of trucks to agencies even if there is some space left." The latter method may be easier to track, but the former may enable the delivery of slightly more cargo, so this variation can impact the effectiveness of the transport plan.

Impact of sensemaking and strategy on dispatching

In earlier parts of this chapter, sensemaking and strategic decision-making were described as integral parts of making a transportation plan. The actual plan is the result of dispatching, so it is worth asking how sensemaking and strategic decision-making impact the dispatching process. Most of the impacts have already been discussed. First, the understanding of problem constraints that comes from sensemaking, along with additional constraints resulting from strategic decisions, constrain the problem space in which dispatch decisions are made (see Sections 2.5.4 and 2.6.1). Second, the sensemaking activity of picturing leads to sets of tasks that drive the task-based dispatching process. Third, picturing also suggests different formats for tracking, which become the starting points for both task-based and location-based dispatching processes. Sensemaking and strategic decision-making thus impact dispatching by influencing tracking, directing search, and constraining the space of possible decisions.

The last important impact of sensemaking and strategy on dispatching is to push toward solutions (transportation plans) that satisfy goals and meet challenges. Goals and challenges are surfaced during sensemaking (see Section 2.5.1) and turned into implementable policies during strategic decision-making (see Section 2.6.1). These policies are implemented during dispatching. Policies impact nearly every step in both dispatching archetypal processes by guiding decision-making.

Planners considered policies when they were selecting cargo to load on vehicles, in both the location-based and task-based archetypes. One team described how they would select cargo: "So whenever we have a shipment we consider whether the item is a priority... and the geographic priority." This is a statement of a policy, and it is evident that most teams followed this kind of a policy. For example, while selecting cargo to send, one person looked through the waiting CMRs and selected those that were a priority. "Ok, let's do priority. Ok? Blankets still priority? Yes. Hygiene kit? Let us see what else is there. Jerry can? Yes. Plastic sheeting?" This is an example of prioritizing by item type. Teams also considered the organization sending the cargo. For example, someone tried not to send too much cargo from one organization, saying, "First rotation was UNICEF. The second... UNICEF. The third probably cannot be UNICEF. But there is nothing else from here today." An alternative policy dictated taking what was available; for example, someone asked, "What do we have in Vail that we can take immediately?" Yet another policy dictated selecting cargo that required onward transport by helicopter, and one team implemented it, choosing to fill the helicopter base, Vail, rather than taking high-priority shelter cargo, in the following discussion: "You're doing shelter again...," "No, I'm taking what is in Mammoth that I can take to Vail...," "But there is WASH there," "but I am taking it to Vail." All of these examples, along with many others, show that policies dictated the selection of cargo in each of the dispatching processes.

Policies also influenced the selection of tasks in the second step of the task-based process. One policy dictated feeding cargo to the helicopters. Someone therefore told the road team, "Your [job is] to make sure to deliver in [the helicopter base] on time." Another policy dictated sending the maximum amount of cargo. In support of that policy, someone suggested focusing on the 40-ton movements, meaning the task of moving cargo from Snow to Mammoth: "40-ton is your chief asset. Allows you most clearance [of cargo]," so "don't [mess] around with your 40-tonners to accommodate your helicopters or 10-tonners." In an attempt to implement a geographic prioritization policy, someone asked, "Can we start sending stuff already to the priority villages?". Clearly, the selection of tasks on which to focus was influenced by policies.

Determining vehicle routes may or may not be constrained by the task selected in the preceding step of the task-based process. If the task is to make deliveries on a path, of course, the route is determined by the task. However, when the task is more broadly defined, like delivering from Vail to surrounding villages, there is some flexibility in routing, and thus it may be guided by policies. For example, in an implementation of a policy to serve close helicopter destinations first, someone checked the distance to a proposed destination, asking, "Where is Northstar - is it more or less than 75 [km]?" Thus, policies can also influence the choice of vehicle routes.

Finally, policies influence how cargo is pictured. Unlike the selection of cargo, tasks, and routes described above, cargo picturing is not a choice; rather, it influences later choices. For example, policies influence picturing in that the lists of cargo, in the location-based process, are often ordered by priority. One team said to create a list for each of the "5 locations, sorted by priority and then where it's going." If the list is ordered by priority, the next activity, cargo selection, will probably also focus on higher-priority cargo. In another example, one team wanted to make sure they used a geographic policy in addition to other policies, so they amended their tracking format: "We can add a column... for geographic priority." Policies thus have a more subtle influence on dispatching as well, in their impact on the picturing that underlies the dispatching process.

Considering the impact on dispatching from sensemaking and strategic decision-making provides a more complete picture of problem-solving in the humanitarian transportation planning problem. There are two major ways in which dispatching is impacted. First, the space of possible solutions (sets of dispatch decisions) is constrained by an understanding of the physical elements of the problem, such as the transportation network and the transport demand, and by strategic decisions that assign resources to particular tasks. Second, the direction of search within the problem space – the specific decisions made within the dispatching process – is directed or guided by picturing, by tasks that result from picturing and sensemaking, and, most importantly, by goals and policies formulated through sensemaking and strategic decision-making.

2.6.3 Insights and implications of the solving models

This model of problem-solving in humanitarian transportation planning supports existing theories about how humans move through problems. Problem-solving is often conceived as search through a space of possible moves toward a solution (Jonassen, 2000; Newell and Simon, 1972). Human searches tend to be limited, looking for a good-enough rather than optimal solution (Simon, 1956), using heuristics as a shortcut to a solution (Tversky and Kahneman, 1974), or generating just a single possibility (Hale et al., 2006; Klein, 1993). The dispatching processes found in this study represent heuristic search routines within a problem space defined through sensemaking and further refined by strategic decisions.

The findings from this study on transportation planning illuminate how humans solve problems in a particularly urgent and ill-defined context. They use greedy search strategies directed by policies, and they simplify the problem space using assignments and tasks. Each of these insights is explored in the paragraphs below. The general findings are depicted conceptually in Figure 2-6. Teams had some understanding of the problem space of possible solutions, which may be restricted or simplified based on strategic decisions which eliminate some set of solutions (such as assigning a vehicle to a specific region). Their decisions represent movement through that space of solutions. At each decision point, the teams saw some space of possibilities for their next "move" (assigning a vehicle to a movement, or cargo to a vehicle). In most cases, they considered only a subset of all possible moves: for example, they might look only at one location at a time, or at moves that accomplish a particular task. Within that space of possibilities, the selection of the move was guided by policies, such as prioritization policies. The following paragraphs explain and expand upon the concepts depicted in this conceptual diagram.

Greedy search directed by policies

The dispatching processes found in humanitarian transportation planners' behavior resemble greedy search strategies. Greedy algorithms proceed by selecting the best choice at each step. In each of the dispatching processes, the choices at each step are governed by policies derived from goals, such as selecting cargo of the highest-priority item type. As a result, the choices at each step are those that best fit the ultimate objective. Greedy search has long been considered a model of human problem solving (e.g. Newell and Simon, 1972), so

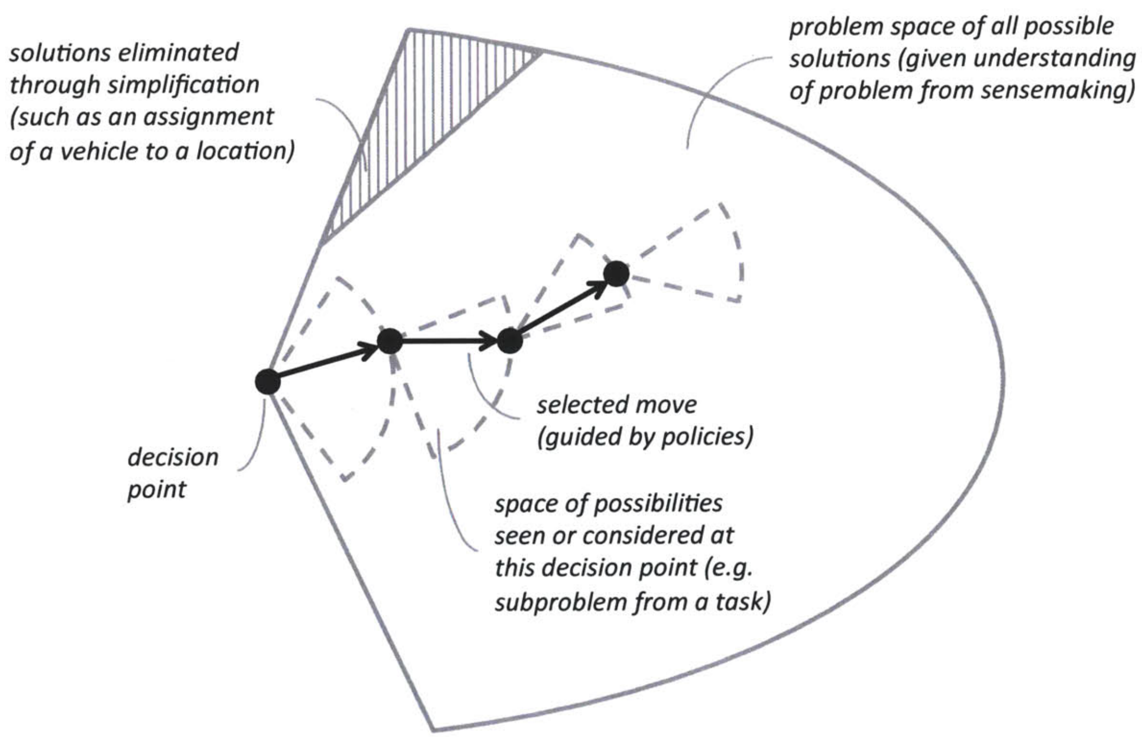


Figure 2-6: Conceptual depiction of search process teams followed. At each decision point, they saw some space of possibilities for their next move (assignment of vehicle and cargo to a movement), and chose a move based on guiding policies. The search occurred within a problem space defined and simplified through their understanding of the problem.

it is not surprising that humanitarian logisticians used such a strategy. In greedy search, however, the details are important. The specific choice of objectives or the formulation of the problem can impact the quality of the resulting solution.

The two archetypes – location-based and task-based – represent different kinds of search within the same space of possible solutions. In the location-based archetype, cargo is pushed forward from wherever vehicles are available, and policies govern which cargo is pushed. The selection of the objective thus impacts only the type of cargo moved, but has little influence on where and when it is moved. Of course, movements may be directed through strategic decisions, such as assigning vehicles to certain areas, but this is outside the pure dispatching process. The task-based dispatching process, on the other hand, does more than simply push cargo forward. It begins with a more global view of the need for cargo movement, and continues by selecting a task from that global view; this choice is governed by some policy. Subsequent choices of vehicle routes and cargo to send are also governed by policies. As a result, this task-based process may have more flexibility to satisfy the demand for cargo transport in accordance with the objectives of the problem.

Another interesting aspect of this model of human problem-solving is the use of policies to govern each solution step. Many policies were formulated by teams, and each policy, if used consistently in dispatching, could lead to different kinds of transportation plans. The evidence from this study suggests that no single policy was used consistently by any team, which makes sense in a multi-objective context like humanitarian aid. However, even if one single policy were used, it is not clear how well the resulting transportation plan would satisfy the objective on which that policy was based. Probably, there are better policies or combinations of policies that would lead to better solutions. Nevertheless, using policies to govern each search step probably insures a better result than a blind or less-informed search, and such policies are certainly easier for humans to implement than complex calculations about what choices have the best chance of meeting their multiple objectives. Thus, one interesting insight from this study is that people transformed multiple goals into a set of policies that could more easily direct search through a problem space.

A second insight from this study is the specific nature of the policies that direct search. One of the weaknesses of greedy algorithms is their inability to "look ahead" to future solution steps. Policies that dictate selecting high-priority cargo do not look ahead, but at least one of the policies developed by the teams does look ahead to some extent: the policy to feed cargo to helicopters. By selecting cargo intended to keep the helicopters busy, teams are thinking about future problem steps, and quite possibly ensuring that the long-term plan delivers more cargo.

To sum up, humanitarian transportation planners use search strategies that resemble greedy heuristics, consistent with existing theories of human problem-solving. Both location-based and task-based dispatching processes exhibit greedy search strategies, but the task-based process is more responsive to overall need for cargo movement while the location-based process is more opportunistic about using vehicles. This greedy search is guided by policies developed by teams based on objectives, which is a simple method of implementing goal-directed search. The specific nature of those policies will impact the quality of the resulting solution.

This discussion prompts us to ask which policies lead to better solutions, but the study described in this chapter did not enable us to answer this question. It was practically impossible to measure the success of each team's transportation plan, because they were only required to report the results verbally within a meeting. As a result, teams delivered plans with varying degrees of completeness and feasibility, whose quality could not be measured. In addition, it would be difficult to claim that any given decision process was the main contributor to a team's success in planning. There were too many confounding factors that contributed to the quality of plans, such as leadership or team dynamics or the use of a number of different decision-making processes or policies. As a result, teams' plans were not analyzed for their quality, and the performance of decision processes and policies was not examined within this study. Instead, Chapter 4 of this thesis uses models to examine the performance of the decision-making methods found in this study.

Simplification through assignments and tasks

Simplification appears to be an important characteristic of the problem-solving models derived from humanitarian transportation planning. In dispatch decision-making, planners focused on one location at a time or one task at a time, thus working within a simplified view of the problem. Strategic decisions, too, resulted in simpler formulations, especially those assigning vehicles to bases or corridors.

Humans simplify in solving a wide variety of problems, and the simplifications take different forms. In search-oriented problems, humans make limited searches (Simon, 1956), and in urgent problems, including crisis response, humans may search only as far as the first option (Hale et al., 2006; Klein, 1993). In judgment under uncertainty, in inventory management, and in many other problems, humans use different kinds of heuristics as shortcuts (Tversky and Kahneman, 1974; Croson and Donohue, 2002; Schweitzer and Cachon, 2000) or miss important aspects of the problem (Sterman, 1989). There are also many problem-specific models incorporating various kinds of simplifications (Jonassen, 2000; Lipshitz, 1993), such as heuristics specific to the traveling salesman problem (MacGregor and Chu, 2011). Given this wide array of examples of simplification in human decision-making, it is not surprising to find the same tendency here; what is interesting is the form those simplifications take.

Tasks are one of the central concepts around which simplification occurred in transportation planning. In the task-based dispatching process, each iteration focused on a single task rather than the entire problem, thus enabling decisions within a simpler sub-problem. However, by selecting those tasks from a global map of cargo movement demand, a more global perspective is maintained. Recall that tasks were also a key element of the sensemaking frame, and there they had a simplifying function as well (see Section 2.5.4). In this problem, it appears, simplification occurs through selecting a series of subproblems, conceived as tasks, which are considered one after another in a greedy search strategy.

There is an element of simplification in strategic decision-making as well, most clearly in the assignment of vehicles to bases or corridors. By giving a vehicle a particular assignment, the problem space is reduced in an easily understood way. Such assignments do not naturally arise from optimization algorithms, but human planners may make this kind of simplification in order to make problems easier to solve. For example, human planners for the UPS network made pickup and delivery routes the same, while an optimization model improved it with an asymmetric solution (Armacost et al., 2002). There may be a complicated relationship between vehicle assignments and the quality of the solution. As the UPS example suggests, assignments reduce the solution space and may eliminate the best possible solution. However, human problem solvers might not find that optimal solution anyway; in these cases, vehicle assignments may limit the problem space in ways that also eliminate some very bad solutions, thereby directing search toward pretty good, if not optimal, transportation plans. In this problem, assignments have a simplifying function, and may also direct search toward better solutions.

Simplification is thus a key element of human problem-solving in the humanitarian transportation planning problem, as in many other problem-solving environments. In this study, simplification centered around tasks and assignments. Assignments reduce the problem space by limiting the possible space of solutions, and they do so in a way that is easily understood by planners. In fact, the the assignments usually allocate vehicles to tasks, such as delivering from a base or along a busy corridor. Tasks, then, are central to the way people conceive and solve the problem. Tasks can be easily understood, because they are simple and intuitive to state. They are a product of sensemaking, and they play a central role in assignments (strategic decision-making) and in dispatching. Tasks, then, represent a simple construct around which the complex humanitarian transportation planning problem can be conceived and simplified.

2.7 Conclusions

This chapter set out to accomplish two goals. The first goal was to precisely describe human problem-solving in humanitarian transportation planning, in order to identify strengths and weaknesses that could ultimately lead to better planning methods. The second goal was to understand more generally how humans solve urgent problems that require serious effort in both sensemaking and solving. Humanitarian transportation planning provided an extreme case in which to learn how humans make sense of and solve complex problems in ill-structured environments.

Towards the first goal, this research has provided a detailed description of human problem-solving in a simulated but realistic transportation planning problem. Figure 2-7 summarizes key aspects of the sensemaking and solving processes found in this chapter. Planners made sense of the problem by moving (unconsciously) between exploring, understanding, and solving activities (as depicted earlier in Figure 2-3). Through this sensemaking process, problem-solvers surfaced challenges and goals, saw physical constraints and flows, and accumulated a set of tasks or an idea of what to do. This frame, shown in the top row of Figure 2-7, represented their understanding of the problem and guided further sensemaking and solving. Thus, sensemaking appeared to be an ongoing process of understanding, formulating, and simplifying the problem.

Solving was distinguished from sensemaking because it involved fixing aspects of the

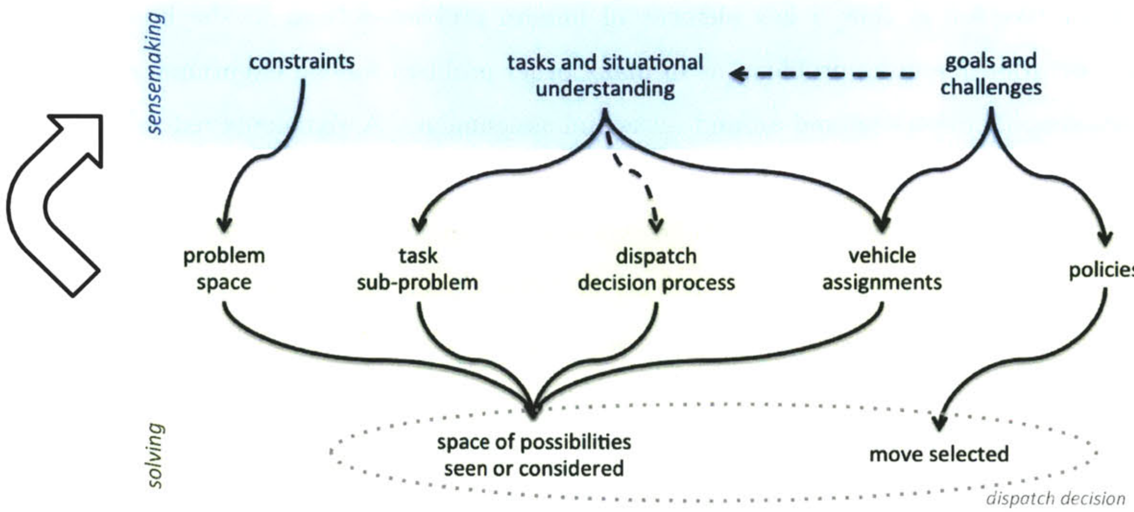


Figure 2-7: Summary of key findings. The sensemaking frame, depicted in the top row, influences dispatch search processes, strategic decisions, guiding policies, and the problem space and subproblems within which solving occurs (middle row). These elements in turn influence the space of possible dispatch decisions seen or considered and the move (dispatch decision) ultimately selected.

transportation plan. Solving occurred on two levels. Strategic decisions determined policies and vehicle assignments. Dispatch decisions, involving the allocation of cargo to vehicles and vehicles to movements, followed location-based or task-based heuristic processes: cargo was loaded and vehicles routed either from a location or along a given task, then the process was repeated at another location, for another task, or at another time. These two processes were described earlier in Figure 2-5, and the sequence of decisions was depicted conceptually in Figure 2-6.

Figure 2-7 highlights the specific mechanisms by which sensemaking influenced solving. Strategic decisions, including vehicle assignments and guiding policies, were driven by elements of the sensemaking frame: goals and tasks. The dispatch decision processes (location-based or task-based) were influenced by the situational understanding held by each team. Constraints and tasks defined the problem space of possible solutions, or subproblems within it. These higher-level concepts are directly influenced by the sensemaking frame, as shown in the second row in Figure 2-7.

The dispatch decisions that make up a transportation plan were made by repeatedly choosing a move from within a space of possibilities (such as which cargo to load on a given truck at a given time). The bottom row of Figure 2-7 depicts how these decisions were influenced by strategic decisions and, ultimately, sensemaking. The space of possibilities seen or considered was influenced by the team's understanding of constraints (the problem space), by the tasks they had in mind (and the resulting sub-problems they considered), by the way they made dispatch decisions (location-by-location or task-by-task), and by strategic decisions like vehicle assignments (which, in turn, were influenced by tasks and goals). The selection of a specific move within that space of possibilities was governed by policies derived from the goals and challenges that emerged from sensemaking. Thus, dispatch decisions were guided by policies and made within a sub-problem influenced by sensemaking and strategic decisions.

In assessing the strengths and weaknesses of human problem-solving on this task, it is helpful to contrast the human approach with that of optimization models which, given a formulation, can search the solution space for the best solution. One issue in building models for this type of problem, which highlights a potential strength of human problem-solvers, is that the formulation itself is not clear. An optimization formulation requires only an understanding of the physical constraints, which in this problem are actually available at the start of the problem, even though they are not well understood by the problem-solvers. (In real humanitarian contexts, these constraints might be changing as infrastructure problems arise and are solved, and humans might be able to predict the trajectory or understand the implications of this dynamic. In this scenario, however, constraints were unchanging, so this potential human advantage was not explored.) In addition to constraints, an optimization model requires an objective, and the right objective was by no means clear for this problem, as in most humanitarian contexts. A variety of objectives, or goals, are stated by the teams, and they seem to consider most of them as they make their plans. There is not enough understanding of the goals of humanitarian aid delivery to formulate an objective that emulates those of the decision-makers in this problem. Moreover, goals appear to depend quite strongly on the specific humanitarian context, so they may be difficult to formulate ahead of time. Here, then, is a strength of human problem-solvers: their ability to understand the goals and challenges of a given humanitarian context through sensemaking and transform them into implementable policies to guide problem-solving.

There are also a number of potential weaknesses in human problem-solving processes, which are especially evident in comparison with optimization approaches. Instead of searching the entire problem space for the best solution, humans rely on a greedy search guided by policies. The quality of the resulting solution depends to a large extent on how well the policies are able to guide the search. Some policies probably lead to better solutions than others, and task-based dispatching may have more flexibility to reach better solutions than location-based dispatching. Policies and search heuristics are therefore potential weaknesses of human problem-solving approaches. With further analysis to develop better policies and heuristics, training or decision support could enable better transportation planning in practice. A second weakness is in the "picturing" activities that function both as exploring activities in sensemaking and as the start of the dispatching processes. Teams struggled to understand and manage data on what they needed to move and where they had already moved it. Different pictures of data led to different understanding of what to do, most strikingly evident in the difference between location-based and task-based dispatching, which are driven by fundamentally different pictures (lists at each location and a global map of cargo movement demand, respectively). If different picturing formats drive better or worse decision-making, as my findings suggest, this is a potential weaknesses of human decision-making processes. Decision support systems, for tracking and understanding cargo movement demand, might be developed to address this weakness.

A second goal of this research, beyond identifying ways to improve humanitarian transportation planning, was to understand more generally how humans solve complex, ill-defined problems in urgent environments. Sensemaking and solving would both be essential in such a problem, thus providing an opportunity to understand how these processes interacted. Humanitarian transportation planning was an extreme case of such a problem. It was urgent, in that logisticians felt pressure from partner agencies, donors, and the media to deliver aid to those in need; it was complex, in that the planning problem was a difficult combinatorial search problem; and it was ill-defined, in that the problem formulation itself was not clear. The formulation was unclear in part because of the difficulty of understanding the physical constraints of the problem, but perhaps more importantly because of the dynamic, information-poor, and multi-objective environment of humanitarian response. It turned out that the lack of clear objectives was a very important challenge in solving this problem, but the dynamic environment and lack of information did not matter very much, possibly because this particular problem scenario did not emphasize those aspects. In this scenario, the main challenges were the complexity of the planning problem itself, the difficulty of understanding the physical constraints of the problem, and multiple unclear objectives. Translating these challenges into more general language, the solution consisted of a complex series of moves, the allowed moves were not obvious, and it was unclear what the objectives were and how to evaluate how well a solution satisfied an objective. With unclear formulations and complex solutions, serious effort was required in both sensemaking and solving. My study of humanitarian transportation planners suggests general characteristics of human problem-solving in this type of problem.

As expected, sensemaking and solving are both essential, and they are intertwined and inter-dependent activities. Sensemaking is a process of understanding, formulating, and simplifying the problem space in which solving proceeds, including the definition of goals. Solving involves two kinds of decisions: strategic decisions that further restrict the problem space and outline goal-based policies to guide search, and decisions about moves (here called dispatch decisions) that are made through a greedy search directed by policies. Moreover, two important insights emerge from this model which deserve further attention: the development and implementation of objectives and the centrality of tasks.

Goals or objectives are essential in successful problem-solving, but little is known about how they are understood and implemented by human problem-solvers, especially when it is unclear what the goals are and how to evaluate how well they are satisfied. My findings show that goals are surfaced during sensemaking, through two main pathways. Goals are formulated to address challenges or undesirable situations, or in response to dilemmas that require a goal to decide. Multiple goals may form part of the team's understanding of the problem. These goals are translated into implementable policies, and these rules direct later decision-making by guiding it in the direction dictated by the goal. This type of policy gets around the problem of evaluating how well a given solution satisfies a goal by simply sending search in the right direction. This study of humanitarian transportation planners thus offers a glimpse into the process of formulating and implementing goals in unclear, multi-objective problems.

A second insight from this research is the centrality of tasks in human problem-solving. I have used the word "task" loosely, defined by examples such as sending cargo from a location or making deliveries in a region. Tasks are short-term pieces of work that can be simply and intuitively described. They are sub-problems, a common concept in problem-solving, but they are a specific kind of sub-problem that can be easily understood by human problem-solvers, because they are intuitive and do not rely on complex constructions like network diagrams. These simple tasks appear to be central to both sensemaking and solving. Problems are understood in terms of a set of tasks, along with the other components of the sensemaking frame. Tasks give problem-solvers a sense of what to do and appear to drive problem-solving forward. If sensemaking is a process of formulating, the formulation includes tasks as a central component. It is then not surprising that tasks are also central in solving processes. Strategic decisions allocate vehicles to tasks, thereby further restricting the formulation. Most importantly, one of the two archetypal dispatching processes centers around tasks: solvers select an important task, work on it, then select another important task. In essence, sensemaking leads to a set of tasks to accomplish, and solving attempts to accomplish those tasks. As sensemaking and solving proceed, tasks are accomplished and new tasks are formulated, eventually leading to a complete solution to the problem. This study of transportation planning suggests that simple tasks may be a key driver of intertwined sensemaking and solving in human problem-solving.

Through this study of humanitarian transportation planning, I have developed a general model of human problem-solving in complex, ill-structured problems, highlighting how sensemaking and solving are intertwined, the process of formulating and implementing goals, and the centrality of simple tasks in both formulating and solving. I have also identified specific strengths of human transportation planners, particularly their ability to formulate goals, along with key weaknesses, including data management and the policies that govern dispatch planning. By building upon the strengths and supporting the weaknesses of human problem-solvers, humanitarian aid delivery can be improved in practice.

References

Armacost, A. P., Barnhart, C., and Ware, K. A. (2002). Composite variable formulations for express shipment service network design. *Transportation Science*, 36(1):1–20.

BBC (2010). Logistical nightmare hampers aid effort. BBC News, 22 Jan 2010.

Ben-Zion, U., Cohen, Y., Peled, R., and Shavit, T. (2008). Decision-making and the newsvendor problem: an experimental study. *Journal of the Operational Research Society*, 59(9):1281–1287.

Bendoly, E., Donohue, K., and Schultz, K. L. (2006). Behavior in operations management: Assessing recent findings and revisiting old assumptions. *Journal of Operations Management*, 24:737–752.

Bolton, G. E. and Katok, E. (2008). Learning by doing in the newsvendor problem: a laboratory investigation of the role of experience and feedback. *Manufacturing and Service Operations Management*, 10(3):519–538.

Campbell, D. T. and Stanley, J. C. (1966). *Experimental and Quasi-Experimental Designs for Research.* Houghton Mifflin.

Carroll, J. S. and Johnson, E. J. (1990). *Decision Research: A Field Guide.* Sage.

Caunhye, A., Nie, X., and Pokharel, S. (2012). Optimization models in emergency logistics: A literature review. *Socio-Economic Planning Sciences*, 46:4–13.

Corbin, J. and Strauss, A. (2008). *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory.* Sage, third edition edition.

Crawford, S. and Wiers, V. C. (2001). From anecdotes to theory: A review of existing knowledge on human factors of planning and scheduling. In MacCarthy and Wilson (2001), chapter 2.

Croson, R. and Donohue, K. (2002). Experimental economics and supply-chain management. *Interfaces*, 32(5):74–82.

de la Torre, L., Dolinskaya, I. S., and Smilowitz, K. (2011). Disaster relief routing: Integrating research and practice. *Socio-Economic Planning Sciences*, (in press).

Eisenhardt, K. M. (1989). Building theories from case study research. *Academy of Management Review*, 14(4):532–550.

Eisenhardt, K. M. and Zbaracki, M. J. (1992). Strategic decision making. *Strategic Management Journal*, 13:17–37.

Emerson, R., Fretz, R., and Shaw, L. (1995). *Writing Ethnographic Fieldnotes.* University of Chicago Press.

Freeman, A. (2010). The logistics response training 8 - surviving brinland. Logistics Cluster Blog, 18 Oct 2010.

Gigerenzer, G., Todd, P. M., and the ABC Research Group (1999). *Simple Heuristics That Make Us Smart.* Oxford University Press.

Glaser, B. and Strauss, A. (1967). *The discovery of grounded theory.* Aldine de Gruyter New York.

Golden, B. R. (1992). The past is the past – or is it? the use of retrospective accounts as indicators fo past strategy. *Academy of Management Journal*, 35(4):848–860.

Grant, T. J. (1986). Lessons for o.r. from a.i.: A scheduling case study. *Journal of the Operational Research Society*, 37(1):41–57.

Grünert, T. and Sebastian, H.-J. (2000). Planning models for long-haul operations of postal and express shipment companies. *European Journal of Operational Research*, 122:289–309.

Hale, J. E., Hale, D. P., and Dulek, R. E. (2006). Decision processes during crisis response: An exploratory investigation. *Journal of Managerial Issues*, 18(3):301–320.

Jackson, S., Wilson, J. R., and MacCarthy, B. L. (2004). A new model of scheduling in manufacturing: Tasks, roles, and monitoring. *Human Factors*, 46(3):533–550.

Jonassen, D. (2000). Toward a design theory of problem solving. *Educational technology research and development*, 48(4):63–85.

Kahneman, D., Slovic, P., and Tversky, A., editors (1982). *Judgment under uncertainty: heuristics and biases.* Cambridge University Press.

Kalkanci, B., Chen, K.-Y., and Erhun, F. (2011). Contract complexity and performance under asymmetric demand information: an experimental evaluation. *Management Science*, 57(4):689–704.

Kim, D., Barnhart, C., Ware, K., and Reinhardt, G. (1999). Multimodal express package delivery: A service network design application. *Transportation Science*, 33(4):391–407.

Klein, G. (1993). A recognition-primed decision (rpd) model of rapid decision making. In Klein et al. (1993), chapter 6.

Klein, G. (1998). *Sources of power: how people make decisions.* MIT Press.

Klein, G., Orasanu, J., Calderwood, R., and Zsambok, C. E., editors (1993). *Decision Making in Action: models and methods.* Ablex.

Klein, G., Phillips, J. K., Rall, E. L., and Peluso, D. A. (2007). A data-frame theory of sensemaking. In Hoffman, R. R., editor, *Expertise Out of Context: Proceedings of the sixth international conference on naturalistic decision making*, chapter 6, pages 113–155. Lawrence Erlbaum.

Landgren, J. (2005). Supporting fire crew sensemaking enroute to incidents. *International Journal of Emergency Management*, 2(3):176–188.

Landgren, J. and Nulden, U. (2007). A study of emergency response work: Patterns of mobile phone interaction. In *Proceedings of the 2007 SIGCHI conference on Human Factors in Computing Systems CHI 2007.*

Langley, A. (1999). Strategies for theorizing from process data. *Academy of Management Review*, 24(4):691–710.

Latham, G. P. and Locke, E. A. (1991). Self-regulation through goal setting. *Organizational Behavior and Human Decision Processes*, 50(2):212 – 247. ¡ce:title¿Theories of Cognitive Self-Regulation¡/ce:title¿.

Lipshitz, R. (1993). Converging themes in the study of decision making in realistic settings. In Klein et al. (1993), chapter 5.

Lipshitz, R., Klein, G., Orasanu, J., and Salas, E. (2001). Taking stock of naturalistic decision making. *Journal of Behavioral Decision Making*, 14:331–352.

MacCarthy, B. L. and Wilson, J. R., editors (2001). *Human Performance in Planning and Scheduling.* Taylor and Francis.

MacGregor, J. N. and Chu, Y. (2011). Human performance on the traveling salesman and related problems: A review. *The Journal of Problem Solving*, 3(2).

Maitlis, S. and Sonenshein, S. (2010). Sensemaking in crisis and change: Inspiration and insights fro weick (1988). *Journal of Management Studies*, 47(3):551–580.

McKay, K. N., Safayeni, F., and Buzacott, J. (1995). 'common sense' realities of planning and scheduling in printed circuit board production. *International Journal of Production Research*, 33(6):1587–1603.

McKay, K. N. and Wiers, V. C. (2001). Decision support for production scheduling tasks in shops with much uncertainty and little autonomous flexibility. In MacCarthy and Wilson (2001), chapter 8.

Miles, M. B. and Huberman, A. M. (1984). *Qualitative Data Analysis.* Sage.

Miller, C. C., Cardinal, L. B., and Glick, W. H. (1997). Retrospective reports in organizational research: A reexamination of recent evidence. *Academy of Management Journal*, 40(1):189–204.

Miyake, N. (1986). Constructive interaction and the iterative process of understanding. *Cognitive Science*, 10(2):151–177.

Mohr, L. (1982). *Explaining Organizational Behavior.* Jossey-Bass.

Moldoveanu, M. (2009). Thinking strategically about thinking strategically: The computational structure and dynamics of managerial problem selection and formulation. *Strategic Management Journal*, 30:737–763.

Moxnes, E. (1998). Not only the tragedy of the commons: misperceptions of bioeconomics. *Management Science*, 44(9):1234–1248.

Muhren, W. J. and Van de Walle, B. (2009). Sensemaking and information management in humanitarian disaster response: Observations from the triplex exercise. In *Proceedings of the 6th International ISCRAM Conference, Gothenburg, Sweden, May 2009.*

Newell, A. and Simon, H. A. (1972). *Human Problem Solving.* Prentice-Hall.

Pettigrew, A. M. (1990). Longitudinal field research on change: Theory and practice. *Organization Science*, 1(3):267–292.

Rudolph, J. W., Morrison, J. B., and Carroll, J. S. (2009). The dynamics of action-oriented problem solving: Linking interpretation and choice. *Academy of Management Review*, 34(4).

Schweitzer, M. E. and Cachon, G. P. (2000). Decision bias in the newsvendor problem with a known demand distribution: experimental evidence. *Management Science*, 46(3):404–420.

Simon, H. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, 69(1):99–118.

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review*, 63(2):129–138.

Smith, W. and Dowell, J. (2000). A case study of co-ordinative decision-making in disaster management. *Ergonomics*, 43(8):1153–1166.

Spradley, J. P. (1980). *Participant Observation*. Holt, Rinehart and Winston.

Sterman, J. D. (1989). Modeling managerial behavior: misperceptions of feedback in a dynamic decision making experiment. *Management Science*, 35(3):321–339.

Tversky, A. and Kahneman, D. (1974). Judgment under uncertainty: heuristics and biases. *Science*, 185(4157):1124–1131.

Vernon, C. (2001). Lingering amongst the lingerie: An observation-based study into support for scheduling at a garment manufacturer. In MacCarthy and Wilson (2001), chapter 7.

Volkema, R. J. (1983). Problem formulation in planning and design. *Management Science*, 29(6):639–652.

Walsh, B. (2006). Fema says katrina taught it a lesson. Times-Picayune, 13 May 2006.

Webster, S. (2001). A case study of scheduling practice at a machine tool manufacturer. In MacCarthy and Wilson (2001), chapter 4.

Weick, K. E. (1988). Enacted sensemaking in crisis situations. *Journal of Management Studies*, 25(4):305–317.

Weick, K. E. (1993). The collapse of sensemaking in organizations: The mann gulch disaster. *Administrative Science Quarterly*, 38(4):628–652.

Weick, K. E. (1995). *Sensemaking in Organizations*. Sage Publications, Inc.

Weick, K. E., Sutcliffe, K. M., and Obstfeld, D. (2005). Organizing and the process of sensemaking. *Organization Science*, 16(4):409–421.

WFP (2009). Simulating an emergency response in brindisi. Logistics Cluster Blog, 8 December 2009.

Woods, D. D. (1993). Process-tracing methods for the study of cognition outside of the experimental laboratory. In Klein et al. (1993), chapter 13, pages 228–251.

Yin, R. K. (2009). *Case Study Research: Design and Methods.* Sage, fourth edition edition.

Chapter 3

Using expert preferences to assess trade-offs among multiple objectives for humanitarian aid delivery

3.1 Introduction

Recent disasters have highlighted the importance of logistics in emergency response. Supply chain challenges hampered aid efforts in the South Asian tsunami (CNN, 2005), Hurricane Katrina (Walsh, 2006), and the Haiti earthquake (BBC, 2010). Emergency response entails many logistical challenges, including procurement of scarce supplies and services, storage and transport of goods despite destroyed infrastructure, and reaching affected regions through damaged or congested roads and airports. This paper addresses the transportation of humanitarian aid cargo to affected communities, under constrained resources and high demand, and focuses on understanding the trade-offs among the multiple goals of aid delivery.

Numerous scholars and practitioners have developed modeling approaches to support and improve aid delivery (Caunhye et al., 2012; de la Torre et al., 2011). However, a key problem with existing models is the lack of an appropriate objective. The maximization of profit, the clear goal of commercial endeavors, is irrelevant for the non-profit entities that typically provide humanitarian aid. Humanitarian aid delivery models have, therefore, employed other objectives, often maximizing some measure of service. Many models explore multi-objective formulations, making assumptions about penalty functions or objective weights to trade off conflicting goals. However, little empirical validation has been provided for these measures.

Empirically-based objectives are difficult to define and use in operational decision-making, even though humanitarian organizations have a common goal to alleviate human suffering (de la Torre et al., 2011). For example, it is difficult to determine whether loading more food, more shelter, or more medicine on a truck will have a greater impact on human well-being. Should planners focus on delivering fewer supplies faster, reducing costs, or delivering more cargo but with less speed? Balancing these goals and understanding how each contributes to the alleviation of suffering is difficult, as is modeling such objectives quantitatively so that optimization models can be utilized.

To address these challenges, modelers must operationalize humanitarian objectives in a way that captures the key goals and assesses the relative importance of each. Toward this end, we measure the preferences of experienced humanitarian logisticians, to learn how experts trade off the multiple goals of humanitarian aid delivery. We quantify the importance of each objective and formulate a piece-wise linear utility function for humanitarian aid delivery. The survey results and utility functions should enable the development of optimization models that make better trade-offs among the multiple objectives of humanitarian aid delivery.

3.1.1 Context: transportation planning for aid delivery

This chapter develops objective functions to guide the planning of transportation for humanitarian aid delivery, although the issues considered herein are likely relevant for other humanitarian logistics problems as well. In the literature, transportation planning for aid delivery has been studied as a last-mile delivery problem and as a routing problem (see De la Torre et al., 2011, for a review, which we do not repeat here). In practice, this problem is encountered by numerous organizations, but we study it from the perspective of a United Nations-managed organization called the Logistics Cluster. The Cluster often provides transportation services to multiple aid organizations operating in the emergency theater (e.g. UNICEF, International Medical Corps, Save the Children, and many others).

The Cluster typically manages a fleet of vehicles that includes both trucks and helicopters. Organizations submit requests to move cargo within the region: for example, from an international airport to a set of villages in the affected area. Often, the Cluster receives many more requests than they have the capacity to fulfill. Cluster logisticians must plan the routes for the trucks and helicopters and determine which cargo to carry and how to prioritize shipments.

The Cluster's challenge illustrates the complexity and importance of developing objectives for aid delivery. Transportation decisions affect the amount and type of aid distributed, the areas to which aid is provided, and the speed with which it arrives. Each of these outcomes has some impact on human well-being. Objective functions should consider the importance of the various consequences of transportation decisions, such as the types of aid delivered and the spatial and temporal distribution of deliveries, alongside more traditional service- and cost-based criteria. This chapter seeks a way to measure and model the importance of each of the multiple objectives of humanitarian aid delivery.

The Logistics Cluster provides an ideal context in which to study the operational objectives of humanitarian aid delivery. Because the Cluster is an inter-agency group, it is more representative of the humanitarian community than any single organization. The Cluster has provided services many times and in many contexts (in 2011 alone, the Cluster provided logistics services in 10 countries (Logistics Cluster, 2012)). A community of experienced logisticians from multiple agencies are deployed by the Cluster where and when they are needed. Thus, the Cluster has extensive experience and an inter-agency perspective on humanitarian aid delivery, which makes it an excellent context in which to understand the broader objectives of aid transportation planning. The objective functions developed in this research should therefore be useful beyond the Cluster scenario in which they are built.

3.1.2 Approaches to modeling humanitarian objectives

Models of humanitarian aid delivery employ a variety of objectives in attempting to represent the goals of aid transportation (see de la Torre et al., 2011; Holguin-Veras et al., 2010, for reviews). Many of these models make trade-offs between conflicting goals. Trade-offs are manifested on both an operational level, in determining which communities to serve first and what commodities to bring, and at a conceptual level, in considering conflicts among goals like efficiency, effectiveness, and equity.

At a conceptual level, humanitarian aid delivery (as with other public services) exhibits conflicting goals of efficiency, effectiveness, and equity (Huang et al., 2012; Marsh and Schilling, 1994; Savas, 1978). One or more of these goals is considered, explicitly or not, in the objective functions of most aid delivery models. Efficiency objectives typically minimize operations costs (Tzeng et al., 2007; Balcik et al., 2008) or travel times (Tzeng et al., 2007; Lin et al., 2011; Huang et al., 2012; Campbell et al., 2008). Effectiveness objectives maximize some measure of service, often the amount of demand satisfied (Lin et al., 2011) and/or the speed with which demand is satisfied (Huang et al., 2012; Balcik et al., 2008). Equity objectives encourage models to provide service equally to all recipients, by minimizing the latest arrival time (Campbell et al., 2008), minimizing the sum of arrival times (Campbell et al., 2008), maximizing the smallest demand satisfaction rate (Tzeng et al., 2007), minimizing the disparity in demand satisfaction (Lin et al., 2011), or imposing penalties on inequitable deliveries (Huang et al., 2012).

Some of these models explicitly consider trade-offs between efficiency, effectiveness, and equity. By considering multiple objectives, studies have highlighted the cost, in terms of one objective, of optimizing under another objective (Campbell et al., 2008; Huang et al., 2012; Tzeng et al., 2007; Doerner et al., 2007). Some models generate multiple solutions for decision-makers to choose among (Nolz et al., 2010; Tzeng et al., 2007; Doerner et al., 2007; Viswanath and Peeta, 2003); in choosing, decision-makers implicitly weight the importance of each objective. Other models explicitly weight the importance of each objective and optimize accordingly (Lin et al., 2011; Nolz et al., 2010). However, to our knowledge, these weights are based on assumptions about the relative importance of each objective, with little empirical validation.

At the operational level, limited transportation resources necessitate prioritization of deliveries among communities and commodities. In most cases, these trade-offs are modeled using penalty functions, based on the deprivation cost of time without a commodity (Holguin-Veras et al., 2010; Perez et al., 2010), late delivery to a location (Balcik et al., 2008) and of particular commodities (Lin et al., 2011). Penalty functions are a useful way to structure models that prioritize deliveries, but to date, the parameters of the penalty functions have simply been assumed (with the important exception of a deprivation cost for water, in Holguin-Veras et al. 2010).

A key problem, at both the conceptual and operational levels, is the lack of a way to value the trade-offs between conflicting goals and objectives. As described in the preceding paragraphs, trade-offs have been incorporated into models by weighting multiple objectives or incorporating penalty costs into a single objective, but these weights and penalty costs have simply been assumed, rather than measured (again, except for the deprivation cost for water, in Holguin-Veras et al. 2010). Assumed penalty functions and objective weights enable the development of important insights. However, we should also attempt to measure the appropriate penalty functions and deduce the weights that will lead models to make better trade-offs between conflicting goals and objectives.

3.1.3 Approaches to valuing trade-offs

Our challenge is to define a multi-criteria measure of the quality or performance of humanitarian aid delivery plans, which values trade-offs between the criteria. This measure should capture important trade-offs identified in practice and by existing models, including deliveries of various commodities to various locations, and the importance of efficiency, equity, and effectiveness. A natural form for such a measure is a multi-attribute utility function. There are several possible approaches to defining such a function.

One approach is to measure aid delivery effectiveness based on the utility of those receiving the aid – the beneficiaries. For this approach to work, it is necessary to find a utility function that adequately represents the preferences of beneficiaries. Holguin-Veras et al. (2010) and Perez et al. (2010) propose this approach, attempting to minimize the suffering, as measured by the deprivation cost, incurred by people in need. The deprivation costs (disutility functions) have proven difficult to measure, though this seems a promising approach. In the related domain of healthcare, a preference measurement approach is used to weigh the relative benefit of various health interventions, by comparing the number of quality-adjusted life years (QALYs) each intervention provides (see Gold et al., 2002; Morrow and Bryant, 1995; Dranove, 2003). To determine how much to adjust each year, based on the quality of life, individuals' or communities' preferences for different health states are measured and converted to weights, yielding a utility function for each individual or community. Utility functions like these could be maximized in aid delivery models.

The challenge in applying approaches like QALYs or deprivation costs is that measuring the preferences of beneficiaries of humanitarian aid is challenging for numerous reasons. Beneficiaries in different emergencies have different preferences. There are so many possible locations and types of emergencies that it is impractical to measure preferences ahead of time, and in the midst of an emergency, preference measurement is practically impossible. The ability to gather beneficiary preferences would be valuable, but we leave this approach to future research.

A second approach is to model the utility functions of humanitarian experts, whose job it is to make the decisions on how to prioritize aid deliveries. Experienced aid workers should have in mind the preferences of the beneficiaries in their own preferences. Moreover, they may be better positioned than beneficiaries to determine some aspects of aid distribution. For example, beneficiaries may not realize the importance of adequate sanitation to prevent the spread of disease. Certainly such experts may also have other influences on their preferences, such as the need to raise money from donors. Nevertheless, measuring expert preferences seems the most direct route to a practical objective for humanitarian aid.

In health care, a similar approach has been taken in the development of disability-adjusted life years (DALYs). QALYs, which incorporate the health preferences of individuals or communities, were not suitable for allocating resources across diverse sets of communities or populations. DALYs were developed for this purpose, by measuring the preferences of a group of health experts (Gold et al., 2002; Jamison et al., 1993). In this research, we work toward a similar measure for humanitarian aid, one that can be used across communities to determine the best resource allocation decisions.

This research seeks to find a multi-attribute utility function that values trade-offs between key outcomes of humanitarian aid delivery, including conceptual measures of efficiency, effectiveness, and equity, along with operational trade-offs like delivery of one commodity or another. Section 3.2 describes the selection and use of the conjoint analysis method of measuring preferences. In Section 3.3, we apply conjoint analysis to the problem of aid transportation planning, and we develop the performance criteria by which plans will be judged. Section 3.4 details the survey design, and Section 3.5 discusses the results, showing how experts assessed trade-offs among the performance criteria. Section 3.6 maps the survey results to objective functions for humanitarian aid delivery models. Section 3.7 provides conclusions and directions for future work.

3.2 Measuring Preferences with Conjoint Analysis

Our goal is to find a function that can quantitatively value aid delivery plans, by measuring expert preferences across such plans. Utility functions are a natural method of capturing expert preferences, and can be defined over a wide range of plan attributes. Later sections consider which specific plan attributes to include in the analysis. In this section, we consider multiple methods of measuring expert preferences, then describe the conjoint analysis method, selected for this research, in more detail.

3.2.1 Selection of preference measurement method

Measuring preferences for aid delivery will require the estimation of a multi-attribute utility function. There are several methods that could potentially be used to measure such a function, including conjoint analysis, the analytic hierarchy process, and methods associated with multi-attribute utility theory.

Multi-attribute utility theory supports complex decisions by measuring a decision-maker's utility function in an in-depth interview, which looks for indifference points in tradeoffs among the attributes of the potential decision outcomes (Keeney, 1977; Dranove, 2003). The analytic hierarchy process (AHP) breaks decisions down into a series of pairwise comparisons, and calculates preference scores based on the decision-maker's evaluation of each comparison (Forman and Gass, 2001; Meissner and Decker, 2009). Conjoint analysis uses a survey, in which respondents rank or rate several possible decision outcomes, to estimate a utility function over the attributes of the decision outcomes (Green and Srinivasan, 1978; Green et al., 2001; Orme, 2006).

This research required a method that would take limited time from busy humanitarian logisticians, that would ask intuitive questions with no need for pre-training, and that would scale easily to multiple respondents. In these respects, conjoint analysis offers several advantages over the other methods. The number of questions required to estimate a utility function is smaller than in AHP, enabling shorter surveys. The survey questions are intuitive, since they are asked to make choices between different alternatives, as they would do in practice. In contrast, the interviews of multi-attribute utility analysis are much more complex. Finally, conjoint analysis surveys scale naturally from single to multiple respondents, whereas multi-attribute utility analysis would require an interview of each respondent. For these reasons, conjoint analysis was selected for this research.

It is important to note that conjoint analysis assumes linear, additive preferences, meaning that nonlinear preferences cannot be captured. However, the method remains relatively flexible in that it only restricts utility functions to be piece-wise linear, so nonlinear effects can be approximated. Piece-wise linear functions are a first-order approximation to true utility functions, but they capture important elements of human preferences, showing the direction and magnitude of change in preferences as some quantity varies. The ranges of values for which utility was estimated (discussed later) were chosen to lie in the middle of the range rather than at its extreme points, to avoid areas where preferences were likely to be nonlinear (for example, when all cargo is delivered, prioritization is unimportant). In these middle-range values, linear preferences seem likely to provide a good approximation to true preferences: delivering a little more cargo, a little faster, should provide a little more utility. Furthermore, while a more flexible model might be desirable to represent preferences accurately, models with nonlinear objectives are more difficult to solve. Piece-wise linear objective functions represent a good balance between accuracy and implementability.

3.2.2 Introduction to conjoint analysis

Conjoint analysis is typically used to measure consumer preferences for products (Green et al., 2001; Orme, 2006). Consumer preferences for a product are assumed to be a function of the attributes of that product. A product is decomposed into a set of *attributes* (such as weight), each with a set of *levels* (e.g. light and heavy). These attributes and levels can be combined to create a number of potential product profiles (e.g. a light product at a cost of \$1, or a heavy product at a cost of \$2). A consumer is given a survey asking him to rank or select from among several product profiles. Based on his responses, it is possible to estimate his "part-worth utilities", meaning how much each attribute-level contributed to his preference for each product profile.

To estimate all of an individual's part-worths, the respondent would have to evaluate all possible product profiles that can be created from combinations of attribute-levels for the product. Conjoint analysis provides ways to reduce the number of evaluations required to estimate useful part-worths. The product profiles are carefully selected based on the same principles used in design of experiments. Questionnaire designs can be fixed, in which the questions are selected ahead of time (Kuhfeld et al., 1994), or adaptive, in which the questions are selected in real time based on the answers already provided by the respondent (Toubia et al., 2007; Sawtooth, 2009a).

Once the questionnaire is completed, the part-worths must be estimated, using one of several possible estimation methods. Each question in the survey provides some information about which product profiles are preferred over others. Statistical methods combine information from all the questions and estimate how much each attribute-level contributed to the total utility of the product profiles, based on the assumed utility model (described in the next section). Multiple estimation methods can be used, but the most effective method appears to be hierarchical Bayes estimation (Lenk et al., 1996; Sawtooth, 2009b).

There are multiple formats for the survey given to respondents. Respondents can be given partial or full product profiles (only a few attributes or all the attributes). The task they are asked to perform also differs. Sometimes respondents are asked to rank a set of profiles in order of preference. In choice-based surveys, they are asked to choose which product they would purchase among a set of profiles. In metric paired-comparison surveys, they are shown two profiles and asked how much they prefer one to another.

Before describing the survey developed for this research, we first describe the assumptions underlying conjoint analysis methods.

Model and assumptions

There are many models for conjoint analysis that make different assumptions, but here we provide an overview of the model used in this survey (for more details, see, e.g., Green and Srinivasan, 1978; Abernethy et al., 2008; Grissom et al., 2006). Recall that a product is decomposed into a set of attributes, each with a set of possible levels. A product profile, then, is made up of levels for each of its attributes. A row vector $\mathbf{x} = (x_1, x_2, ...x_n)$ defines the levels of attributes in a given profile.

The components of $\mathbf{x}$ are defined differently depending on the assumed model of preference over each attribute. Green and Srinivasan (1978) describe three models of preference. In the ideal vector model, the utility function over each attribute is linear, so that preference for the attribute scales linearly as the level of the attribute increases. The second model assumes an ideal point for each attribute, and preference decreases as we move away from the ideal level. The third model, the part-worth function model, is the most flexible: it assumes no specific form for the utility function over each attribute. The utility of several discrete levels of the attribute is found, and then linear interpolation can be used to create a piece-wise linear utility function for the attribute. In this latter model, used in this research, each of the components of **x** is a dummy variable indicating whether or not a particular level is present in this profile. (For example, $x_1 = 1$ if the profile is heavy and 0 if it is not heavy, while $x_2 = 1$ if the profile is light and 0 if it is not light. Similarly, $x_3 = 1$ if the profile costs \$1 and 0 otherwise, while $x_4 = 1$ if the profile costs \$2 and 0 otherwise.)

Conjoint analysis typically assumes that consumer preferences for products are a linear, additive function of the attributes of that product. The assumption of additive preferences requires the attributes to be mutually preferentially independent. The preference for one attribute, when all others are fixed, must not depend on the levels of the fixed attributes. This assumption is taken into account in the design of conjoint surveys, by ensuring that the attributes are independent. Under the assumption of linear, additive preferences, a product's utility is $U(\mathbf{x}) = \sum w_i x_i = \mathbf{x} \cdot \mathbf{w}$, where $\mathbf{w} = (w_1, w_2, ..., w_n)$ is the part-worth vector. The part-worth vector is what we are trying to estimate: it represents the partial utilities of each of the attribute-levels of the product.

3.3 Designing a Conjoint Survey for Aid Delivery

In this section, we consider the specific criteria to be included in our multi-criteria performance measure, and the emergency context in which it will be derived. In addition, we describe how these elements are integrated into the design of the conjoint survey, which will measure the utility functions of experts.

3.3.1 Context: survey scenario

The conjoint analysis method is here used to measure preferences for aid delivery plans, rather than consumer products. The survey asks respondents to select a preferred plan from among several plan profiles. Which plan is preferred depends very much on the disaster environment: the severity and scale of the emergency, the terrain, the climate, the population, and many other variables. It was necessary to provide some of this information in order to enable respondents to select among transportation plans. We tried to balance this goal with the desire for generalizable results, by providing a scenario specific enough to enable selection among plans, yet general enough to draw broader conclusions.

The scenario focuses on aid deliveries in the first week after a major earthquake. The affected areas are cold and mountainous, and the earthquake has destroyed much of the infrastructure. There is an order of priority for types of aid: shelter is first priority, health second, water and sanitation third, and food fourth priority. There is also an order of priority for locations: some locations (first priority) are in more urgent need than others (second priority). No further information about the scenario is provided. Participants are told to evaluate plans based on the limited information they have.

This scenario is relatively general (for aid delivery problems) because it provides a *given* order of priority for item types and locations, rather than asking respondents to determine the priorities based on the scenario. Therefore, it should generalize to other scenarios in which priority item types and locations are well defined, which is common in emergency response. However, it is specific to the *first phase* of a *major emergency response.* In later phases, and in smaller emergencies, the goals of aid delivery might be different, or resources might be valued differently. The same study design could be used to investigate other response phases, but this investigation is specific to the first phase.

3.3.2 Performance criteria: survey attributes and levels

The selection of criteria to include in a multi-criteria performance metric is crucial to the usefulness of the resulting function. Recall that our goal is to value trade-offs between various objectives of humanitarian aid delivery. This section describes the considerations involved in selecting objectives to include, then describes how they are operationalized within the conjoint survey, in the design of the survey's attributes and levels.

In determining which objectives, or performance measures, to study, we considered performance measures used in practice and in models, along with the requirements of conjoint analysis surveys.

To find measures that match the objectives of expert humanitarian aid workers, we looked for the most important characteristics by which experts evaluated the "goodness" of different aid delivery plans. An extensive field research study helped to identify the informal performance measures employed by expert humanitarian logisticians (see Chapter 2 of this thesis). Through a grounded theory analysis of observations of 10 expert teams creating aid delivery plans, we distilled five key performance metrics by which people evaluated their evolving plans: the total cargo delivered, the total delivered of each type of item (shelter, health, water and sanitation, and food), the total delivered to high- and low-priority locations, the speed of delivery, and the cost of the operation.

Next we consider the usefulness of these five measures of performance: do they capture the important trade-offs identified in practice and in past modeling efforts? Because they are derived from practice, we believe they do capture trade-offs that matter to expert logisticians. Moreover, they include most of the objectives utilized in past humanitarian aid delivery models (see Section 3.1.2). Therefore, valuing trade-offs among these measures should provide useful objective weights for future modeling efforts. Finally, these measures will enable us to explore trade-offs between efficiency, effectiveness, and equity. Instead of formulating one objective specifically designed to measure each of these three goals, the five practice-derived metrics unpack each goal into some of its component measures. Efficiency is measured by cost, effectiveness by the total cargo delivered and the speed of delivery, and equity by the prioritization of deliveries across locations. We believe this is a more useful approach, because we can measure finer-grained trades among the component measures and still gain some intuition about the relative importance of efficiency, effectiveness, and equity. Note, however, that the measure we map to equity does not correspond directly to equity measures used in past models (Marsh and Schilling, 1994; Huang et al., 2012; Lin et al., 2011; Campbell et al., 2008), which typically focus on equalizing distribution of services rather than prioritizing deliveries across locations. We chose to use the measure found in practice rather than introduce model-based constructs into the survey. Furthermore, as we will describe in the following paragraphs, we use the issue of prioritization to probe the importance of equitable distribution.

To satisfy the assumptions of conjoint analysis, the performance measures must be independent, with a limited number (3, in this survey) of mutually exclusive levels that span the range of possibilities to be studied (Sawtooth, 2011). In some cases, measures that were conducive to use as objective functions in optimization models were not appropriate for conjoint analysis nor intuitive to decision-makers. For example, in optimization models, a useful way to measure prioritization by item type is to measure the fraction of satisfied demand for each type of item. However, this measure proved non-intuitive to survey respondents and it was not independent of the total cargo delivered (over all item types) and therefore unsuitable for conjoint analysis. Instead, we measure the fraction of capacity allocated to each type of item, which leads to a nonlinear objective function. We strove to choose measures both intuitive to decision-makers and appropriate for conjoint analysis, but occasionally the analytical tractability of the measures was less than desired (this issues is addressed further in Section 3.6).

In conjoint analysis, respondents evaluate profiles described in terms of several *attributes*, each of which can take on one of several *levels*. In our survey, the profiles are aid delivery plans, and the attributes are the performance measures by which they are evaluated, i.e., the objectives whose trade-offs we wish to investigate. The levels are discrete values of these performance metrics. The following paragraphs describe the transformation of the five metrics identified earlier into measures both intuitive to humanitarian logisticians and appropriate for conjoint analysis. The complete set of attributes and levels is provided in Table 3.1. In the survey, each attribute description included a "[?]" button which, when clicked, accessed a more extensive description of the attribute-level (provided in Appendix A.1).

Total cargo delivered The total cargo delivered was simply described as the percent of requested cargo that was delivered during one week. The next step was to select three discrete levels that span the range of possibilities, i.e. from 0% to 100%. We feared that, in the survey, levels of 0% or 100% would provide skewed data, because 0% might be simply unacceptable to respondents (generating no data), while 100% would render useless the prioritization attributes (prioritization is irrelevant if all the cargo is delivered). Therefore, levels were chosen towards the middle of the range, with a high level of 80%, a middle level of 60%, and a low level of 40%.

Prioritization by item type and by location The two prioritization metrics – by item type and by location – were more difficult to describe. One way to show how a plan prioritized item types was to list the amount of each item type delivered, but this seemed cumbersome. Instead, three "prioritization schemes" were developed. One level would be a lexicographic prioritization scheme, which sends all priority-1 cargo first, then all priority-2 cargo, and so on. The opposite extreme would be an evenly mixed prioritization scheme, which attempted to deliver the same amounts of all cargo types. Between these two extremes, a "weighted mix" scheme would deliver all four types of items, but more of the higher-priority types. In order to ensure that the prioritization attributes were independent

1. Total cargo delivered
    (a) Deliver 80% of cargo requested for this week
    (b) Deliver 60% of cargo requested for this week
    (c) Deliver 40% of cargo requested for this week
2. Prioritization by item type
    (a) Priority-1 (shelter) items first: Load vehicles with Priority-1 items before any other cargo.
    (b) More high-priority items: Load vehicles with a mix of items, but more of the higher-priority types.
    (c) Even mix of items: Load vehicles with an even mix of item types.
3. Prioritization by location
    (a) Priority-1 locations first: Send vehicles to high-priority locations before serving other locations.
    (b) More high-priority locations: Send vehicles to all locations but send more to the high-priority locations.
    (c) Even mix of locations: Send vehicles to all locations in equal proportions.
4. Speed of delivery
    (a) 1-3 days: Complete most deliveries in 1-3 days.
    (b) 2-6 days: Complete most deliveries in 2-6 days.
    (c) 4-7 days: Complete most deliveries in 4-7 days.
5. Cost
    (a) $0.5 million cost for this week
    (b) $2.0 million cost for this week
    (c) $3.5 million cost for this week

Table 3.1: Attributes and levels

of the total deliveries, they were described not in terms of the total cargo they delivered of each item type or to each location, but instead based on how vehicles were loaded with cargo (item prioritization) and dispatched to destinations (location prioritization). This description also proved more intuitive to humanitarian logisticians.

Speed of delivery Speed of delivery was also challenging to describe in a compact manner, because deliveries are made throughout the week. Two clear options were to state an average delivery time or give a range of days in which most deliveries occurred. The latter was selected because it was more intuitive for respondents. The 'fast' level concentrated deliveries in the first 1-3 days, the 'slow' level concentrated deliveries in the last 4-7 days, and the in-between level spread them throughout the week.

Cost Finally, the cost was easily described as the operational cost for one week of operation, in dollars. The selection of the three levels was more challenging, because costs can vary widely for different kinds of responses by different organizations. The range of costs was centered on the estimated cost for a one-week response in the scenario upon which the questions are based. This estimated cost includes a large number of trucks and several helicopters. Then, the high and low levels were designed around this center, roughly corresponding to a response with many more helicopters (high) and one with no helicopters (low). Humanitarian experts confirmed that these cost levels made sense for this scenario.

The five attributes were carefully designed to be as independent as possible, in order to satisfy the assumptions of conjoint analysis. Independence requires that the preference for one measure, when all others are fixed, does not depend on the values of the fixed measures. The five attributes can take any value no matter the values of the other measures, and the value of one attribute does not provide any information about the values of the other levels. For example, prioritization by item type was described in terms of the amount of cargo loaded on each truck, which does not affect the total amount of cargo delivered, how quickly it was delivered, where it was sent, nor how much the truck movement cost. Of course, in any given instance of the problem, some combinations of these attributes woudl be infeasible. For example, in some networks, delivering to high-priority locations might require much longer trips, and as a result less cargo could be delivered. However, in this survey, no such specifics were provided, and respondents were told that all these combinations were feasible.

The choice to provide three levels for each attribute was made because it seemed the minimum number required to understand the general structure of respondent preferences. It is desirable to have a small number of levels because it reduces the number of required questions and enables better estimation of the more limited number of parameters. It is also desirable to have the same number of levels across all performance measures, because attributes may artificially appear more important if they have more levels than other attributes (Orme, 2006; Sawtooth, 2011). At least three levels were required to distinguish whether there were differences in how much people cared about going from a low to a medium, and from a medium to a high, level of an attribute; moreover, three levels were easy for people to consider. These concerns drove the choice to use three levels for each attribute.

3.4 Research Design

This section describes the design of the survey: the sample of respondents, the survey format and question design, and the estimation method.

3.4.1 Sample

Because the survey's purpose is to develop an aid performance metric based on expert preferences among delivery plans, it was necessary to find a sample of respondents who qualify as "experts". Such people should have in mind the preferences of the beneficiaries and understand the implications of different decisions (such as the importance of sanitation in preventing disease). It is difficult to determine whether people meet these criteria, so we sought humanitarian logisticians with significant field experience. It was hoped that people with field experience would have learned the implications of aid delivery decisions, would have worked with beneficiaries, and would be reasonably good at their jobs.

There is a small community of expert humanitarian logisticians, many of whom are posted in remote parts of the world. Instead of attempting a broad, representative sample of this community, whose members would be hard to define, we chose to survey two specific groups of logisticians brought together for training (see Table 3.2).

Group 1 consisted of the facilitators at a logistics response training run by a major

	Respondents	Avg. Yrs. Exp.	Description
Group 1	18	11.7	Logistics response training facilitators
Group 2	12	8.3	Humanitarian logistics masters students

Table 3.2: Groups of respondents

humanitarian aid organization. The 18 facilitators all had extensive experience in humanitarian logistics (average of 11.7 years), and they came from many different humanitarian organizations, mainly large international aid agencies. Group 2 consisted of students in a humanitarian logistics masters program. The 12 respondents had an average 8.3 years of experience in humanitarian logistics, and came from various aid organizations, both large and small.

Group 1, in particular, is representative of the expert community who currently make aid delivery decisions, because the training at which the survey was administered is directed at people who will fill this role in future emergencies. Group 2, on the other hand, spans a more diverse set of agencies who may have different opinions about the best ways to distribute aid.

In both cases, the survey was introduced to the entire group. They were requested to complete the survey online within the next few days. They had the opportunity to ask questions, but only a few clarifying questions were asked. The survey was completed by nearly all facilitators in Group 1, and all students in Group 2.

3.4.2 Selection of conjoint method and survey format

There are many possible questionnaire formats, which offer different advantages and disadvantages. One of the key differences is the way profiles are described: in terms of all the attributes (full profile) or only a subset of the attributes (partial profile). Another difference is whether respondents are asked to rank profiles, rate how much they prefer one to another (metric paired-comparison), or select a favorite from among a set of possible profiles (choice-based). Some questionnaire designs are fixed, meaning the questions are determined ahead of time, while others are adaptive, meaning the questions are determined in real time, based on the respondent's previous answers, to elicit the most possible information. Finally, each questionnaire format has advantages and disadvantages in the part-worths that can be estimated from it.

Some of the most commonly used formats are adaptive conjoint analysis (ACA), choice-based conjoint (CBC), and adaptive choice-based conjoint (ACBC). ACA asks respondents to choose among two partial profiles, rating how much they prefer one to another. Both choice-based methods provide full-profile options and respondents select among them. Choice-based methods are popular because they are intuitive for respondents, asking them to make choices among products as they would in real purchase decisions. However, each question offers less information than in metric paired-comparison questions (as in ACA), in which respondents indicate how much one profile is preferred to the other. As a result, it is more difficult to estimate individual-level preferences from choice-based surveys than from other formats like ACA. However, recent developments in estimation methods make it possible to do so, and innovations in questionnaire design, like adaptive methods, make it possible to elicit more information from each respondent.

In this research, the questionnaire format was selected to meet several criteria. First, it was important that the questions be very clear, because the "products" to be evaluated are quite complex. Respondents probably have not encountered this type of choice before. While they might have created transportation plans, they probably have never evaluated several possible plans in terms of their attributes (whereas in surveys about computer purchases, for example, most respondents have probably selected a computer based on its speed, memory, brand, etc.). Because of the complexity of the choices, full-profile descriptions were desired, in order to ensure that the trade-offs were clear. Choice-based methods also seemed to offer a more intuitive question format.

A second concern was that the questionnaire be engaging and of manageable length, because the target group of respondents were likely very busy, and were not obligated to complete the questionnaire. Adaptive methods allow shorter questionnaires, because they collect more information from each question. The adaptive choice-based conjoint (ACBC) approach, in particular, appears to offer a more engaging format, keeping respondents interested despite slightly longer questionnaires (Sawtooth, 2009a). These initial criteria all pointed to ACBC as the best approach. However, choice-based methods generally are weaker in estimating individual-level preferences. Still, hierarchical Bayes estimation methods have enabled the estimation of individual-level preferences in ACBC surveys, even with small sample sizes (Sawtooth, 2009c). For these reasons, ACBC was selected for this research. We

Survey section	No. questions
Request consent	-
Introduction and example question	1
Build-your-own ideal plan	1
Introduction to screening questions	-
Screening questions (6) and optional unacceptable rule (1)	6-7
Introduction to choice tasks	-
Choice tasks (max 9) and holdout questions (2)	(max) 11
Demographic questions (5) and feedback	6

Table 3.3: Survey overview

use the Adaptive Choice-Based Conjoint (ACBC) format provided by Sawtooth Software, one of the most commonly used software packages for administering conjoint surveys.

3.4.3 Survey structure and question design

Table 3.3 provides an overview of the structure of this survey, along with the number of questions in each section. Each of the survey sections is described in more detail in the paragraphs below.

Adaptive CBC questionnaires, as implemented in the Sawtooth software package, have three main sections. First, there is a "build-your-own" section, with only one question that asks the respondent to select his most preferred level for each attribute. It also serves as a useful introduction to all the attributes and levels the respondent will see later in the survey.

Second, there is a "screening" section, in which the respondent evaluates individual profiles as either acceptable or unacceptable. In this survey, 6 screening tasks are included, with 3 profiles per task. This is less than the recommended number (Sawtooth, 2011), but the change enabled a shorter survey with (we expected) little loss because the screening questions seemed less likely to provide useful data. Screening questions make sense in purchase decisions, in which one might choose not to buy a product. However, in the case of aid transportation planning, any plan is better than none, so few plans should be deemed unacceptable. After a few screening questions, one "unacceptable rule" question was included. The software scans recent answers for any attribute levels the respondent may consider unacceptable, and asks him to confirm that any one such level is indeed unacceptable.

The third section consists of choice tasks, in which the respondent is shown three profiles and asked to choose among them. The respondent answers a series of such questions, adapted by the software to extract the maximum amount of information. The number of choice tasks can vary depending on the respondent's answers. In this survey, a maximum of 18 profiles can be evaluated, resulting in a maximum of 9 required choice tasks (as recommended by Sawtooth, 2011).

In addition to these three main sections, the survey also contains a request for consent to participate in research, introductory pages for each section, a few demographic questions, and an opportunity to provide feedback. In total, the respondent answers a maximum of 20 choice or acceptability questions and 6 short demographic/feedback questions.

Question design for choice and screening tasks

Question design involves selecting the sets of profiles the respondent sees in each of the screening and choice tasks. The adaptive software designs questions as respondents proceed through the questionnaire. In Sawtooth's ACBC method (Sawtooth, 2009a), questions are designed by varying attributes around the respondent's ideal profile, which they selected in the first section of the survey. The algorithm randomly selects the number of attributes to vary (within a range set by the user, in this case between 2 and 3) and which attributes to substitute.

There are two main goals in question design: to ensure the respondent sees each level enough times to gather preference data on it, and to maximize statistical efficiency (i.e. gather the most useful data for understanding all of the respondent's preferences). The question selection algorithm works to ensure greater level balance (the former goal) by increasing the probability that a level will be selected when it has been included relatively fewer times already. The latter goal of statistical efficiency is generally achieved (in other methods) by choosing orthogonal designs in which each attribute is varied independently. However, in adaptive CBC, attributes are instead varied around the respondent's ideal design. While this results in a less statistically efficient design, it enables collection of better data because respondents are more engaged in evaluating product concepts close to those they prefer (Sawtooth, 2009a).

The number of questions (discussed earlier and shown in Table 3.3) and the question design parameters control the data gained from the survey. This survey's design attempted to limit the number of questions while still gathering enough data to estimate preferences at the individual level, with a relatively small sample size. In order to ensure that these parameters generated good surveys, several tests were performed. The authors filled out a number of surveys, each with different types of preferences, then analyzed the data to ensure that it revealed the intended preferences. In the process of analysis, we verified that the estimation procedure converged. In a second test, the survey designs were examined to ensure that each attribute level appeared at least 3 times, as recommended by Sawtooth (2011).

Holdout questions

A series of holdout questions are included within the conjoint survey. Holdout questions are identical to choice task questions, but they are not utilized in preference estimation. Instead, they are used for model validation: we can measure how well the estimated model predicts respondent preferences, as evidenced by their choices in the holdout tasks. The "hit rate" is the number of correctly predicted choices divided by the total sample size. In addition to holdout tasks, it is common to repeat one of the holdout choices in order to measure the test-retest reliability of each respondent.

Two holdout tasks were included in the survey. Each holdout task was designed so that the options have unequal utility; otherwise, a random predictor would have the same hit rates as the estimated utility function (Johnson and Orme, 2010). It is difficult to know when designing the survey which options will be preferred by respondents, so instead, each option was designed to be preferred by respondents with particular (and different) kinds of preferences.

The first holdout task was repeated to provide a measure of test-retest reliability. In order to keep the survey as short as possible, this question was first introduced as an example question during the introduction to the survey, then repeated later during the choice tasks section. The example question was a two-choice rather than a three-choice question (for simplicity), so the choice task included the same two choices, along with a dominated choice (for respondents with rational preferences). This test-retest design is not ideal, because it contains only two choices and because it appears so early in the survey, but it does enable a rough estimate of test-retest reliability without adding additional questions to the survey.

Introductory materials and question text

The question text and introductory materials were designed carefully to ensure that respondents understood the tasks they were being asked to complete. The text was intended to be as brief and as clear as possible, so that respondents would understand the questions and instructions, yet would not be burdened with unnecessary information. The text used to describe the attributes and levels was described in Section 3.3.2. The remaining survey text can be found in Appendix A.2.

Several versions of the survey were pilot-tested with small groups of respondents. Respondents provided detailed feedback to the researcher on points of confusion, length of the survey, and any other reactions. They were also questioned by the researcher to ensure their understanding of the questions matched the intended meaning. The survey was revised after each set of feedback and tested again, until respondents reported no confusion and a manageable survey length. Pilot test respondents included both university colleagues and humanitarian logisticians.

3.4.4 Estimation method

After the survey has been completed by respondents, their part-worth utilities can be estimated. This study utilizes the hierarchical Bayes (HB) estimation method, which appears to be the most effective estimation method available (Sawtooth, 2009b). The method estimates individual part-worth utilities by borrowing information from all respondents, thus enabling more accuracy with fewer questions (Lenk et al., 1996). In this section, we describe the estimation method as implemented by Sawtooth Software. Our intention here is to provide enough details to understand the intuition of the algorithm, but additional information can be found in Sawtooth (2009b) and Allenby and Rossi (2006).

The goal is to estimate the part-worths of all individuals, given the data from each individual's survey responses. The algorithm works on two levels, incorporating both the distribution of part-worths over all the respondents and the choices made by each individual during the survey. In the higher level, we assume the individual part-worths are distributed normally. In the lower level, we assume a logit model describes the choices made by individuals. The parameters of both models (all individuals' part-worths along with their distribution) are estimated iteratively by drawing estimates of each parameter in turn, and converging toward the most probable set of parameters.

This study uses the algorithm implemented by Sawtooth Software (Sawtooth, 2009b). Analysis of the estimation process showed that the algorithm converged. The algorithm produced estimates of both average and individual part-worth utilities.

3.5 Survey Results and Discussion

The results of conjoint analysis surveys are part-worth utilities: the amount of utility contributed by each attribute-level to the overall utility of a profile (aid delivery plan). Put another way, the overall utility of a plan is the sum of the part-worth utilities of each of its attribute-levels. Figures 3-1 and 3-2 show the average part-worths for Group 1 and Group 2, respectively, with individual utilities for each respondent in gray behind the averages. Part-worth utilities are estimated only for the three discrete levels of each attribute, but these graphs include lines connecting the levels, to aid in reading the graph. Utilities were estimated for each individual separately, along with a group average. The average is not meant to represent a group utility function; rather, it is simply a useful way to summarize the individual preferences of these groups of experts.

Utilities can be compared within an attribute (delivering 80% is better than delivering 60%, which is much better than delivering 40%), but they cannot be compared across attributes (delivering 80% may or may not be preferred to plans costing \$0.5 million). Utilities are interval scaled data, with no natural zero point, and here they are arbitrarily scaled so that they sum to zero for each attribute.

It is possible to compare utilities across attributes by considering how much the attribute could contribute to the utility of a plan. The size of the difference in utility between two levels of an attribute indicates how much utility can be gained by "moving up" one level. Intuitively, those attributes that cover a larger range of utility are more important. For example, it is clear in Figure 3-1 that total cargo delivery is much more important than cost, because total delivery has a range of about 175 "utils", while cost has a range of less than 50 utils. While this method of comparing across attributes is intuitive, note that it depends on the specific levels selected for each attribute. An attribute will appear to have more importance if its levels cover a larger range: for example, 40% to 80% should have a larger jump in utility than 50% to 70%. As a result, comparisons across attributes are

Attribute	Preferred Level	Att. Range	Avg. Importance	
			Group 1	Group 2
Total cargo delivered	80% (highest)	40% - 80%	0.33	0.24
Prioritization by item type	Weighted mix	P1 first - Even mix	0.21	0.18
Prioritization by location	Weighted mix	P1 first - Even mix	0.17	0.21
Speed of delivery	1-3 days (fastest)	1-3 days - 4-7 days	0.18	0.15
Cost	$0.5 million	$0.5 m. - $3.5 m.	0.10	0.21

Table 3.4: Average attribute importances

only valid for the ranges of levels chosen in each study. (When designing attributes for a study, an effort is made to cover the entire range of possible values, in order to make these comparisons more intuitive.) Table 3.4 lists the average importances of each attribute among each group of respondents (the “importance” is the range of an attribute divided by the sum of ranges over all attributes).

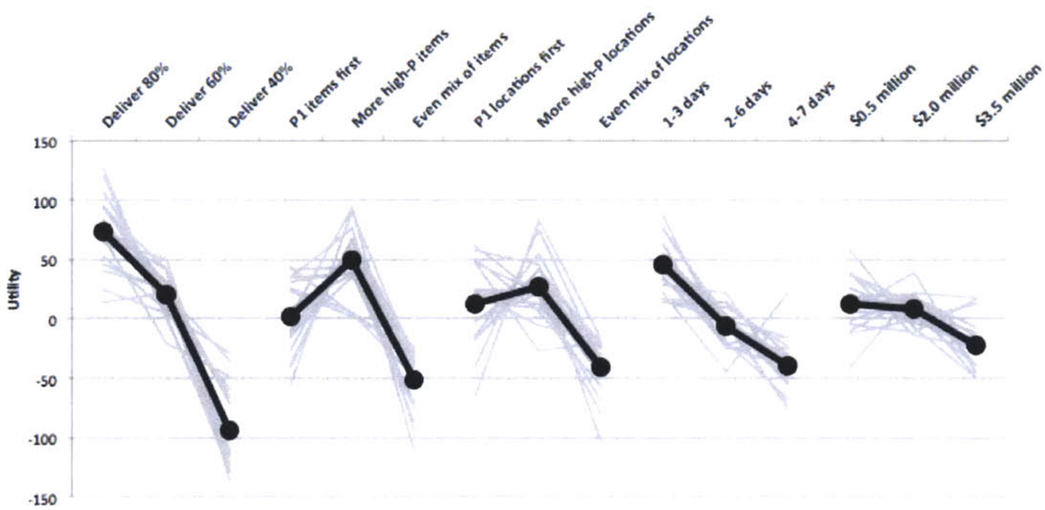


Figure 3-1: Group 1 average (black) and individual (gray) part-worth utilities

The results from Group 1 are shown in Figure 3-1. We focus first on utilities for different levels within the same attribute. As expected, Group 1 respondents, on average, prefer to deliver larger total quantities, with 80% and 60% both much preferred to 40%. They prefer faster speeds of delivery, and lower costs, with $3.5 million slightly less preferred and the other two levels roughly equivalent. Less intuitively, in both prioritization attributes, the “weighted mix” prioritization schemes are preferred to either extreme (priority-1 first, or even mix). One might not expect a decision-maker to send low-priority items before all high-priority items are delivered. However, respondents explained that they preferred to send a weighted mix of cargo because people need more than one type of aid to survive: someone who receives shelter but no sanitation may be safe from exposure but not from disease. In prioritizing across communities, people wanted to send more aid to those in more urgent need, but reasoned that a sending a little aid to the lower-priority communities might be much better than nothing.

Comparing across attributes, the most important concern was increasing the total quantity delivered, especially jumping from 40% to 60%. The least important concern was the cost, with \$3.5 million only slightly lower in utility than the other cost levels. This is consistent with comments from respondents that, in the wake of a major emergency, cost was unimportant because donors would provide whatever funding was needed. There is little difference in importance between the two prioritization attributes and the speed of delivery, though prioritization by item type does appear slightly more important than prioritization by location.

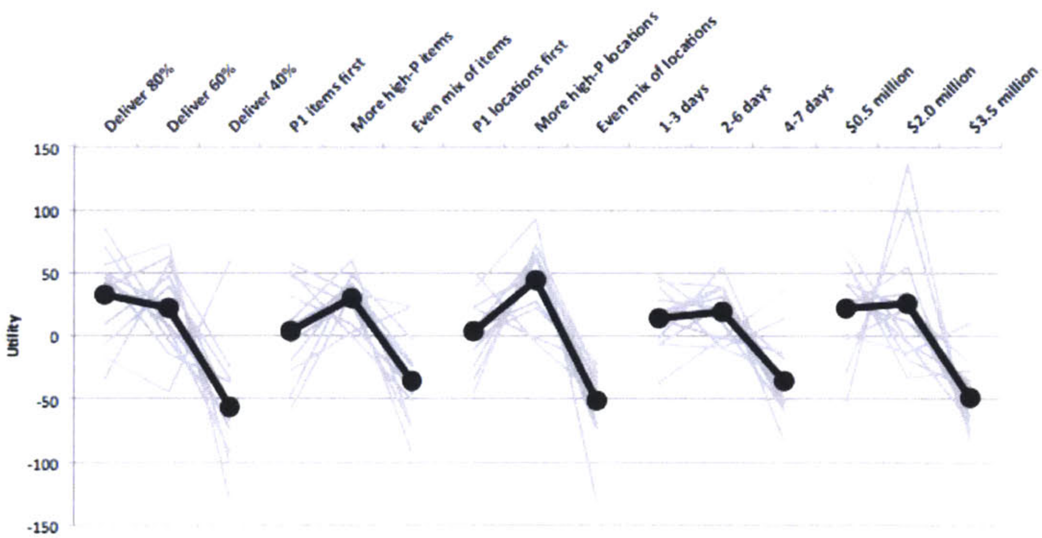


Figure 3-2: Group 2 average (black) and individual (gray) part-worth utilities

The results from Group 2 are shown in Figure 3-2. Within attributes, the order of the levels is the same as for Group 1. Group 2 prefers to deliver more total cargo, and also prefers the “weighted mix” prioritization schemes. The two faster speeds of delivery have roughly the same utility, with the slowest option less preferred. Similarly, the two lower cost levels have roughly the same utility, with the \$3.5 million level much less preferred.

Across attributes, there is less difference in importance than in Group 1. Speed of delivery appears least important, along with prioritization by item type, because they span the smallest ranges of utilities. The more important attributes are total deliveries (especially the jump from 40% up to 60%), prioritization by location, and cost (especially avoiding costs of $3.5 million).

Group 2 provides an interesting contrast to Group 1. Group 2 included more participants from smaller aid organizations, which may account for some of the most striking differences: the higher importance of cost for Group 2, and the importance of prioritizing by location. Discussions with members of Group 2 suggested that the highest cost, $3.5 million, was an extraordinarily high logistics cost for their smaller organizations. On the other hand, many Group 1 respondents came from larger organizations and remembered spending $3.5 million per week in recent emergency responses. It makes sense that cost would be more important to smaller aid organizations who cannot count on the same level of donor support, and who are unaccustomed to the large volumes of cargo commonly handled by larger organizations. Group 2 also appeared to place more importance on prioritization by location. Respondents commented that larger organizations tended to focus on moving large amounts of cargo, while smaller organizations paid more attention to which cargo and where it was needed (though moving less cargo overall).

We have considered only the averages so far, but there is also important information in the variation around the average utilities. The variation around the average utility values can be seen in both the individual utility estimates (shown in gray in each figure) and in the standard deviations around each utility estimate. Table 3.8 provides the average utility values and standard deviations for both groups.

The individual utilities suggest that some patterns of preference are less consistent across individuals. In particular, there appears to be more variation in the prioritization and cost attributes, while the total delivery and speed attributes show more consistent patterns, with a few individual exceptions. This variation could be interpreted in at least two ways: either there is more real variation in preferences across individuals, or these attributes are more confusing and respondents interpreted them differently, or even incorrectly. Some individuals exhibited what we might consider irrational preferences. However, this variability may also be attributable to real preferences. Some respondents mentioned they would prefer a more expensive plan because they would have the option to scale it down or redirect funds, whereas a cheaper plan left little flexibility. Respondents also mentioned they might prefer slower deliveries so that those handing out aid at the final destination would have the capacity to distribute it. It is worth considering both the averages and some of the important exceptions in drawing conclusions about aid delivery preferences.

These results also shed some light on the conceptual trade-off between efficiency, effectiveness, and equity. The five measures in this paper represent components of each of these three goals (see Section 3.3.2). Across both groups, the most important objective was to deliver aid cargo, suggesting that effectiveness is the most important of the three goals. Speed of delivery, often used in optimization models as a measure of effectiveness, was only somewhat important, suggesting that the amount of cargo delivered is a better proxy measure for effectiveness (though both measures matter). For Group 1, cost was the least important, suggesting that efficiency is a secondary goal. Prioritization of deliveries by location, one measure of equity, was less important than total cargo delivery but more important than cost. Group 2 provides an interesting contrast, in that effectiveness, cost, and equity appear to be valued similarly. For the group of Logistics Cluster decision-makers (Group 1), it is clear that effectiveness is the primary concern, followed by equity, and last by efficiency.

3.5.1 Assessing validity with holdout questions

One measure of the validity of the estimated preference models is their ability to predict choices made by respondents. Holdout questions were included in the survey, as described in Section 3.4.3, so it is possible to measure how well the estimated utility functions predict respondent choices on the holdout questions. Recall that holdout questions are identical to the other choice tasks but are not utilized in preference estimation.

The estimated preference model predicts that the respondent will choose the option with the highest utility. The "hit rate" is the number of correctly predicted holdout choices divided by the total sample size. The hit rate obtained from a conjoint survey can be assessed in comparison with the hit rate that would be obtained with random selection of choices. For example, for three-choice tasks, random selection would provide a 33% chance of correctly predicting respondent choices. Hopefully, the hit rate from an estimated model would be higher than 33%.

Table 3.5 shows the test-retest reliability, hit rates, and minimum expected hit rates for

	Group 1	Group 2
Test-Retest Reliability	0.89	0.50
H1 Hit Rate	0.89	0.67
H1 Min. (Random) Hit Rate	0.50	0.50
H2 Hit Rate	0.53	0.42
H2 Min. (Random) Hit Rate	0.33	0.33

Table 3.5: Hit rates and test-retest reliability

Groups 1 and 2. Before considering the results, it is important to note a few complications due to the design of the holdout questions. The test-retest reliability was measured on a two-choice question, based on the agreement in response between the two-choice example and the first holdout question. The first holdout question, H1, was essentially a two-option question, because the third option was designed to be dominated for most respondents. As a result, the two holdout questions have different minimum expected hit rates, because random choice yields different rates depending on the number of questions.

Based on the test-retest reliability, it appears that Group 1 was much more reliable than Group 2. The hit rates for Group 2 are correspondingly lower than those for Group 1. In Group 1, the estimated model predicted correct choices on H1 89% of the time (where a random model would have managed 50%), and on H2 53% of the time (where a random model would have managed 33%). Clearly, the estimated model is much better than random choice, and therefore has some value in characterizing respondent preferences. In Group 2, the estimated model predicted choices on H1 67% of the time, and on H2 42% of the time, again performing better than random choice but not as well as the models for Group 1.

While the hit rates show that the models capture some elements of respondent preferences, they are not perfect predictors of choices. There may be several reasons for this. First, the test-retest reliability shows that some respondents had inconsistent preferences, making prediction difficult. Second, the holdout questions appeared relatively early in the survey, while respondents were getting used to the questions and clarifying their own preferences. Third, the limited number of questions in the survey created a smaller dataset that limited the accuracy of models.

3.6 Mapping Survey Results to Objective Functions

Based on the results of the conjoint survey, it is possible to develop objective functions that value the importance of each of the five attributes in the conjoint survey. In this section, we develop two different forms of objective functions. First, we use the part-worth utilities to develop a piecewise linear utility function over each of the five attributes, which estimates the value of any given aid delivery plan. This requires mapping the survey attribute levels to a performance metric that can be calculated based on any aid delivery plan. Second, we develop a different form of objective function, using the attribute importances as weights in a weighted-sum objective function. These two objective functions represent different ways of interpreting the findings of the conjoint survey.

3.6.1 A utility function over five attributes

In this section, we develop a piecewise linear utility function that can be used as an objective function in optimizing aid delivery plans. We seek a function of the form:

$$U(\mathbf{x}) = u_1(x_1) + u_2(x_2) + u_3(x_3) + u_4(x_4) + u_5(x_5) \tag{3.1}$$

Each u_i is a utility function over attribute i, and each x_i is the value of a performance metric that represents attribute i. Each of the attributes in the survey is equated to a performance metric x_i that can be calculated based on a given aid delivery plan. The survey data enable estimation of $u_i(x_i)$. Each attribute-level maps to a discrete value of the performance metric x_i. The survey results provide a corresponding utility value $u_i(x_i)$ for each attribute level. Interpolating between and extrapolating beyond these points defines a piecewise linear utility function $u_i(x_i)$.

Table 3.6 shows the complete set of attributes and their corresponding performance metrics, which were developed to match the attribute descriptions in the survey. (The remainder of this section describes the mapping of attributes and levels to performance metrics.) The right-most column in Table 3.8 shows the value of the performance metric corresponding to each attribute-level. (The survey respondents had access to expanded definitions of each attribute, beyond those listed in Table 3.1. These expanded definitions, provided in Appendix A.1, provide the details necessary to associate each attribute-level with a value of the performance metric.)

Attribute	Performance Metric x_i
Total cargo delivered	$\frac{\text{total delivered}}{\text{total requested}}$
Prioritization by item type	wtd. "shortfall Manhattan distance" from $\mathbf{r} = [\frac{\text{P1 d.}}{\text{total d.}}, \frac{\text{P2 d.}}{\text{total d.}}, \frac{\text{P3 d.}}{\text{total d.}}, \frac{\text{P4 d.}}{\text{total d.}}]$ to the ideal mix $\hat{\mathbf{r}} = [0.5, 0.3, 0.15, 0.05]$.
Prioritization by location	wtd. "shortfall Manhattan distance" from $\mathbf{s} = [\frac{\text{P1 d.}}{\text{total d.}}, \frac{\text{P2 d.}}{\text{total d.}}]$ to the ideal mix $\hat{\mathbf{s}} = [0.7, 0.3]$.
Speed of delivery	average day of delivery
Cost	operations cost

Table 3.6: Attributes with associated performance metrics

Figure 3-3 shows the estimated utility functions $u_i(x_i)$ for each attribute, based on the Group 1 results. (We work only with the Group 1 average results, for simplicity.) In most cases, the slopes of the lines have an intuitive interpretation, representing the value gained per additional fraction of cargo delivered, or lost per additional day to deliver or increase in cost. Table 3.7 provides the slopes and intercepts for each utility function.

Total cargo delivered The first attribute, total cargo delivered, is measured by the total cargo delivered divided by the total cargo requested for delivery. This is a straightforward interpretation of the attribute levels, which state that some percent of requested cargo was delivered. These attribute levels, in turn, are easily equated with a discrete value of the performance metric, equal to the percentage delivered (see the first rows of Table 3.8). The survey results provide a utility value for each of these attribute levels, thus giving the utility of three discrete values of the performance metric. These are plotted in Figure 3-3a, along with lines interpolated between the points and extrapolated beyond them over the domain of the performance metric.

The lines in Figure 3-3a represent a piecewise linear utility function estimated based on each pair of points. The utility u_1 is a function of the fraction x_1 of requested deliveries completed: $u_1(x_1) = m_1 x_1 + b_1$. The slope m_1 and intercept b_1 can be found based on each pair of points; in this function, the parameters will be different for $x_1 < 0.6$ and $x_1 > 0.6$.

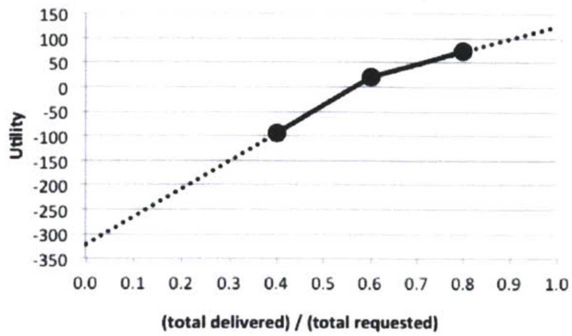
Attribute and Model	Domain	Intercept $b_i =$	Slope $m_i =$	Interpretation of slope
Total delivered $u_1 = m_1 x_1 + b_1$	$x_1 < 0.6$ $0.6 < x_1$	-322 -138	571 264	Value per additional 1% of cargo delivered (of all cargo requested)
Prioritization by item type $u_2 = m_2 x_2 + b_2$	$c_2 < 0$ $0 < c_2$	49 49	-1.395 -0.795	Value per 1-unit increase in shortfall Manhattan distance
Prioritization by location $u_3 = m_3 x_3 + b_3$	$c_3 < 0$ $0 < c_3$	27.6 27.6	-0.94 -0.33	Value per additional 1% of capacity allocated to high-priority locations
Speed of delivery $u_4 = m_4 x_4 + b_4$	$x_4 < 3.9$ $3.9 < x_4$	161 136	-43 -37	Loss in value per additional (average) day to deliver
Cost $u_5 = m_5 x_5 + b_5$	$x_5 <$ \$2 mil \$2 mil $< x_5$	14 49	-3 -20	Loss in value per additional \$1 million in cost

Table 3.7: Slopes and intercepts for estimated utility functions.

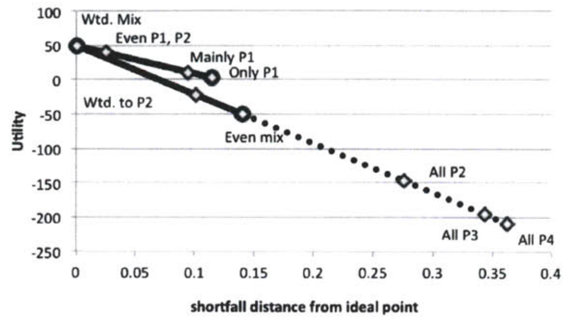
Table 3.7 gives the slopes and intercepts for this (and other) utility functions. The slope has an intuitive interpretation: it represents the value gained per additional 1% of cargo delivered (of all cargo requested).

Prioritization by item type The second attribute describes the allocation of deliveries among different types of items, each of which has a different priority. The attribute levels describe prioritization in terms of the percent of vehicle space allocated to each type of item (specific percentage values were given in the expanded definitions of the attributes, shown in Appendix A.1). This is measured by dividing the amount delivered of each type of item by the total delivered over all priority levels, resulting in a four-dimensional vector, $\mathbf{r} = [\frac{\text{P1 delivered}}{\text{total delivered}}, \frac{\text{P2 d.}}{\text{total d.}}, \frac{\text{P3 d.}}{\text{total d.}}, \frac{\text{P4 d.}}{\text{total d.}}]$.

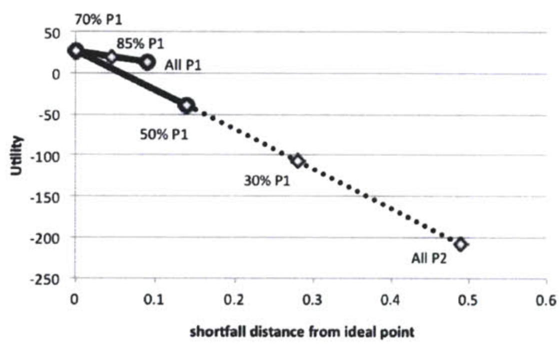
Because the survey gave only three utility values for three instances of this vector, it is not possible to estimate linear coefficients describing the utility of delivering each of the four types of cargo. Instead, we map the four-dimensional vector onto a one-dimensional scale. We take the Manhattan distance between a given point $\mathbf{r}$ and a reference point $\hat{\mathbf{r}}$, which we choose as the most preferred prioritization level, $\hat{\mathbf{r}} = [0.5, 0.3, 0.15, 0.05]$. However, we do not want to penalize solutions for delivering more than the ideal amount of cargo,



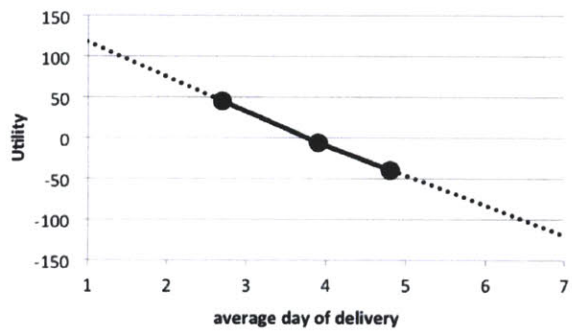
(a) Total delivered



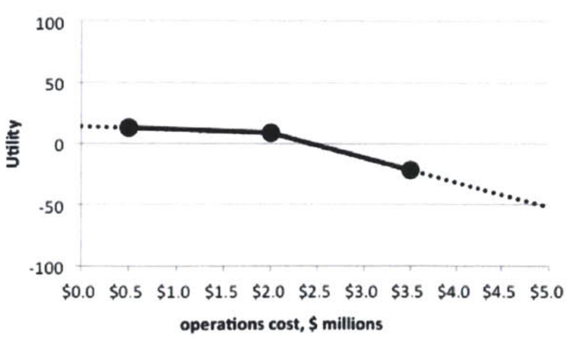
(b) Prioritization by item type



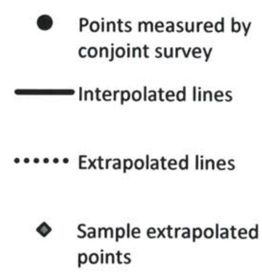
(c) Prioritization by location



(d) Speed of delivery



(e) Cost



(f) Legend

Figure 3-3: Utility functions for each attribute. Black points are measured by the conjoint survey, black lines are interpolated, and dotted lines extrapolated from the measured points.

	Group 1		Group 2		Value of x_i
	Avg. Util.	SD	Avg. Util.	SD	
Deliver 80%	73.3	30.8	32.5	36.5	0.8
Deliver 60%	20.4	17.5	22.8	34.8	0.6
Deliver 40%	-93.7	33.9	-55.3	47.4	0.4
Priority-1 items first	2.2	30.4	4.4	38.1	[1.00 0.00 0.00 0.00]
More high-priority items	49.2	26.8	30.2	18.1	[0.50 0.30 0.15 0.05]
Even mix of items	-51.4	25.6	-34.6	36.9	[0.25 0.25 0.25 0.25]
Priority-1 locations first	12.4	34.9	4.8	28.9	[0.9, 0.1]
More high-priority locations	27.6	26.8	45.3	29.3	[0.7, 0.3]
Even mix of locations	-40.0	24.1	-50.1	32.0	[0.5, 0.5]
1-3 days	45.2	21.6	14.4	24.3	2.7
2-6 days	-6.1	15.7	20.3	21.5	3.9
4-7 days	-39.1	23.8	-34.7	34.3	4.8
$0.5 million	12.9	22.7	22.3	37.2	0.5
$2.0 million	8.8	13.5	26.1	50.2	2.0
$3.5 million	-21.7	21.1	-48.4	30.0	3.5

Table 3.8: Average part-worth utilities with standard deviations, with corresponding values of attribute performance metric x_i.

so only the negative portion of the Manhattan distance is counted, resulting in a "shortfall Manhattan distance". Moreover, it is more important to deliver high-priority cargo than low-priority cargo, so the "shortfall Manhattan distance" is weighted accordingly, again using the vector $\hat{\mathbf{r}}$. Let r_p be component p of $\mathbf{r}$. Then, the item type prioritization measure is $x_2 = \sum_p \hat{r}_p \max[0, \hat{r}_p - r_p]$.

The scale x_2 captures the distance from the ideal mix of item type deliveries, but it does not capture the "direction" of the difference. People preferred lexicographic prioritization to evenly distributed deliveries, so it is necessary to determine whether $\mathbf{r}$ delivers more high-priority item types (closer to lexicographic) or fewer than the ideal point. To do so, we use a weighted difference $c_2 = \hat{\mathbf{r}} \cdot (\mathbf{r} - \hat{\mathbf{r}})$: when $c_2 > 0$, $\mathbf{r}$ is closer to lexicographic; otherwise, it is closer to even or it delivers more low-priority item types.

The utility function $u_2(x_2)$ is plotted in Figure 3-3b; the sign of c_2 determines whether a given point lies on the upper (toward lexicographic) or lower (toward even) line. The three points measured in the survey are shown as black circles. A piecewise linear utility function was extrapolated from each pair of measured points. Several sample points along the extrapolated lines are shown as gray diamonds, to provide a sense of how this scale maps to various possible values of the prioritization vector $\mathbf{r}$.

Prioritization by location The third attribute describes the allocation of deliveries among different locations, just as the second attribute described the allocation of deliveries among types of items. The attribute levels describe prioritization in terms of the percent of vehicles allocated to high- and low-priority locations (specific percentage values were given in the expanded definitions of the attributes, shown in Appendix A.1). Again, this is measured by dividing the amount delivered to each location by the total amount delivered over all priority levels, resulting in a two-dimensional vector $\mathbf{s} = [\frac{\text{delivered to P1 locations}}{\text{total delivered}}, \frac{\text{delivered to P2 locations.}}{\text{total delivered}}]$. The metric is analogous to that described above for item prioritization, and it measures the weighted "shortfall Manhattan distance" from a given point $\mathbf{s}$ to the ideal point $\hat{\mathbf{s}}$: $x_3 = \sum_p \hat{s}_p \max[0, \hat{s}_p - s_p]$. Again, the weighted difference $c_3 = \hat{\mathbf{s}} \cdot (\mathbf{s} - \hat{\mathbf{s}})$ distinguishes between points closer to lexicographic ($c_3 > 0$) or closer to even ($c_3 < 0$).

The utility function $u_3(x_3)$ is plotted in Figure 3-3c. The sign of c_3 determines whether a point lies on the upper (toward lexicographic) or lower (toward even) line.

Speed of delivery The fourth attribute, the speed of delivery, is measured by the average day of delivery for all *delivered* cargo. The average day of delivery can be calculated by summing the amount of cargo delivered on each day multiplied by the day, then dividing the quantity by the total cargo delivered. The expanded definitions of each attribute (see Appendix A.1) state explicit percentages of cargo delivered in each date range, so the average day of delivery can be calculated for each attribute level (see Table 3.8). The survey provides part-worth utilities for each of these attribute levels. The resulting utility function for speed of delivery is plotted in Figure 3-3d, and the slopes and intercepts are given in Table 3.7. The slope represents the loss in value for delivering one day later (on average).

Cost The fifth attribute, the cost, is simply the operational cost of a plan. The attribute levels map directly to this metric. The utility function for cost is plotted in Figure 3-3e, and the slopes and intercepts are given in Table 3.7. The slope represents the change in value for each additional \$1 million in cost.

The performance metrics in Table 3.6 were developed to provide the "best fit" with the conjoint survey questions. As mentioned earlier, we privileged the design of a useful survey over the design of attributes that mapped to convenient objective functions. This led to several performance metrics that are nonlinear: the "total delivered" is a decision variable, and it appears in the denominator of several metrics (both priority metrics and the speed of delivery).

It is often more convenient to have linear objectives. The maximum possible amount of cargo that can be delivered can be substituted for the actual cargo delivered. This substitution does not yield exactly the same measure, but is an equally reasonable interpretation of the survey questions. This quantity can be obtained by first solving a maximum flow problem, then using the result as the denominator in those utility functions that require the total cargo delivered.

3.6.2 A weighted-sum objective based on attribute importances

This section develops a simpler interpretation of the conjoint survey findings. Recall from Section 3.5 that the importance of attributes can be compared by considering how much the attribute could contribute to the utility of a plan, i.e. the difference in utility between the highest and lowest levels of the attribute. The "importance" of an attribute is this difference divided by the sum of the differences between high and low utilities of all attributes. Note, however, that the importance of an attribute depends on the levels selected in the conjoint survey; if its levels cover a larger range, it will have a larger difference in utilities and appear more important. Nevertheless, importances are a useful way to compare how each attribute contributes to plan utility. Table 3.4 lists the average importances of each attribute among each group of respondents.

To develop a weighted-sum objective function, we simply interpret the attribute importances as weights in a weighted-sum objective function. We seek a function of the form:

$$U(\mathbf{y}) = w_1 y_1 + w_2 y_2 + w_3 y_3 + w_4 y_4 + w_5 y_5 \tag{3.2}$$

Each w_i is the importance of attribute i, listed in Table 3.4. Each y_i, like x_i in the previous section, is a metric that can be calculated based on a given aid delivery plan. Here, however, each y_i is a linear function such that maximizing y_i increases the "goodness" of attribute i

(whereas each x_i represented the value of attribute i). In addition, each y_i must be scaled so that its range is the same as the other y_i functions (e.g. from 0 to 1). For example, $y_1 = \frac{\text{total delivered}}{\text{total requested}}$ has a range from 0 to 1, and its maximization improves the total amount of cargo delivered. Similarly, y_5 might be defined as the negative operations cost divided by its maximum value, so that its maximization reduces the cost and it ranges from 0 to 1.

This interpretation of the conjoint findings is more flexible, yet less precise, than the utility function developed earlier. Attribute importances can be interpreted more broadly than part-worth utilities, which are tied to specific levels of attributes. In this interpretation, each y_i can be defined as needed in the specific problem being studied, and the importances can be used as weights for each of these objective functions, in a multi-objective optimization problem. For future modeling efforts, the results of this survey, and particularly the importance weights for each attribute (Table 3.4), can be used to value trade-offs between efficiency, effectiveness, and equity.

3.7 Discussion and conclusions

The purpose of this study was to develop an objective function that convincingly represented the objectives of humanitarian aid. In particular, we sought to value the trade-offs between multiple objectives of aid delivery, to enable the development of penalty functions or objective weights that better model the real objectives of humanitarian aid. Measuring the preferences of expert humanitarian logisticians seemed the most direct route to determining the importance of humanitarian objectives. Five performance criteria were selected, based on the ways experts evaluated plans, and because they enable us to explore important trade-offs such as weighing the importance of efficiency, effectiveness, and equity, or the prioritization of different commodities. A conjoint analysis survey was developed and used to estimate the preferences of two groups of experts over these five attributes of aid delivery plans. The survey results quantify the ways experts traded off five objectives of aid delivery. In addition, we map the survey results to two forms of objective functions, a piecewise linear utility function and a weighted-sum objective function. The results of this paper should enable the development of optimization models that make better trade-offs between multiple objectives of humanitarian aid.

One important contribution of this chapter is in identifying important characteristics of average expert preferences across the five attributes considered here. Delivering aid is the most important objective, but the other attributes contribute significantly to the utility of plans, and therefore should not be left out of the objective function. Cost is not very important, unless it is extremely high, and even then it is less important than most other elements of a plan. This is an important insight, especially for modelers accustomed to working in the commercial sector with cost-based objectives. A "weighted mix" prioritization scheme is preferred to both lexicographic prioritization and even distribution, which was by no means clear from the start, since priority statements seem to suggest lexicographic prioritization while humanitarian missions often state a goal of equitably distributing aid. These insights point to the most important concerns of humanitarian aid delivery planners.

In addition to identifying preference patterns that were consistent across respondents, we have also identified preference patterns that differed among respondents. Cost, while less important than most other concerns, appeared to be more important to respondents from smaller organizations (Group 2). Other respondents described counter-intuitive preferences: some preferred later deliveries in order to spread them over time, and some preferred higher costs to generate more flexible resources. These preference patterns seem to apply in specific situations, and are important to understand alongside the average patterns.

Future work must address which of the estimated utility functions – group-level averages, or even the 30 individual part-worth utilities – represents the "right" objective for any given emergency response. Specific organizations, or particular operational constraints, might lead to a preference for one plan or another. Nevertheless, the Group 1 average utility function found in this paper represents the preferences of a group of experts specifically tasked with managing aid deliveries in emergency response, while the Group 2 results provide a contrasting example of the preferences of a different set of logisticians. Quantifying how these experts value trade-offs between multiple objectives is an important step forward.

Future work should also seek to clarify the conditions under which these various preferences apply, and develop families of objective functions suited to different types of preferences or specific situations. This work might also be extended by surveying humanitarian experts who design programs, in addition to logisticians. Our preference-derived utility functions might be compared with those based on explicit models of human suffering, or on the preferences of beneficiaries. Finally, it would be extremely useful, though challenging, to measure beneficiary preferences. Combining beneficiary and expert preferences might enable even better design of humanitarian interventions through a deeper understanding of the goals and opinions of those we seek to serve.

Developing a set of metrics that both met the requirements of conjoint survey techniques (readily understandable by humanitarian logisticians and independent) and led to well-formulated objective functions was a challenging goal for this work. Our approach privileged the design of the survey over the form of the resulting objective functions. The mapping of survey attributes to objective functions suggested in this paper is reasonable, but requires approximations; other interpretations might be developed as well. However, given our current lack of knowledge of the right penalty functions or weights for multi-objective problems, the results of this survey should prove useful in future modeling efforts. The implications of using these objective functions to develop transportation plans will be explored in future work.

A utility-based planning approach is likely to be most useful in the first week or two of an emergency, in which there is limited information on needs and priorities. As information becomes available, it may be possible to develop approaches that take into account more specific understanding of needs. For example, one community might urgently require water, even though in general it is low priority; or, if information on individual beneficiaries were available, it might be possible to track urgent medical needs of specific patients. The utility function developed in this paper measures performance in the aggregate, and is intended for use when this type of specific information is unavailable. Future work could explore ways to integrate aggregate and specific information.

The utility functions developed in this chapter are an important step forward. We have developed a framework for valuing trade-offs between objectives of humanitarian aid, including the identification of five key attributes of aid delivery plans. We have estimated part-worth utilities that represent the preferences of expert humanitarian logisticians, and mapped them to objective functions for use in transportation planning models. Finally, we have discovered some important characteristics of expert preferences across five key attributes of aid delivery plans, learning how they trade off aspects of efficiency, effectiveness, and equity. With these insights, it is possible to develop more appropriate objective functions for the optimization of humanitarian aid delivery.

References

Abernethy, J., Evgeniou, T., Toubia, O., and Vert, J.-P. (2008). Eliciting consumer preferences using robust adaptive choice questionnaires. *IEEE Transactions on Knowledge and Data Engineering*, 20(2):145–155.

Allenby, G. M. and Rossi, P. E. (2006). Hierarchical bayes model. In Grover, R. and Vriens, M., editors, *The handbook of marketing research: uses, misuses, and future advances*, chapter 20. Sage Publications, Inc.

Balcik, B., Beamon, B. M., and Smilowitz, K. (2008). Last mile distribution in humanitarian relief. *Journal of Intelligent Transportation Systems*, 12(2):51–63.

BBC (2010). Logistical nightmare hampers aid effort. BBC News, 22 Jan 2010.

Campbell, A., Vandenbussche, D., and Hermann, W. (2008). Routing for relief efforts. *Transportation Science*, 42(2):127–145.

Caunhye, A., Nie, X., and Pokharel, S. (2012). Optimization models in emergency logistics: A literature review. *Socio-Economic Planning Sciences*, 46:4–13.

CNN (2005). Much tsunami clothing aid 'wasted'. CNN, 5 Oct 2005.

de la Torre, L., Dolinskaya, I. S., and Smilowitz, K. (2011). Disaster relief routing: Integrating research and practice. *Socio-Economic Planning Sciences*, (in press).

Doerner, K., Focke, A., and Gutjahr, W. (2007). Multicriteria tour planning for mobile healthcare facilities in a developing country. *European Journal of Operational Research*, 179(3):1078–1096.

Dranove, D. (2003). *What's your life worth? : health care rationing – who lives? who dies? who decides?* Prentice Hall.

Forman, E. H. and Gass, S. I. (2001). The analytic hierarchy process - an exposition. *Operations Research*, 49(4):469–486.

Gold, M. R., Stevenson, D., and Fryback, D. G. (2002). Halys and qalys and dalys, oh my: Similarities and differences in summary measures of population health. *Annual Review of Public Health*, 23(115-34).

Green, P. E., Krieger, A. M., and Wind, Y. J. (2001). Thirty years of conjoint analysis: Reflections and prospects. *Interfaces*, 31(3):pp. S56–S73.

Green, P. E. and Srinivasan, V. (1978). Conjoint analysis in consumer research: Issues and outlook. *Journal of Consumer Research*, 5:103–123.

Grissom, M., Belegundu, A., Rangaswamy, A., and Koopman, G. (2006). Conjoint-analysis-based multiattribute optimization: Application in acoustical design. *Structural and Multidisciplinary Optimization*, 31(1):8–16.

Holguin-Veras, J., Perez, N., Jaller, M., Destro, L., and Wachtendorf, T. (2010). On the need to reformulate humanitarian logistics modeling: Deprivation costs, and material convergence (draft version). In *Proceedings of the 89th Transportation Research Board Annual Meeting.*

Huang, M., Smilowitz, K., and Balcik, B. (2012). Models for relief routing: Equity, efficiency and efficacy. *Transportation Research Part E: Logistics and Transportation Review*, 48:2 – 18.

Jamison, D., Bobadilla, J.-L., and Hecht, R., editors (1993). *The World Bank, World Development Report, 1993: Investing in Health.* Oxford University Press, New York.

Johnson, R. and Orme, B. K. (2010). Including holdout choice tasks in conjoint studies. Sawtooth Software Technical Paper Series.

Keeney, R. L. (1977). The art of assessing multiattribute utility functions. *Organizational Behavior and Human Performance*, 19:267–310.

Kuhfeld, W. F., Tobias, R. D., and Garratt, M. (1994). Efficient experimental design with marketing research applications. *Journal of Marketing Research*, 31:545–557.

Lenk, P. J., DeSarbo, W. S., Green, P. E., and Young, M. R. (1996). Hierarchical bayes conjoint analysis: Recovery of partworth heterogeneity from reduced experimental designs. *Marketing Science*, 15(2):pp. 173–191.

Lin, Y., Batta, R., Rogerson, P., Blatt, A., and Flanigan, M. (2011). A logistics model for emergency supply of critical items in the aftermath of a disaster. *Socio-Economic Planning Sciences*, 45(4):132–145.

Logistics Cluster (2012). Overview of logistics cluster operations 2010-2011. http://www.logcluster.org/about/logistics-cluster/background-information/overview-of-logistics-cluster-operations-2010-2011.

Marsh, M. T. and Schilling, D. A. (1994). Equity measurement in facility location analysis: A review and framework. *European Journal of Operational Research*, 74:1–17.

Meissner, M. and Decker, R. (2009). An empirical comparison of cbc and ahp for measuring consumer preferences. In *Proceedings of the International Symposium on the Analytic Hierarchy Process.*

Morrow, R. H. and Bryant, J. H. (1995). Health policy approaches to measuring and valuing human life: Conceptual and ethical issues. *American Journal of Public Health*, 85(10):1356–1360.

Nolz, P., Doerner, K., Gutjahr, W., and Hartl, R. (2010). A bi-objective metaheuristic for disaster relief operation planning. *Advances in Multi-Objective Nature Inspired Computing*, pages 167–187.

Orme, B. K. (2006). *Getting Started with Conjoint Analysis: Strategies for product design and pricing research.* Research Publishers.

Perez, N., Holguin-Veras, J., Mitchell, J. E., and Sharkey, T. C. (2010). Integrated vehicle routing problem with explicit consideration of social costs in humanitarian logistics. Submitted August 2010. http://transp.rpi.edu/ HUM-LOG/Doc/IRP.pdf.

Savas, E. S. (1978). On equity in providing public services. *Management Science*, 24(8):800–808.

Sawtooth (2009a). ACBC technical paper. Sawtooth Software Technical Paper Series.

Sawtooth (2009b). The CBC/HB system for hierarchical bayes estimation version 5 technical paper. Sawtooth Software Technical Paper Series.

Sawtooth (2009c). Which conjoint method should I use? Sawtooth Software Technical Paper Series.

Sawtooth (2011). Sawtooth software help files. SSI Web 7.

Toubia, O., Hauser, J. R., and Garcia, R. (2007). Probabilistic polyhedral methods for adaptive choice-based conjoint analysis: Theory and application. *Marketing Science*, 26(5):596–610.

Tzeng, G., Cheng, H., and Huang, T. (2007). Multi-objective optimal planning for designing relief delivery systems. *Transportation Research Part E: Logistics and Transportation Review*, 43(6):673–686.

Viswanath, K. and Peeta, S. (2003). Multicommodity maximal covering network design problem for planning critical routes for earthquake response. *Transportation Research Record: Journal of the Transportation Research Board*, 1857(-1):1–10.

Walsh, B. (2006). Fema says katrina taught it a lesson. Times-Picayune, 13 May 2006.

Chapter 4

Modeling and evaluating human decision-making approaches for humanitarian transportation planning

4.1 Introduction

Humanitarian supply chains are generally managed by people, with very few decision support tools. Without decision support, humans may struggle to manage the complexity in some logistics problems, such as the transportation planning problem studied in this thesis. On the other hand, decision support tools may require time and information to set up, both of which are in short supply in the first week after an emergency. To improve transportation planning in humanitarian supply chains, it is necessary to find the right balance between the complexity of decision support and the flexibility of human planners. In this chapter, we compare models of human decision-making processes to an optimization approach, to identify specific strengths and weaknesses of human planning approaches. Optimization approaches may create better plans, but we investigate how much could be gained with optimization compared with simpler decision support tools that could be implemented quickly and easily in the field. The ultimate goal is to find the best ways to improve transportation planning, whether by developing complex optimization-based planning approaches or simpler decision tools that correct specific weaknesses in human planning approaches.

In an ethnographic study, discussed in Chapter 2, we described the decision-making processes of teams of experienced humanitarian logisticians as they planned aid deliveries after a (simulated) emergency. We found that human decision-making resembles a greedy search process, guided by decision rules in the form of policies that prioritize deliveries of high-priority items or to high-priority locations. In greedy search, the search space and the policies that guide search have a major influence on the quality of the resulting solution. Our observations suggested that the teams struggled to solve this complex problem in the short time available, and many teams' plans were incomplete at the end of the day. Unfortunately, we could not evaluate the effectiveness of the specific decision processes and policies employed by the teams, based on the ethnographic study. The quality of each team's plan was influenced by too many confounding factors, including the mix of policies and decision processes employed, the dynamics of the team, and the computer skills of the team members. Moreover, because most teams used multiple types of policies and processes, either together or one after another, we could not sort out the effectiveness of each individual policy or process.

In this chapter, we attempt to understand the effectiveness of the human decision-making approaches discovered in the ethnographic study by developing algorithms that model the human processes and policies. The human decision-making approaches have an algorithmic character, and we develop heuristics based on their approaches. The processes found in the ethnographic study are not precise enough to describe every aspect of a computable heuristic, so additional assumptions are made in line with the researcher's experience in the ethnographic study. These models are *not* intended to replicate human behavior: as Chapter 2 showed, humans tried many strategies, started over, made mistakes, and generally moved through the problem in an unorganized manner. In contrast, these heuristics represent pure implementations of the decision-making processes observed within this unorganized human behavior, and their solutions show what would have happened had a team stuck to a single process and policy throughout their work. Models derived from human behavior perform better than humans themselves, in part because humans exhibit much more variation in their implementation (Bowman, 1963). By studying these pure versions of human approaches, we can identify their strengths and weaknesses.

The heuristic models are evaluated against one another and in comparison with a mixed-integer linear programming optimization approach. The optimization approach provides a useful standard against which to compare the models of human behavior, enabling us to identify the "gap" between the heuristic solutions and the best possible solution. It is not our goal to develop heuristics that minimize this gap, but rather to understand how well or poorly the human approaches perform and how much can be gained by using optimization. Note, however, that these performance evaluations are made in the "model world" described by the problem formulation, for which the optimized solution is (by definition) the best possible solution (Taylor and Iwanek, 1980). In reality, the optimized solutions (or heuristic solutions) may perform more poorly as the situation changes, but we do not consider these dynamics in this chapter. Instead, we use the "model world" to understand the fundamental strengths and weaknesses of different decision-making approaches.

In order to assess the performance of heuristic solutions and to guide the optimization approach, we require a performance measure that can convincingly evaluate the effectiveness of various transportation plans. Chapter 3 of this thesis developed a utility function, based on the preferences of expert humanitarian logisticians, that evaluates the utility of plans based on five key attributes: the total cargo delivered, prioritization of types of items, prioritization of destinations, speed of delivery, and cost. This utility function is employed as the objective in our optimization model and as a performance measure for the heuristic models.

Heuristics may perform well or poorly in different problems, depending on the specific cargo deliveries requested, network structure, and many other problem elements. In this chapter, heuristics are evaluated on two small stylized problems, each with several variations, in order to gain intuition about the heuristics' performance on specific, relevant problems, and show how performance changes as key problem dimensions are varied. In addition, the heuristics are evaluated on the full problem solved by the teams in the ethnographic study, which we call the "Snowland" scenario, to investigate heuristic performance on a realistic example of the type of problem these heuristics are trying to solve.

The remainder of this chapter is organized as follows. First, the remainder of this introduction provides a generic description of the transportation planning problem we study in this chapter. Please refer to Chapter 2 for a richer description of the problem faced by the teams in the ethnographic study. Section 4.2 reviews relevant literature, including studies of human behavior in operations management and existing models of humanitarian transportation planning. Section 4.3 develops the mixed-integer linear programming optimization model, and Section 4.4 describes the heuristic models of human decision-making processes. Section 4.5 describes the test problems and the strategy for evaluating the heuristics, and Section 4.6 provides the results of those evaluations. We conclude in Section 4.7.

4.1.1 The humanitarian transportation planning problem

We study the problem of planning shipments of cargo from various origins to various destinations, using a fleet of ground and air transportation assets. This section provides a brief description of this problem; a richer description was provided in Chapter 2 of this thesis. The problem arises when the Logistics Cluster provides common transportation services to its partner organizations. The partners submit Cargo Movement Requests (CMRs), requesting that their cargo be moved from an origin to a destination. The Cluster manages a fleet of trucks and helicopters, and must plan shipments to satisfy the demand for transportation. In practice, based on observations of teams solving this problem, the objective is not always clear. The goal is to deliver all the cargo as quickly as possible. However, demand for transport generally outstrips supply, so the prioritization of deliveries is important. Cargo is prioritized according to its type (e.g., shelter, health, water, food), its destination, or operational considerations such as utilization of vehicles.

The Cluster's fleet may contain vehicles of various types, each with capacities (weight and volume) and capabilities to travel on certain routes. They must operate within a network of road and air routes, each of which may be accessible only to certain types of vehicles. The problem is to allocate these vehicles to deliver a set of cargo movement requests from their respective origins to their respective destinations, where the cargo movement requests each contain different types of items with different weight and volume. In addition, the problem requires linking deliveries within a multi-modal network: many commodities must travel on multiple types of vehicles (e.g. first by truck, then by helicopter) to reach their destinations.

While this thesis focuses on the problem faced by the Logistics Cluster, other large humanitarian organizations face similar problems in managing fleets of vehicles and prioritizing deliveries, so this model should be useful beyond the specific context of the Logistics Cluster.

4.2 Literature Review

Two distinct sets of literature are relevant to this study. The first set of literature provides an understanding of how humans solve operational problems. The second focuses on building mathematical models to solve problems similar to the transportation planning problem. Each of these literatures is reviewed below.

4.2.1 Behavioral operations literature

Operations management problems have been studied extensively using mathematical models, but much less attention has been directed toward understanding how these problems are conceived and solved by humans. In other fields, such as psychology and economics, human decision-making is often compared to "rational" models of decision-making, which depend on mathematical models to determine the right choice. Researchers have shown that human behavior rarely matches such rational models, due at least in part to human cognitive limitations (Simon, 1955). People use reasoning heuristics to make judgments under uncertainty, which create problematic biases (Tversky and Kahneman, 1974; Kahneman et al., 1982). Even when mathematical models are based on the behavior of humans, the models perform better than humans. Expert judgments about uncertain outcomes, such as predictions of student performance based on application materials, are consistently inferior to linear regression models, even when the regression coefficients are randomly assigned (Hastie and Dawes, 2001). Humans are good at identifying the cues that matter to a prediction, but models are better at making the prediction based on the cues. In part, humans fail because they are sensitive to variation, while models do not have this weakness (Bowman, 1963). In general, research has shown that humans make worse decisions than rational models on a wide variety of problems, for a wide variety of reasons.

Most of the recent literature on human behavior in operations management settings (see Bendoly et al., 2006, for a review) focuses on inventory problems, including newsvendor models (Schweitzer and Cachon, 2000; Ben-Zion et al., 2008; Bolton and Katok, 2008), multi-period inventory models (Sterman, 1989; Croson and Donohue, 2002), and design of supply contracts (Kalkanci et al., 2011). Most of the effort is directed toward understanding when and why humans make poor decisions, and designing interventions to correct them. These efforts have yielded some general understanding of human failures in inventory problems. People tend to rely on simple rules or simple reasoning strategies (Schweitzer and Cachon, 2000; Sterman, 1989; Kalkanci et al., 2011; Moxnes, 1998). Anchoring and adjustment heuristics, in particular, are found in newsvendor (Schweitzer and Cachon, 2000) and multiperiod inventory (Sterman, 1989) problems. Evidence is also found for certain biases such as the gambler's fallacy and learning from small samples (Bolton and Katok, 2008), and bias toward the mean and influence from results in the previous round (Ben-Zion et al., 2008). More complex problems seem to require a larger reliance on simple rules (Kalkanci et al., 2011). Complex problems also inhibit learning, in part because people misinterpret the causes of poor results because they do not understand the problem structure and the performance feedback they receive (Sterman, 1989; Moxnes, 1998). They may use a simple rule that would be better suited to a less complex but similar version of the problem (Moxnes, 1998). Interventions to support better learning or counteract the biases have met with limited success (Ben-Zion et al., 2008; Bolton and Katok, 2008; Croson and Donohue, 2002).

Many of the findings in this literature are very problem-specific. In the newsvendor problem, for example, Schweitzer and Cachon (2000) find that people make newsvendor decisions by attempting to reduce inventory error (rather than profit loss), combined with anchoring and insufficient adjustment. Ben-Zion et al. (2008) and Bolton and Katok (2008) replicate these results, and find evidence of slow learning - progress toward but not to the optimal solution - over many rounds of solving the problem. They also find a series of biases that decline as learning occurs: the gambler's fallacy and learning from small samples (Bolton and Katok, 2008), and bias toward the mean and influence from results in the previous round (Ben-Zion et al., 2008). In the multiperiod inventory problem of the beer game, Sterman (1989) finds evidence of anchoring and adjustment from the desired inventory level. The key dynamic that contributes to failures in inventory management (the bullwhip effect) is the failure to account for delays and products in the supply line. People do not see this as the reason for their poor performance, and therefore do not learn. Croson and Donohue (2002) review later literature on the beer game, which experiments with various interventions intended to counteract these tendencies that produce the bullwhip effect. In contracting, Kalkanci et al. (2011) shows that people use simple rules to determine complex contracts, and that these lead to solutions which are no better than those of simple contracts. Finally, in a very complex bioresource management problem, Moxnes (1998) shows that even experts do not understand the dynamics and feedback structure of the problem, and use greedy rules that would work well in a simpler problem but lead to poor results in this one. The human weaknesses highlighted in this literature are usually specific to the dynamics of a particular problem situation; as a result, the interventions to correct these human weaknesses are also problem-specific.

The literature on human decision-making in inventory management shows that humans fail to make rational decisions for a variety of reasons, and these reasons are often specific to the decision being made. The transportation planning problem is quite different from the inventory models discussed thus far, so humans probably have different problem-specific heuristics. A solution requires not just one decision, such as an order quantity, but a long series of vehicle dispatch decisions that must be linked in a feasible manner. Many of the heuristics identified in inventory problems, such as anchoring and adjustment, would not be relevant; instead, we expect to find different kinds of heuristics in search-oriented problems like this one.

This type of search-oriented problem has been studied in an older literature that asks whether humans or algorithms are better at solving particular problems. Because most of this work was performed in the seventies and eighties, and algorithms have progressed significantly since then, some of the findings may no longer be valid. Researchers ran experiments that pitted humans, often with decision support tools, against state of the art algorithms (generally heuristics), and compared the speed and quality of their solutions. Studies conflict as to whether humans or heuristic algorithms are better problem-solvers, and as might be expected, this depends on the heuristic. This literature has included problems of similar complexity to the transportation planning problem, including routing problems (Krolak et al., 1972; Hill, 1982) and facility network design problems (Robinson and Swink, 1995; Swink and Robinson, 1997; Taylor and Iwanek, 1980). A few studies investigate what characteristics make problems easier or harder for humans to solve (Swink and Robinson, 1997; Taylor and Iwanek, 1980). My work updates this literature by taking a more in-depth look at when and *how* human solutions differ from optimized solutions, in a particularly complex routing problem.

The traveling salesman problem has been studied more recently. Humans are surprisingly good at solving traveling salesman problems when they are represented visually: they find solutions less than 15% worse than optimal with performance degrading only linearly with increasing problem size. Various heuristic mechanisms have been proposed to explain this performance, including clustering, avoiding crossed arcs, and utilizing the convex hull (MacGregor and Chu, 2011), but again these are specific to visually represented traveling salesman problems, and do not apply to our multi-modal, multi-vehicle routing problem.

In general, study of human behavior in solving operations problems has yielded the general insight that humans make worse decisions than models. This applies to models that are mathematically derived and models derived from human behavior (Bowman, 1963; Hastie and Dawes, 2001). However, the reasons for poor human decision-making are very problem-specific. The decision-making literature has not focused problems similar to the humanitarian transportation planning problem (or indeed multi-modal planning problems in general), so the specific weaknesses in human decision-making remain unclear. Understanding how humans make decisions is an important first step toward supporting better decision-making. This paper studies looks to identify specific weaknesses (or strengths) in human decision-making on a particularly complex transportation planning problem.

4.2.2 Humanitarian transportation planning literature

There is an extensive literature on mathematical models for transportation planning problems outside the humanitarian context, many of which are similar to the problem studied here. The most relevant set of models outside the humanitarian context deal with service network design for freight transportation, in which cargo from multiple clients is consolidated onto vehicles (Crainic, 2000; Crainic and Kim, 2007; Wieberneit, 2008). This type of problem arises in industries such as less-than-truckload carriers (Farvolden and Powell, 1994; Powell and Sheffi, 1983; Powell, 1986; Powell and Sheffi, 1989) and postal or express package services (Armacost et al., 2002; Grünert and Sebastian, 2000; Kim et al., 1999). Each such carrier operates a network of hubs and a fleet of vehicles, and must determine how and when both vehicles and cargo move within this network.

The problem can be written as a mixed-integer linear program, but the size of the problem prohibits this approach from being useful in most realistic problem instances. The primary challenge addressed by the literature is managing the scale of the problem. Researchers have introduced heuristic methods or reformulations, but these are generally very specific to the structure of the problem and difficult to transfer to other, similar problems. None of the models in the literature are directly applicable, though some of the formulations were useful in developing a mixed-integer linear programming formulation of the transportation planning problem. In particular, Crainic (2000) distinguishes between models that solve for a frequency of service (Armacost et al., 2002; Powell and Sheffi, 1983; Powell, 1986) and models that solve for a specific schedule (Farvolden and Powell, 1994; Grünert and Sebastian, 2000; Kim et al., 1999). This work follows the latter type of formulation, which utilizes a time-space network to model time explicitly.

In recent years, increasing attention been paid to logistics problems in the humanitarian context. Studying humanitarian logistics requires an understanding of the unique context of humanitarian action. To that end, a number of overview papers identified the key challenges faced in humanitarian logistics, described the activities and actors in humanitarian supply chains, and compared them to commercial supply chains with a view toward cross-learning (Long and Wood, 1995; Beamon, 2004; Van Wassenhove, 2006; Kovacs and Spens, 2007; Tatham and Kovacs, 2007; Thomas and Kopczak, 2006; Tomasini and Van Wassenhove, 2009; Apte, 2009; Ergun et al., 2010). The key challenges stem from the specific context of humanitarian logistics, including operating conditions, uncertainty in needs, coordination, performance measurement, information technology, and last mile transportation (Kovacs and Spens, 2007; Van Wassenhove, 2006; Beamon, 2004; Gustavsson, 2003; Ergun et al., 2010). Several papers identify distinct phases of the relief supply chain, along the lines of those delineated by Kovacs and Spens (2007): a pre-disaster preparation phase and a post-disaster immediate response phase transitioning later to reconstruction. Some authors provide descriptions of specific emergencies from which they derive key challenges and lessons learned (Larson et al., 2006; Van Wassenhove, 2006; Holguin-Veras et al., 2007; Russell, 2005). A broad range of papers describe problems and solutions in humanitarian logistics from a practice orientation, either identifying issues across organizations or citing specific organizational or disaster contexts (Chaikin, 2003; Chomilier et al., 2003; Fenton, 2003; Gustavsson, 2003; Kaatrud et al., 2003; Molinaro and Blanchet, 2003). These contextual analyses of humanitarian logistics identify a broad range of challenges, some of which are amenable to modeling while others present issues better addressed by organizational or information technology solutions.

Modeling is particularly suited to addressing problems in meeting uncertain needs and planning inventory and transportation. Caunhye et al. (2012), in their review of optimization models in humanitarian logistics, identify the main problems studied in the recent literature: facility location in advance of a disaster, often combined with inventory pre-positioning or distribution; and relief distribution after a disaster, considering resource allocation, commodity flow, or (more commonly) both. De la Torre et al. (2011) review models that specifically address transportation and routing for relief distribution, and Altay and Green (2006) review older literature on disaster response. Both recent reviews provide useful summaries and discussions of the humanitarian logistics modeling literature, which we also summarize in the following paragraphs.

One set of models examines inventory management, facility location, and other resource positioning problems. One well-studied problem is inventory pre-positioning, in which inventory is allocated in the pre-disaster phase in anticipation of future emergency responses (Akkihal, 2006; Duran et al., 2011; Mete and Zabinsky, 2010; Rawls and Turnquist, 2010; Zhu et al., 2008). While inventory management before a disaster has been examined by a number of researchers, only a few have looked at inventory management after a disaster or in long-term humanitarian operations (Beamon and Kotleba, 2006a,b). Beyond inventory, a few models study the positioning of a wider range of resources, such as warehouses, capacity, fleet, and points of distribution (Salmerón and Apte, 2010; Pedraza Martinez et al., 2010; Jaller and Holguin-Veras, 2011).

More relevant to this research are those models that examine the routing of vehicles in the post-disaster context. An extensive review is provided by de la Torre et al. (2011), so it is not repeated here. Instead, the characteristics of these models are summarized, and those most relevant are identified. Many variations on transportation planning problems for humanitarian aid delivery have been studied by operations researchers. Most are based on vehicle routing formulations (e.g. Balcik et al., 2008; Huang et al., 2012; Campbell et al., 2008; Lin et al., 2011; Doerner et al., 2007; De Angelis et al., 2007) and network flow formulations (e.g. Haghani and Oh, 1996; Perez et al., 2010; Yi and Ozdamar, 2007; Yi and Kumar, 2007). Others combine routing decisions with other types of decisions, such as resource allocation or facility location (Barbarosoglu et al., 2002; Ukkusuri and Yushimito, 2008; Tzeng et al., 2007). Most models are deterministic, but some include stochasticity (Mete and Zabinsky, 2010; Van Hentenryck et al., 2010; Shen et al., 2009a,b).

This paper employs a multi-commodity, multi-modal network flow formulation over multiple time periods, which models the flow of both vehicles and cargo throughout the network. Similar formulations have been employed previously, but with different objective functions and other variations. Clark et al. (2007) explore various formulations and show how the behavior of the solutions changes with different elements. Among other conclusions, they show that in general the cheapest and closest areas are served first, that including prioritization of deliveries improves the equity of solutions, and that limited truck capacity affects how air transportation is utilized. Their objective function privileges the minimization of unsatisfied demand over other components of the objective function. Haghani and Oh (1996) minimize cost plus time until demand is satisfied. They explore two solution algorithms, but provide little discussion of the solutions. Ozdamar et al. (2004) minimize unsatisfied demand, and discuss a re-planning model to update solutions as information changes. The solution methodology is based on Lagrangian relaxation. Yi and Ozdamar (2007) add evacuation of wounded and health center service levels to the problem, again minimizing unmet demand along with unserved wounded. The results for an Istanbul earthquake scenario show that the model balances service rates of medical facilities, incorporating delays in delivering commodities, but occasionally neglects affected people in remote areas. Yi and Kumar (2007) solve a similar problem using ant colony optimization. Perez et al. (2010) formulate an interesting objective function based on deprivation costs, which increase with time as demand remains unsatisfied. They incorporate this objective into multi-commodity network flow formulations, noting how the objective makes them challenging to solve. Holguin-Veras et al. (2010) provide small example problems to illustrate how the deprivation cost objective helps to ensure equity in service among different locations, resulting in reduced deprivation costs compared to simpler objectives.

Our goal is not to develop new formulations or solution methods, but to explore the implications of the new objective function developed in Chapter 3 (and to examine human decision approaches in comparison with proven optimization approaches). The models mentioned in the preceding paragraphs employ a wide variety of different objectives, across categories of efficiency, equity, and effectiveness. Efficiency objectives typically minimize operations costs (Tzeng et al., 2007; Balcik et al., 2008) or travel times (Tzeng et al., 2007; Lin et al., 2011; Huang et al., 2012; Campbell et al., 2008). Effectiveness objectives maximize some measure of service, often either the amount of demand satisfied (Lin et al., 2011), the speed with which demand is satisfied, or both (Huang et al., 2012; Balcik et al., 2008). Equity objectives encourage models to provide service equally to all recipients, by minimizing the latest arrival time (Campbell et al., 2008), minimizing the sum of arrival times (Campbell et al., 2008), maximizing the smallest demand satisfaction rate (Tzeng et al., 2007), minimizing the disparity in demand satisfaction (Lin et al., 2011), or imposing penalties on inequitable deliveries (Huang et al., 2012). Chapter 3 of this thesis developed an objective function that trades off multiple objectives based on measured preferences of humanitarian logisticians; it goes beyond the models described above by using empirical measurements to determine the trade-off between multiple objectives. This chapter explores the implications of using such an objective function in a transportation planning optimization problem. (See Chapter 3 for further discussion of objectives in the humanitarian modeling literature.)

4.3 Optimization formulation

In this section, we develop a mixed-integer linear programming formulation of the humanitarian transportation planning problem described in the introduction. The formulation models the flow of both vehicles and cargo through the network.

4.3.1 Structure, Data, and Decision Variables

We are given a set $\mathcal{R}$ of cargo movement requests $\{r_1, r_2, ..., r_n\}$, each of which has a quantity of cargo to be transported Q_r, an origin O_r, a destination D_r, and an available time T_r, meaning the cargo is available at the beginning of period T_r at the location O_r. Each request contains a single type of commodity, so each request also has a unit weight W_r and a unit volume V_r. Each type of commodity is given an item priority level P_r based on the type of cargo, and a destination priority level P'_i based on its final destination.

$\mathcal{R} = \{r_1, ..., r_n\}$	set of cargo movement requests
Q_r	quantity of cargo in request r
O_r	origin of request r
D_r	destination of request r
T_r	time at which request r becomes available at its origin
W_r	unit weight of cargo in request r
V_r	unit volume of cargo in request r
P_r	priority of type of item in request r
P'_i	priority of destination i

Time is discretized, and each period is equal to some fraction of a work day. We assume no work overnight, so one period represents the last part of one work day and the next represents the first part of the next work day. It is necessary for modeling purposes to know which periods occur just after an "overnight", so we define the set $\mathcal{D}$ of periods that represent the morning of a new day. We assume that all travel times are integer multiples of a single period, and all cargo and vehicle arrivals and departures occur at the beginning of a period. We plan transportation over a horizon of N_t periods.

N_t	number of periods in planning horizon
$\mathcal{D}$	set of periods that represent the morning of a new day

There is a set of possible modes $m \in \mathcal{M}$ of transportation: for example, one set includes 40-ton trucks m_1, 10-ton trucks m_2, and helicopters m_3. Each mode m describes a single vehicle type, so each mode has a weight capacity per vehicle C_m^W and a volume capacity per vehicle C_m^V. Each vehicle also has a cost F_m, in dollars per ton-km for road vehicles and in dollars per rotation for helicopters. The number of vehicles of type m initially at each node i is given by Z_i^m. Let N_v represent the total number of vehicles available over all modes over time (this quantity is defined because we use it later as a number greater than or equal to the number of vehicles available in the scenario).

Vehicles can become available (i.e. be added to the scenario) at any period t throughout the planning horizon. The number of vehicles of type m that are newly available in period t is given by T_m^t. All vehicles of type m become available at a single node O_m, but can subsequently make unrestricted movements.

Because helicopters and trucks have different restrictions, it is necessary to note which modes are helicopters. Therefore, the set $\mathcal{M}_H$ is defined as the set of helicopter transport modes.

$\mathcal{M} = \{m_1, ..., m_n\}$	set of possible transportation modes
$\mathcal{M}_H$	set of helicopter transport modes
C_m^W	weight capacity per vehicle of type m
C_m^V	volume capacity per vehicle of type m
F_m	cost per ton-km or rotation for vehicles of type m
Z_i^m	number of vehicles of type m initially at location i
N_v	total number of vehicles available over all types
O_m	arrival location for vehicles of type m
T_m^t	number of vehicles of type m arriving in period t

We use the type of network shown in Figure 4-1. In this example, the nodes P (for the major arrival port), A, and B represent the hubs within the network, s is a source node, s' is a sink node, and the remaining nodes represent final destination points. All the nodes labeled D represent final destinations, some of which are reachable by truck or by helicopter only, some in short flights and some in long flights. Any network of this type can be used within this formulation.

Let $\mathcal{N}$ be the set of nodes in the network, and let $\mathcal{N}_p$ be the set of nodes representing physical locations, i.e. all the nodes except the source s and the sink s'. Storage is available

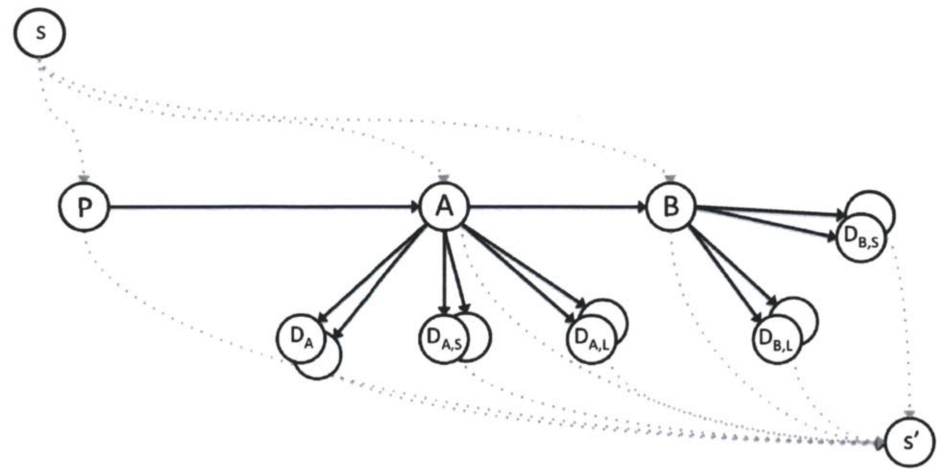


Figure 4-1: Network

at some nodes; let the storage capacity at node i be given by C_i. Some nodes serve as helicopter bases; let B_i be 1 if they do and zero otherwise. (Truck bases are unnecessary because fuel is available everywhere.) Deliveries to some nodes are prioritized over others, so let P_i' be the priority of delivering to node i. Let $\mathcal{A}$ be the set of arcs in the network, and let $\mathcal{A}_p$ be the set of arcs representing physical movement, i.e. all the arcs except those to and from the source s and the sink s'.

Each arc representing physical movement, i.e. each $(i, j) \in \mathcal{A}_p$, can be traveled by some subset of all the types of vehicles (e.g. helicopter, 40T truck, or 10T truck). A_{ij}^m describes which routes can be traveled by which modes, and is equal to 1 if arc (i, j) can be traveled by vehicles of type m, and 0 otherwise. Each such arc also has a time to traverse L_{ij}^m, which is an integer number of periods. The source and sink arcs, $(i, j) \in \mathcal{A} \backslash \mathcal{A}_p$, have no capacity limits and require no 'vehicles'; they take no time to traverse, i.e. $L_{ij}^m = 0$. All arcs also have a distance used to calculate the cost, which differs based on the mode used to traverse it; this "billable distance" is given by E_{ij}^m.

$\mathcal{N}$	set of nodes, including source and sink nodes
$\mathcal{N}_p$	set of nodes representing physical locations
C_i	storage capacity at node i
B_i	indicates whether node i is a helicopter base (1) or not (0)
P_i	priority of delivering to node i
$\mathcal{A}$	set of arcs in the network, including source and sink arcs
$\mathcal{A}_p$	set of arcs representing physical movement
A_{ij}^m	1 if arc (i, j) can be traveled by vehicles of type m, and 0 otherwise
L_{ij}^m	time to traverse arc (i, j) by vehicles of type m
E_{ij}^m	"billable" distance for mode m to traverse arc (i, j)

Four sets of decision variables are defined. The variables $x_{ij}^{r,m,t}$ represent the movement of cargo through the network, and the variables $y_{ij}^{m,t}$ represent the movement of vehicles (and therefore transport capacity) through the network. The variables $w_i^{r,t}$ and $z_i^{m,t}$ represent the inventory of cargo and vehicles, respectively, at each node in the network.

$x_{ij}^{r,m,t} =$ quantity of request r leaving by mode m on arc $(i, j) \in \mathcal{A}$ at the beginning of period $t \in \{1, ..., N_t\}$

$w_i^{r,t} =$ quantity of request r available in inventory at location $i \in \mathcal{N}_p$ at the beginning of period $t \in \{0, ..., N_t\}$, where $t = 0$ represents the initial conditions

$y_{ij}^{m,t} =$ number of vehicles of type m leaving on arc $(i, j) \in \mathcal{A}$ at the beginning of period $t \in \{1, ..., N_t\}$, $y_{i,j}^{m,t} \in \{0, 1, 2, ...N_v\}$

$z_i^{m,t} =$ number of vehicles of type m available at location $i \in \mathcal{N}_p$ at the beginning of period $t \in \{0, ..., N_t\}$, $z_i^{m,t} \in \{0, 1, 2, ...N_v\}$

with $r \in \mathcal{R}$, the set of requests for transport, and $m \in \mathcal{M}$, the set of possible modes of transportation. While the vehicles are represented by integer variables, note that $x_{ij}^{r,m,t}$ is continuous, so requests can be split.

Note that for cargo and vehicle movements, $x_{ij}^{r,m,t}$ and $y_{ij}^{m,t}$ respectively, t is defined over $t \in \{1, 2, ..., N_t\}$ periods; but for the cargo and vehicle inventory, $w_i^{r,t}$ and $z_i^{m,t}$ respectively, t is defined over $t \in \{0, 1, 2, ..., N_t\}$ periods, where $t = 0$ represents the initial conditions.

4.3.2 Objectives

Multiple objectives are employed in order to balance the multiple goals of aid delivery. In Chapter 3, we found utility functions over five measures of the performance of humanitarian aid delivery plans: the total cargo delivered, the prioritization of deliveries by cargo type, the prioritization of deliveries by location, the speed of delivery, and the operations cost. In this section, we develop an objective function based on the utility functions found earlier. In each case, we first develop a performance measure for each of the five objectives, $f_1, f_2, ..., f_5$, that can be calculated based on any given set of variables describing a plan. The utility contributed by each objective is then $u_i(f_i)$, where the u_i are piecewise linear functions. The utility of a given plan is the sum of the utilities of each of the five objectives:

$$U = u_1(f_1) + u_2(f_2) + u_3(f_3) + u_4(f_4) + u_5(f_5) \tag{4.1}$$

Many of the utility functions depend on the amount of cargo delivered (in total, of various priority types, etc.), but "amount" can be measured in various ways. In this formulation, we use metric tons, but alternatives could include cubic meters or number of items. It would be useful to define a measure that corresponds to the number of people assisted rather than the physical attributes of the cargo, but such a measure has not been defined for the many varieties of cargo transported by the Logistics Cluster.

Total Cargo Delivered First, we measure the total amount of cargo delivered divided by the amount of cargo requested. Let the total metric tons of cargo delivered be given by R:

$$R = \sum_{(i,s') \in \mathcal{A}, r, m, t} x_{i,s'}^{r,m,t} W_r \tag{4.2}$$

Then the fraction of requested cargo delivered is given by

$$f_1 = \frac{R}{S} \tag{4.3}$$

where S is the total tons of cargo requested:

$$S = \sum_r Q_r W_r \tag{4.4}$$

Based on the utilities estimated from the survey described in Chapter 3, the utility function for total cargo delivered is:

$$u_1(f_1) = \begin{cases} 571f_1 - 322, & f_1 < 0.6 \\ 264f_1 - 138, & f_1 > 0.6 \end{cases} \tag{4.5}$$

Prioritization by cargo type The second objective measures the prioritization of deliveries by cargo type. Let the tons of items of priority p delivered be given by R_p

$$R_p = \sum_{r:P_r=p,(i,s')\in\mathcal{A},m,t} x_{i,s'}^{r,m,t} W_r \tag{4.6}$$

and $\tilde{R}_{max}$ be an estimate of the maximum possible deliveries (such an estimate may be obvious from the problem formulation, or it may be obtained by solving a maximum flow problem). The vector $\mathbf{r}$ describes the fraction of deliveries allocated to each of the four types of items:

$$\mathbf{r} = [\frac{R_1}{\tilde{R}_{max}}, \frac{R_2}{\tilde{R}_{max}}, \frac{R_3}{\tilde{R}_{max}}, \frac{R_4}{\tilde{R}_{max}}] \tag{4.7}$$

To create a one-dimensional measure of prioritization by item type, we take the Manhattan distance from $\mathbf{r}$ to the point $\hat{\mathbf{r}} = [0.5, 0.3, 0.15, 0.05]$, the most preferred prioritization vector. This preferred vector corresponds to allocating 50% of deliveries to priority-1 items, 30% of deliveries to priority-2 items, 15% of items to priority-3 items, and 5% of items to priority-4 items. However, there should be no penalty for delivering more than the ideal amount, only for delivering less. Therefore, a "shortfall Manhattan distance" is employed, in which only the shortfall from $\mathbf{r}$ to $\hat{\mathbf{r}}$ is counted, so each component of the distance is given by $\max[0, \hat{r}_p - r_p]$. It is more important to deliver the higher-priority cargo, so the distance is weighted by $\hat{\mathbf{r}}$, resulting in the second performance measure:

$$f_2 = \sum_p \hat{r}_p \max[0, \hat{r}_p - r_p] \tag{4.8}$$

The utility function found in the survey has different slopes for deviations from the ideal vector $\hat{\mathbf{r}}$, depending on whether the deviation is toward the "lexicographic prioritization" vector $[1, 0, 0, 0]$ or toward the "even mix" vector $[0.25, 0.25, 0.25, 0.25]$. The function c_2 is used to distinguish between deviations from $\hat{\mathbf{r}}$ in each of these directions:

$$c_2 = \sum_p \hat{r}_p(r_p - \hat{r}_p) \tag{4.9}$$

When $c_2 > 0$, the deviation is lexicographic, and when $c_2 < 0$, the deviation is towards an even mix or deliveries of lower-priority items. The function measures the difference between the actual (r_p) and ideal ($\hat{r}_p$) deliveries of each priority p, which will be positive for over-deliveries and negative for under-deliveries. The differences are summed, but weighted by $\hat{r}_p$ so that deliveries of more important priorities count more heavily. Thus, when c_2 is negative, the higher priority items are under-delivered, meaning the prioritization vector is deviating from the ideal in the "even mix" direction. When c_2 is positive, the higher priority items are over-delivered, meaning the prioritization vector is deviating from the ideal in the "lexicographic" direction.

Estimating a piecewise linear utility function from the survey results, as described in Chapter 3, results in the utility function

$$u_2(f_2) = \begin{cases} -0.79546 f_2 + 49, & c_2 > 0 \\ -1.39509 f_2 + 49, & c_2 < 0 \end{cases} \tag{4.10}$$

Prioritization by destination The third objective measures the prioritization of deliveries by the priority of their destinations. The structure of the objective function is the same as that of the prioritization of deliveries by item type, described above. Here, let the number of deliveries to destinations of priority p' be given by $R'_{p'}$

$$R'_{p'} = \sum_{(i,s') \in \mathcal{A}: P'_i = p', r, m, t} x_{i,s'}^{r,m,t} W_r \tag{4.11}$$

The vector $\mathbf{s}$ describes the fraction of deliveries allocated to each of the two destination priority-levels

$$\mathbf{s} = [\frac{R'_1}{\tilde{R}_{max}}, \frac{R'_2}{\tilde{R}_{max}}] \tag{4.12}$$

The ideal point in this case is $\hat{\mathbf{s}} = [0.7, 0.3]$, and again a "shortfall Manhattan distance" is used to measure deviations from the ideal point, resulting in the performance measure

$$f_3 = \sum_{p'} \hat{s}_{p'} \max[0, \hat{s}_{p'} - s_{p'}] \tag{4.13}$$

As before, a cutoff function c_3 determines whether the deviation from $\hat{\mathbf{s}}$ is in the lexicographic or even direction:

$$c_3 = \sum_{p'} \hat{s}_{p'}(s_{p'} - \hat{s}_{p'}) \tag{4.14}$$

and the utility function, based on the estimated utilities from the survey, is

$$u_3(f_3) = \begin{cases} -0.331 f_3 + 27.6, & c_3 > 0 \\ -0.9428 f_3 + 27.6, & c_3 < 0 \end{cases} \tag{4.15}$$

Speed of delivery The fourth objective measures the speed of delivery. The fourth performance measure is the average day of delivery of cargo, which is measured here by the following function:

$$f_4 = \frac{\sum_{(i,s') \in \mathcal{A}, r, m, t} t x_{i,s'}^{r,m,t} W_r + t_{\text{final}} \max[0, \tilde{R}_{max} - R]}{\tau \tilde{R}_{max}} \tag{4.16}$$

where R is the total amount of cargo delivered and τ the number of time periods per day. It is computed by finding the average (by weight) time period of delivery, which is converted to a value in days by dividing by the number of time periods per day. The first term of the numerator is the average time of delivery of cargo by weight, and the second term of the numerator ensures that if the maximum possible cargo is not delivered, the shortfall is counted as delivered in the last time period. This ensures that the performance measure cannot be increased simply by delivering less total cargo (and also results in slightly optimistic performance measure, since undelivered cargo would in reality be delivered later than the last time period).

The utility function estimated based on survey responses was extremely close to a line, so a least squares estimate of the line is used in the utility function:

$$u_4(f_4) = -40f_4 + 153 \tag{4.17}$$

Operations cost The final performance measure is the operations cost. The cost structures are different for helicopters and for road vehicles. The first term is for road vehicles, where the cost depends on the weight of the cargo on the truck. The second term is for helicopters, where cost is incurred whether or not the vehicle is loaded.

$$f_5 = \sum_{m \in \mathcal{M} \setminus \mathcal{M}_H, (i,j) \in \mathcal{A}, r, t} x_{i,j}^{r,m,t} W_r F_m E_{i,j}^m + \sum_{m \in \mathcal{M}_H, (i,j) \in \mathcal{A}, t} y_{i,j}^{m,t} F_m E_{i,j}^m \tag{4.18}$$

The utility function estimated from the survey responses gives u_5, below, where cost is in units of millions of dollars.

$$u_5(f_5) = \begin{cases} -3f_5 + 14, & f_5 < 2 \\ -20f_5 + 49, & f_5 > 2 \end{cases} \tag{4.19}$$

4.3.3 Constraints

Flow and inventory balance constraints are included for cargo,

$$w_i^{r,t-1} + \sum_m \sum_{j:(j,i) \in \mathcal{A}, L_{ji}^m < t} x_{ji}^{r,m,t-L_{ji}^m} - \sum_{j:(i,j) \in \mathcal{A}, m} x_{ij}^{r,m,t} = w_i^{r,t} \quad \forall\, i \in \mathcal{N}_p, r, t \geq 1 \tag{4.20}$$

where L_{ij}^m is the time to traverse arc (i, j) by mode m. The first term represents the inventory at node i, and the second and third terms represent cargo entering and leaving (respectively), at the beginning of each time period t. Similar flow and inventory balance constraints are included for vehicles:

$$z_i^{m,t-1} + \sum_{j:(j,i)\in\mathcal{A}, L_{ij}^m < t} y_{ji}^{m,t-L_{ij}^m} - \sum_{j:(i,j)\in\mathcal{A}} y_{ij}^{m,t} = z_i^{m,t} \qquad \forall\, i \in \mathcal{N}_p, m, t \geq 1 \tag{4.21}$$

Capacity constraints for both inventory and flow can be included. Cargo capacity on each arc is restricted according to the vehicle capacity available on that arc, according to both the weight and volume of the vehicles and cargo:

$$\sum_r x_{i,j}^{r,m,t} \cdot W_r \leq y_{i,j}^{m,t} \cdot C_m^W \qquad \forall\, (i,j) \in \mathcal{A}_p : i,j \neq s, s', m, t \geq 1 \tag{4.22}$$

$$\sum_r x_{i,j}^{r,m,t} \cdot V_r \leq y_{i,j}^{m,t} \cdot C_m^V \qquad \forall\, (i,j) \in \mathcal{A}_p : i,j \neq s, s', m, t \geq 1 \tag{4.23}$$

where W_r and V_r are the unit weight and volume of each request r, and C_m^W and C_m^V are the weight and volume capacity of each vehicle of type m, respectively.

Only certain types of vehicles can travel on each arc. Let A_{ij}^m equal 1 if type m vehicles can travel on arc (i,j), and 0 otherwise. Then,

$$y_{ij}^{m,t} \leq A_{ij}^m \cdot N_v \qquad \forall\, (i,j) \in \mathcal{A}, m, t \geq 1 \tag{4.24}$$

ensures that the number of vehicles of type m traveling on arc (i,j) is zero when such vehicles cannot travel on this arc, and unrestricted otherwise (because N_v is the maximum number of vehicles in the scenario; this could be written by mode but is presented this way for simplicity).

The next set of constraints govern the arrival and delivery of each request,

$$x_{s,j}^{r,m,t} = 0 \qquad \forall\, (s,j) \in \mathcal{A} : j \neq O_r, r, m, t \geq 1 \tag{4.25}$$

$$x_{s,O_r}^{r,m,t} = 0 \qquad \forall\, (s,O_r) \in \mathcal{A}, r, m, t \neq T_r, \geq 1 \tag{4.26}$$

$$\sum_{m} x_{s,O_r}^{r,t} = Q_r \qquad \forall\, (s,O_r) \in \mathcal{A}, r, m, t = T_r \tag{4.27}$$

$$x_{i,s'}^{r,m,t} = 0 \qquad \forall\, (i,s') \in \mathcal{A} : i \neq D_r, r, m, t \geq 1 \tag{4.28}$$

$$\sum_{m,t} x_{D_r,s'}^{r,m,t} \leq Q_r \qquad \forall\, (D_r, s') \in \mathcal{A}, r \tag{4.29}$$

where constraints (4.25), (4.26), and (4.27) ensure cargo is available only at its origin O_r starting at the start of the period during which it arrives T_r, and constraint (4.28) ensures that cargo can be "delivered" only from its destination node D_r. Recall that the objective maximizes cargo delivery to destinations, but this constraint can ensure that certain amounts of specific requests are delivered. Finally, constraint (4.29) ensures that no more than the total amount of the request can be delivered.

In a similar manner, the addition and removal of vehicles throughout the planning horizon is restricted:

$$y_{s,j}^{m,t} = 0 \qquad \forall\, (s,j) \in \mathcal{A} : j \neq O_m, m, t \geq 1 \tag{4.30}$$

$$y_{s,O_m}^{m,t} = T_m^t \qquad \forall\, (s,O_m) \in \mathcal{A}, m, t \geq 1 \tag{4.31}$$

$$y_{i,s'}^{m,t} = 0 \qquad \forall\, (i,s') \in \mathcal{A}, m, t \tag{4.32}$$

where constraint (4.30) ensures vehicles can only be added at their mode's origin O_m, and constraint (4.31) sets the number that are added at each time period to T_m^t. Constraint (4.32) ensures that vehicles do not leave the network through the sink node.

We set the initial conditions for both the cargo and vehicle inventories:

$$w_i^{r,t} = 0 \qquad \forall\, i \in \mathcal{N}_p, r, t = 0 \tag{4.33}$$

$$z_i^{m,t} = Z_i^m \qquad \forall\, i \in \mathcal{N}_p, m, t = 0 \tag{4.34}$$

Several additional constraints are included to model specific requirements for helicopters. First, helicopters must refuel after every "rotation". Therefore, helicopters can only make out-and-back trips from a helicopter base. To enforce this restriction, we require that every helicopter trip either begin or end (or both) at a node that is a helicopter base. B_i is 1 if node i is a helicopter base and zero otherwise, so constraint (4.35) requires that for every arc from i to j, either the arc is not used by a helicopter, or one of the nodes i or j is a helicopter base, or both nodes are helicopter bases.

$$(y_{ij}^{m,t})(B_i + B_j - 1)(B_i + B_j - 2) = 0 \qquad \forall\, (i,j) \in \mathcal{A}_p, m \in \mathcal{M}_H, t \tag{4.35}$$

Finally, all decision variables are positive:

$$x_{ij}^{r,m,t}, w_i^{r,t}, y_{ij}^{m,t}, z_i^{m,t} \geq 0 \qquad \forall\, i, j, r, m, t \tag{4.36}$$

The optimization model was implemented in IBM ILOG CPLEX Optimization Studio 12.2. The test problems and results will be described below.

4.4 Models of human approaches

This section develops models of the human approaches identified in Chapter 2 of this thesis. First, we describe the development of a format or framework for describing and evaluating transportation plans. Second, we describe heuristic models of several human approaches.

4.4.1 An Excel-based tool for describing and evaluating plans

In the optimization model described above, the decision variables provide a format in which to concisely describe the output of the optimization: a transportation plan. However, this format is not complete enough to give to a transporter, because it lacks assignment of cargo to specific vehicles, along with other details of interest to real planners. In order to implement the human approaches observed in Chapter 2, we needed to develop a more complete format for transportation plans that was more intuitively intelligible to humans. Such a format was developed within Microsoft Excel, because it is the only piece of software guaranteed to be available to transportation planners in the field. This Excel application can be used as a simple decision support tool for planning in the field; various versions of it were tested as tracking tools in the Logistics Response Team (LRT) training described in Chapter 2, and in classroom exercises.

To describe and evaluate plans, the Excel-based tool represents problem data and solutions, evaluates the feasibility of a given plan, and computes a set of performance metrics. A series of macros, written in Visual Basic for Applications, enable this functionality. The problem data, the solutions, and the performance metrics can be easily inspected and edited by the user.

The tool consists of a set of worksheets that form a relational database describing the problem data, a set of input sheets enabling users to manually enter a plan, output sheets that describe the input plan and show its performance metrics, and VBA macro routines that perform all calculations. The problem data must be entered into the tool first; problem data include vehicles and their characteristics, sites, routes between sites, and cargo movement requests. Input sheets enable users to describe a transportation plan by noting each movement of a vehicle and the cargo it carries. Once movements are entered, the feasibility of the plan can be checked, and messages inform the user exactly where problems occur. Once a feasible plan has been created, the plan can be viewed in the output sheets in various formats and evaluated according to a pre-defined set of metrics. Three plan views provide different perspectives on the input transportation plan. The vehicle view shows all movements of all vehicles, including the cargo they carry. The CMR view shows all cargo movement requests, along with all movements made by any cargo belonging to the CMR. The site view shows all sites and their inventory at a given time. The metrics sheet calculates a series of performance metrics, including the total amount and percent of cargo delivered, broken down by item type priority, destination priority, and requesting organization; the speed of delivery; the cost; and the efficiency in terms of vehicle utilization.

This tool was provided to logisticians participating in the LRT training, and also tested in two classroom exercises. Even without any ability to automatically generate plans, it turned out to be useful as a "tracking" tool for maintaining an understanding of the evolving plan as it was input by hand. As people decided on the dispatch of each individual truck, it was helpful to use the plan views to determine where vehicles would become available, and when cargo would arrive at a given location. The plan was difficult to track by hand, because there were many vehicles moving through a complex network and because many of the cargo movement requests had to be split into several loads, each of which moved on several different modes of transport. The tool enabled planners to track their evolving plan as they created it.

Ultimately, such a tool would be more useful if it could also suggest plans to the user. Within the Excel-based tool, it is possible to add routines that generate plans or that import plans captured in other formats. The models of human approaches, described below, were developed and implemented as VBA macros within this planning tool. In addition, we developed a routine to read in the output of the CPLEX optimization, so that plans generated by the optimizer can be viewed more intuitively and scored in the same way as the human approaches.

4.4.2 Modeling human approaches: heuristics and policies

Chapter 2 of this thesis described an ethnographic study of 10 teams of humanitarian logisticians as they created transportation plans, solving the problem described in this chapter. Six teams had no access to decision support, and two teams had access to the planning tool described in the previous section, which could track a manually input plan, but could not generate any elements of the plan. As a result, all teams had to make planning decisions "by hand", meaning that they had to decide every movement of every vehicle: at what time it would leave, where it would go, and what cargo it would carry.

My study showed there were two archetypal decision processes followed by many teams: the location-based and task-based decision processes described in Chapter 2. In both processes, teams progressed forward in time, first planning the first day, then the second, and so on. In the location-based process, teams focused on a particular location, or node in the network, at the "current" time, meaning the time to which they had planned so far. Within the vehicles and cargo available at that node, they chose cargo to load and a destination to move to, generally sending cargo forward towards its destination. After dispatching (planning movements) for all the vehicles available at that node, they looked at a different node. In this manner, they progressed forward through a plan, making decisions locally at each node and each point in time.

In the task-based process, teams also progressed forward in time; however, the focal unit was not a node but a "task", which here we use to mean moving cargo out of an origin or along a path. For example, teams focused on moving cargo out of the congested port city, or moving cargo along a bottleneck road through the mountains. Before making planning decisions, teams roughly mapped out the need for cargo movement along these tasks, then focused their attention on one task or another, moving vehicles to accomplish that task until movements for all vehicles were planned. Then they moved on to the next time that vehicles became available. Occasionally they re-mapped the need for cargo movement among tasks, or formulated new tasks to address needs. As in the location-based decision process, they stepped through the plan by moving forward in time, but they made decisions based on the need for movement along tasks (from origins or along paths) rather than at each node individually.

A third important element of these two decision processes was how each individual dispatch decision was made. As described in Chapter 2, teams made each decision based on a series of rules or policies describing how to prioritize cargo based on its item type, its destination, or other characteristics. For example, some teams decided to send mostly priority-1 shelter cargo, but also some priority-2 health, priority-3 water and sanitation, and a small amount of priority-4 food. Some teams decided to serve the closest helicopter destinations first, to maximize their capacity. The policies were based on goals formulated by the teams, such as delivering more high-priority cargo, or using helicopters efficiently.

Clearly, the location-based and task-based decision processes are not algorithms, because there were many variations in how they were carried out. However, they do have an algorithmic character. They show people stepping through the plan over time, making the same kinds of decisions over and over. The processes resemble greedy search heuristics, and the policies resemble rules that guide search in the direction of goals. The policies, therefore, probably have a strong impact on the quality of the resulting solution (the transportation plan), but it was impossible to determine which policies are better or worse based on the ethnographic study. To understand the effectiveness of the human decision processes and policies, this section develops models of the decision processes, which can be evaluated against one another and against solutions from the optimization model described in the previous section.

Models cannot replicate the exact processes employed by humans. Instead, they represent an idealized version of those processes, so we can evaluate how effective they would be if implemented consistently. That is the goal of this section. We develop models of the two archetypal decision processes, location-based and task-based, documenting the assumptions necessary to build a heuristic model of each. In addition, we develop variations around these two processes that implement various policies that governed each decision in the process. Table 4.1 shows all the modeled decision processes and policies. It also includes a summary of what was observed in the ethnographic study, and the additional assumptions required to model the process or policy. The following paragraphs describe the models in more detail, and pseudocode is provided in Appendix B.

4.4.3 Location-based decision process

To model the location-based decision process, described above and in Table 4.1, a heuristic algorithm replicates the teams' pattern of stepping forward in time, and focusing on each location in turn. The algorithm moves through the nodes one by one, and at each node, finds the available vehicles and waiting cargo. For each vehicle, it sets a destination for the vehicle, as determined by the cargo selection policy, and loads cargo onto the vehicle, again determined by the cargo selection policy. If there are vehicles but no waiting cargo, the vehicles are moved toward the node with the most waiting cargo. A distinguishing feature of this algorithm is that vehicle movements are only planned one step ahead, meaning movements go only to the adjacent nodes. This represents the behavior of teams in planning movements forward only one step at a time. The algorithm's steps are summarized below. (When a step is governed by a policy, the policy is noted in parentheses; see Table 4.1 and the following sections for definitions of these policies.)

- For each CMR, find a shortest or cheapest path to its destination (determined by airbridge policy)
- At each time step, for each node,
- For each vehicle available at the selected node,
  - Select an adjacent node as the destination for the vehicle: the node with the most cargo waiting to go through it (of type indicated by the cargo selection policy)
  - Fill the vehicle with CMRs going through the selected destination (in order determined by cargo selection policy)
- If there is no waiting cargo, move the vehicle toward the node with the most waiting

Process or Policy	Description, based on observations of teams	Summary of assumptions made in the model	Decisions affected in location-based process	Decisions affected in task-based process
Decision Processes				
Location-based	Teams stepped through the plan over time. They focused on one location at a time, choosing from cargo at that location to load onto vehicles at that location. The selection of cargo determined the destination of the vehicle. After planning one location, teams focused on another location, or the next time at which vehicles became available. Many teams assigned paths to cargo before they planned their movements.	The model assumes the plan is tracked adequately. It assumes the teams move through time sequentially, planning movements for vehicles as soon as they become available. When a truck has no cargo to move, it is moved towards the most waiting cargo. Movements are only planned one step ahead, meaning to an adjacent node. Paths for each cargo movement request are determined at the beginning, and follow the least-expensive path or the shortest path (determined by a policy).		
Task-based	Teams stepped thorugh the plan over time. They focused on one task at a time, meaning an origin or a path. They roughly understood the need for cargo movement in each task, and focused on those that had more urgent needs. For a focus task, they routed vehicles and loaded cargo. After planning one task, they looked at another task, or the next time at which vehicles became available. Many teams assigned paths to cargo before they planned their movements.	The model assumes the plan is tracked adequately. It assumes the teams move through time sequentially, and plan movemetns for vehicles when they become available. The model calculates the need for cargo movement on a defined set of tasks at every new time period. Tasks are selected based on the calculated need (according to a policy). Vehicles are allocated to tasks either proportionally to need or to the neediest first. The model focuses on one task, then another, until all available vehicles have been planned, then moves to the next time when vehicles are available.		
Cargo selection policies				
Prioritize by item type, lexicographic	Teams selected cargo to load by looking at the priority of the type of item. They loaded all priority-1 cargo before any priority-2 cargo, all priority-2 cargo before any priority-3 cargo, etc, within a given location or task.	Given a choice of destinations, vehicles move to the destination to which the most high-priority cargo is going. In selecting cargo to load, higher-priority is loaded first.	Selecting destination for vehicle and cargo to load	Selecting destination for vehicle (in origin tasks) and cargo to load
Prioritize by item type, weighted mix	Teams selected cargo to load by looking at the priority of the type of item. They attempted to load more high-priority items, but some of the lower-priority items as well.	Given a choice of destinations, vehicles move to the destination to which the most high-priority cargo is going. In selecting cargo to load, the model attempts to send 50% priority-1, 30% priority-2, 15% priority-3, and 5% priority-4.	Selecting destination for vehicle and cargo to load	Selecting destination for vehicle (in origin tasks) and cargo to load
Prioritize by destination, lexicographic	Teams considered the priority of destinations, usually by province, two of which were considered high-priority and two others considered lower-priority. They loaded cargo for high-priority destinations before that for lower-priority destinations.	Given a choice of destinations, vehicles move to the destination to which the most high-priority cargo is going. In selecting cargo to load, cargo going to high-priority destinations is loaded first.	Selecting destination for vehicle and cargo to load	Selecting destination for vehicle (in origin tasks) and cargo to load
Prioritize by destination, weighted mix	Teams considered the priority of destinations, usually by province, two of which were considered high-priority and two others considered lower-priority. They loaded more cargo for high-priority destinations, but also sent some cargo to lower-priority destinations.	Given a choice of destinations, vehicles move to the destination to which the most high-priority cargo is going. In selecting cargo to load, the model attempts to send 70% cargo going to high-priority destinations, and 30% cargo going to low-priority destinations.	Selecting destination for vehicle and cargo to load	Selecting destination for vehicle (in origin tasks) and cargo to load
Prioritize cargo with shortest path to destination	Teams often mentioned the goal of delivering as much and as quickly as possible, sometimes manifested as delivering what's there or what will get to its destination most quickly.	The model implements this by prioritizing cargo with the shortest remaining distance to reach its destination. In selecting cargo to load, cargo with the shortest paths are loaded first. Given a choice of destinations, the destination is determined by the path of the cargo with the shortest remaining distance to its destination.	Selecting destination for vehicle and cargo to load	Selecting destination for vehicle (in origin tasks) and cargo to load
Prioritize cargo headed to a particular "bottleneck" node	Teams selected shipments that were needed in order to fill onward transport: helicopters and small trucks. In choosing cargo or tasks to focus on, they chose those that had to go by 10T or helicopter transport.	The model requires that "bottleneck" nodes, to which cargo should be moved for onward transport, are defined ahead of time.Given a choice of destinations, vehicles move to the destination to which the most "bottleneck" cargo is going. In selecting cargo to load, "bottleneck" cargo is selected until the required amount has been sent to the bottleneck each day.	Selecting destination for vehicle and cargo to load	Selecting destination for vehicle (in origin tasks) and cargo to load
Task selection policies				
Select task with most total need	Teams described tasks in terms of the amount of cargo that needed to move along it.	Tasks are selected in order of the most total cargo that needs to move through the task.		Selecting the next task on which to focus
Select task with most need for high-priority item types	Teams described tasks in terms of the amount of high-priority types of cargo that needed to move along it.	Tasks are selected in order of the most high-priority cargo that needs to move through the task.		Selecting the next task on which to focus
Select task with most need for high-priority destinations	Teams described tasks in terms of the priority of reaching certain destinations.	Tasks are selected in order of the most cargo going to high-priority destinations that needs to move through the task.		Selecting the next task on which to focus
Task vehicle allocation policies				
Allocate vehicles to tasks in proportion to need	Teams often decided how many vehicles to allocate across different tasks by looking at how much cargo needed to be moved along each task.	The model allocates vehicles in proportion to the total need (either total, or based on priority) at all tasks for a given time period.		Determining the number of vehicles to allocate to each task
Allocate vehicles to selected task before others	Some teams worked on one task before looking at the need on any others, allocating vehicles to the first one without thinking about the others.	The model allocates enough vehicles to satisfy the total need at each selected task before any others.		Determining the number of vehicles to allocate to each task
Network policies				
Use an airbridge for early cargo, then move to less expensive modes	Most teams discussed whether to use an "airbridge", using helicopters to fly over the mountains in the first few days rather than slowly trucking goods the long way around the mountains.	The model implements this idea by assigning a certain amount of cargo to move along the shortest path to its destination, rather than the least-expensive path. The shortest path employs the helicopter airbridge, while the longer, cheaper path employs trucks. The amount of cargo to be assigned a shortest path, and the type of cargo (e.g. high item priority or destination priority), is defined ahead of time.	Assigning paths to cargo movement requests	Assigning paths to cargo movement requests

Table 4.1: Decision processes and guiding policies

cargo (of type indicated by the cargo selection policy)

4.4.4 Task-based decision process

To model the task-based decision process, described above and in Table 4.1, a heuristic algorithm replicates the teams' pattern of stepping forward in time, and focusing in turn on various tasks. The tasks are not defined by the algorithm but rather must be defined individually for each problem instance. The model is implemented this way because teams did not have a systematic way of defining tasks; rather, task definition was part of the sensemaking process, which is difficult to describe algorithmically (this is left to future research). In these models, tasks are defined for each test problem based on my observations of the teams and the tasks they defined. Tasks are defined as either moving cargo from an origin or along a path, and a type of vehicle is associated with each task (e.g. small trucks, helicopters, etc.) The algorithm computes the need for cargo movement along all tasks at each time step, then selects a task according to the task selection policy, and allocates vehicles to it according to the vehicle allocation policy. For each allocated vehicle, it routes the vehicle to the task's origin. If the task is to move cargo along a path, the vehicle's destination is routed to move along that path, but if the task is to move cargo from an origin, the vehicle's destination is set according to the cargo selection policy. The vehicle is filled with cargo, again according to the cargo selection policy. After all allocated vehicles have been planned, another task is selected, until all vehicles have been planned. The algorithm's steps are summarized below.

- For each CMR, find a shortest or cheapest path to its destination (determined by airbridge policy)
- At each time step,
- Select a task: the task with the most need for cargo movement (of type indicated by the task selection policy)
- Allocate vehicles to the task from all vehicles available at this time, either proportionally or to satisfy this task's need (determined by the vehicle-to-task allocation policy)
- For each vehicle allocated to the task,
  - Route the vehicle to the task's origin

  - If the task is a path, route the vehicle along the path; if the task is an origin, select a node as the destination: the node with the most cargo waiting to go to it (of type indicated by the cargo selection policy)
  - Fill the vehicle with CMRs going through the selected destination (in order determined by cargo selection policy)

- If there are still vehicles available at this time, select another task

4.4.5 Cargo selection policies

Cargo selection policies govern two major decisions in both the location-based and task-based processes: (1) the selection of a destination for each vehicle and (2) the selection of cargo to load on each vehicle. In general, these policies assign a priority level to each set of cargo, and guide decisions such that more cargo of this priority level is delivered. The destination is chosen as the node to which the most high-priority cargo needs to go, and cargo is loaded in priority order (though there are variations in how this is implemented). Each policy defines the priority of cargo loads differently, either based on the type of cargo, its destination, the distance to its destination, or whether it must pass through a bottleneck in the network. Another variation is in whether cargo is loaded strictly in priority order, or whether an attempt is made to send a "weighted mix", with more high-priority cargo but some lower-priority cargo. Each policy is described briefly below, and pseudocode is given in Appendix B (note that the policy-related pseudocode is embedded within the routines for the location-based and task-based processes).

Prioritization by item type The item type policies prioritize cargo according to the type of item. In this scenario, and in practice, the priority of each item type is defined exogenously (e.g. by the host government of the response or the United Nations Office for the Coordination of Humanitarian Affairs). In the Snowland scenario and the others in this thesis, shelter is first-priority, health goods second, water and sanitation equipment third, and food fourth. In determining the destination of a vehicle, the item type prioritization policies select the node with the largest amount of highest-priority cargo waiting to move through it. In loading cargo onto vehicles, there are two variations. In the lexicographic variation, cargo is loaded strictly in priority order; within a given priority level, the cargo is loaded in the order of the list of requests (i.e. in the order requests were received). In the weighted mix variation, the model attempts to allocate, on each truck, the following amount of capacity: 50% shelter, 30% health, 15% water and sanitation, and 5% food (many requests may be split across multiple trucks). (The "amount" of cargo can be defined in various ways, such as the number of items, the weight, or the volume. The teams in the training most often used weight, so weight is used here.)

Prioritization by destination The destination policies prioritize cargo according to the urgency of need at its destination. In practice, these priority levels may or may not be clearly defined (again, for example, by the host government). In the Snowland scenario, location priorities were not defined or sorted into classes, but nearly all the teams considered two provinces high-priority and two provinces low-priority. This model assigns each load of cargo to be high-priority or low-priority depending on its final destination. In determining the destination of a vehicle, there are two variations of the destination prioritization policy. The lexicographic policy simply selects the destination node with the largest amount of highest-priority (by destination, not item type) cargo waiting to move through it. The weighted mix policy selects the first 70% of the vehicles at this node (by capacity, rounded up to the next whole vehicle), and sets its destinations based on the amount of high-priority cargo. The remaining 30% of vehicles at this node (by capacity, rounded down to the next whole vehicle) choose destinations based on the largest amount of low-priority cargo. In loading cargo onto vehicles, the model attempts to load high-priority cargo onto those vehicles sent to high-priority destinations, and low-priority cargo on those vehicles sent to low-priority destinations.

Prioritization by fastest delivery The fastest-delivery policy prioritizes cargo based on how quickly it can be delivered. Instead of assigning a priority level to each load of cargo, the loads are sorted based on the distance from their current locations to their final destinations. In determining the destination of a vehicle, it is set to deliver the first load on the list, the one with the shortest distance to final destination. In loading cargo onto the vehicle, cargo is loaded in order from the list, so that cargo with the shortest remaining distance is loaded first.

Feed bottleneck nodes This policy assigns a priority-level to cargo based on whether it must pass through a pre-defined "bottleneck" node, and the priority-level of that bottleneck node. The policy represents an important dynamic in the observed teams' decision-making: feeding cargo to a helicopter base in order to ensure that the helicopters would never be idle waiting for cargo. The model requires that the bottleneck node(s), such as the helicopter base in the scenario, be defined ahead of time; if multiple nodes are defined, each has a priority level. Cargo is then assigned a priority based on the bottleneck node through which it must pass; cargo that does not go through a bottleneck node is assigned the lowest priority. The policy prioritizes cargo in order of its assigned bottleneck priority, until an assigned amount has been sent toward that node for each day (e.g. the daily capacity of helicopters from the helicopter base). In determining vehicle destinations and in loading cargo, the model attempts to send the assigned amount toward each bottleneck node, progressing through them in priority order.

4.4.6 Task selection and vehicle-to-task allocation policies

Task selection policies are used only in the task-based decision process. They govern which task is selected, from among all tasks at each time step. Tasks are always selected based on the "need" for movement of cargo along a task. There are two types of tasks: origin tasks and path tasks. For origin tasks, the need for cargo movement is simply the amount of cargo waiting at the origin node, at the current time. For path tasks, the need for cargo movement is the amount of cargo waiting at the path's origin with next steps along the path, at the current time. In the total need policy, all loads are counted in the calculation of needs. In the item type policy, only loads of the highest-priority item types (remaining) are counted, while in the destination policy, only loads going to highest-priority destinations are counted. As a result, each of these policies selects tasks in order of the need along each task, where need is either in total, or only of the highest-priority loads.

Vehicle-to-task allocation policies are also used only in the task-based decision process. They govern the way in which vehicles are allocated among tasks at a given time. In the proportional policy, vehicles are allocated proportionally to the need for each task at this time (again, need is calculated according to the task selection policy), but rounding up to whole vehicles favors the earlier-selected tasks over the later. The other vehicle allocation policy is first-come-first-served, in which vehicles are allocated to the first-selected task until all the need is satisfied, then to any other tasks if additional vehicles are available.

4.4.7 Network policies

The airbridge policy governs how CMRs are assigned paths at the start of both the location-based and task-based algorithm. It represents an important behavior observed in the teams: the decision of whether or not to use helicopters to move cargo quickly into remote locations rather than sending it the long way around by truck. The trade-off they discussed was the expense of using helicopters compared to the speed of delivery; their default mode was clearly to send cargo by truck wherever possible. To model this dynamic, the airbridge policy specifies a certain amount of cargo to be assigned a shortest path to its destination, while all remaining cargo is assigned a least-expensive path to its destination. At the start of the algorithm, the complete list of CMRs is sorted in order of priority, according to the cargo selection policy (meaning it is sorted by item type priority, or destination priority, etc.). Starting at the top of the list, CMRs are assigned shortest paths until the total amount of cargo assigned reaches the amount specified in the policy; all remaining CMRs are assigned least-expensive paths.

4.5 Assessing the performance of human approaches

To assess the performance of human approaches, we compare the models of human approaches to each other and to solutions from the mixed-integer program developed earlier. Solutions can be compared in terms of various metrics. Here, the key metric is the utility of the solution (transportation plan), based on the utility function developed in Chapter 3. This utility function is the objective in the optimization model (see Section 4.3.2), and utilities are computed for the solutions to the heuristic models as well. Solutions are also compared in terms of the amount of cargo delivered (again, cargo weight is used for the amount of cargo, following the teams' behavior), broken down by item type and destination priority. These are the key metrics by which teams of humanitarian logisticians evaluated transportation plans, and they are adopted here to assess the performance of plan-generating heuristics.

The effectiveness of heuristics may be influenced by the specific structure of the problem. A two-part strategy is employed to evaluate performance on various types of problems. First, heuristics are evaluated on two small stylized problems, each with several variations. The goal is to gain intuition about the effectiveness of each heuristic on specific types of problems, and how performance changes as key problem dimensions are varied. These stylized problems are based on important characteristics of the Snowland network, and each problem is designed to probe the effectiveness of the heuristic in managing problems with that particular characteristic. We use stylized problems based on real problem characteristics because our goal is to test performance on problems relevant to disaster response. Second, the heuristics are evaluated on the full Snowland problem solved by the teams in the training exercise, to test the performance of the heuristics on realistically scaled problems. The problem is too large to be solved to optimality by the mixed-integer model, so some simplifications are made to obtain an approximate optimal solution for comparison with the heuristic solutions. The following sections describe the test problems in more detail.

4.5.1 Stylized problems

The stylized problems are intended to be simple, in order to probe the effectiveness of heuristics under specific problem situations and to make it easier to understand the advantages and drawbacks of each heuristic. Each problem includes a port node P, and all the inventory is initially located there. There are two final destination nodes, each with a different destination priority. There are 8 cargo movement requests, 4 to each destination node, each of which is of a different item type priority. There are two types of vehicles, trucks and helicopters, each with a fixed volume and weight capacity. In each case, 4-day plans are created.

A key challenge, in practice and in modeling, is to determine the right metric by which to measure the amount of cargo delivered. Should we seek to maximize the weight delivered, or the volume, or the number of items? In reality, the goal is to maximize the number of people to which assistance is provided, but there are no estimates of the number of people assisted per pound or cubic meter of cargo. (One might estimate this for certain items, but doing so across item categories, such as shelter, would be more difficult, and is left to future research.) The teams in the simulation training spoke primarily of maximizing weight, so we follow their lead here. However, in these stylized problems, it is useful to think of the weight as corresponding to the number of people helped. The volume of each type of cargo is varied in some instances, to change the difficulty of delivering priority cargo, but the goal remains to deliver the most weight, or to help the most people.

"Feed Onward Transport" This problem models a system in which cargo must be "fed" to a node for onward transport by some other mode of transportation. In the Snowland scenario, many teams worried about sending cargo by road to the helicopter base for onward transport by helicopter. This problem is a simple representation of that challenge. The stylized problem is described in Figure 4-2. Cargo is at node P initially, and must be transported to either E or F. F is easily reached by truck, but E can only be reached by moving cargo to D by truck, then by helicopter to E.

One variation on this problem looked at high- and low-capacity scenarios, with either four trucks and four helicopters or two trucks and two helicopters. A second variation changed the difficulty of delivering priority cargo. In the easy variation, the easily reached node F was a high-priority destination, and all cargo types had the same volume. In the difficult variation, the hard-to-reach node E was the high-priority location. Moreover, the higher-priority item types had larger volume, making it harder to deliver the same amount (weight) of cargo.

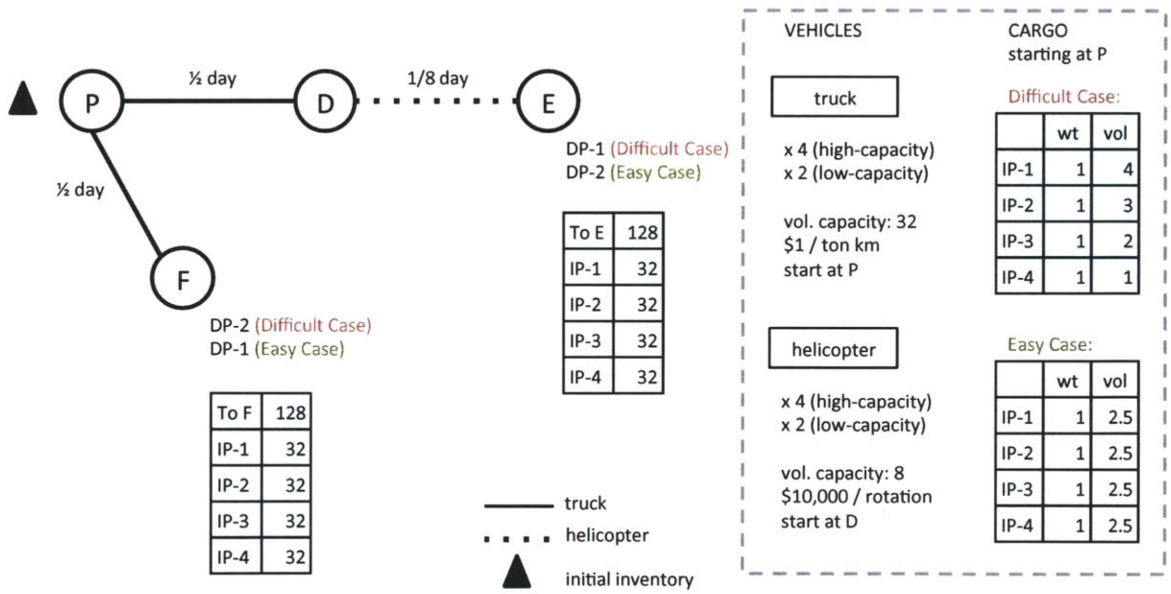


Figure 4-2: The "Feed Onward Transport" instance: deliver cargo for onward transport by helicopter.

"Airbridge" This problem models the trade-off between using helicopters to reach otherwise inaccessible areas or to provide speedy deliveries to areas that can be reached by road, but more slowly. In the Snowland scenario, many teams considered using helicopters to provide such an airbridge in the first few days, before cargo could reach the mountainous areas by road. This problem, described in Figure 4-3, is a simple representation of that challenge. Again, cargo is initially at node P, and must be moved to E and F. F is only accessible by helicopter, while E is accessible by road, but trucks require an entire day to reach it, while helicopters can deliver much more quickly. The same high- and low-capacity scenarios, and easy and difficult prioritization cases, can be created for this problem.

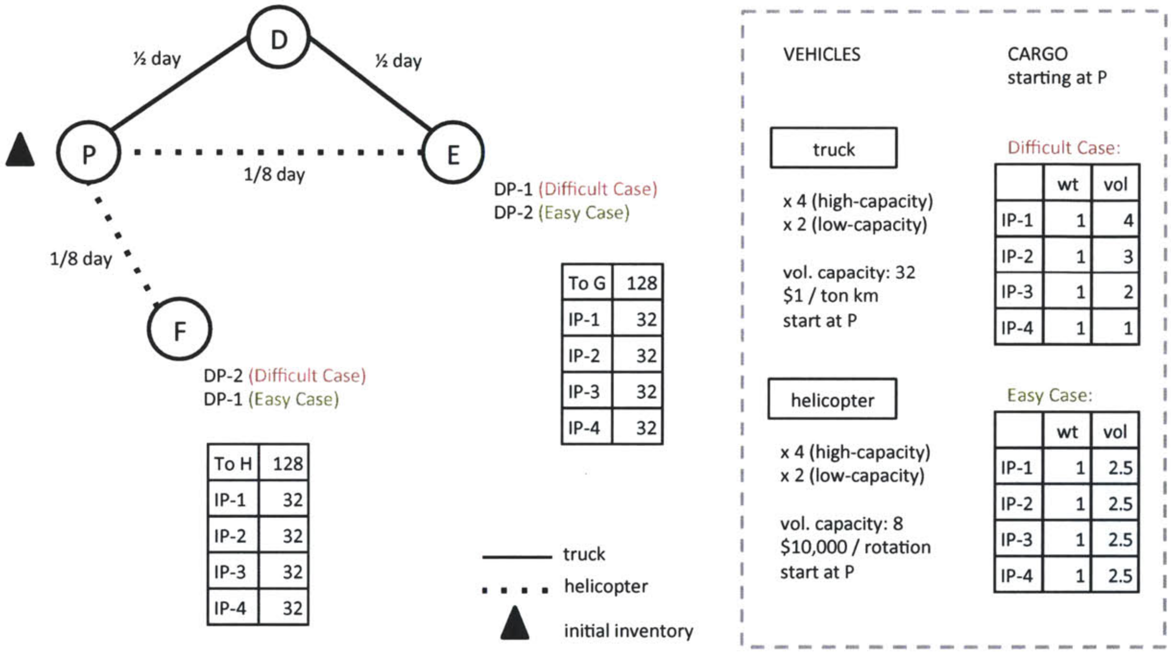


Figure 4-3: The "Airbridge" instance: use helicopters as an airbridge or to deliver to other locations

4.5.2 Realistic problem: Snowland scenario

While the stylized instances are useful for highlighting specific problem characteristics, it is important to test the effectiveness of heuristics on a problem of realistic size and complexity. A natural choice here is the Snowland scenario from which these heuristics were developed. In a sense, using this Snowland problem gives the heuristics the "best shot" at showing their effectiveness, because the heuristics were derived from human behavior in solving exactly this problem. The claim that the Snowland scenario is "realistic" requires some justification. It was developed by a group of very experienced humanitarian logisticians, many with 10-20 years of field experience. They intended to depict the real complexities of operations they had experienced, in order to challenge the training participants with a realistic scenario. In particular, it was developed to be somewhat similar to the Pakistan earthquake in 2005. The Snowland scenario includes many realistic complexities and challenges, which are described below.

"Snowland" The Snowland instance is the same problem solved by the teams in the LRT simulation. Figure 4-4 summarizes the problem. The road network contains good roads accessible to large (40-ton) trucks, bad roads accessible to small (10-ton) trucks, and areas accessible only by helicopter. Cargo enters mainly through the hubs P, H, and A, with small starting inventories at B and D; it must be delivered to hubs and small destinations throughout the network. The amount of cargo destined for each destination is shown, along with its breakdown by item type priority. The double-walled nodes in Figure 4-4 actually include a number of different destinations, but they are depicted as one node in the diagram for clarity. Some of the destinations are priority 1, and some are priority 2. There are five types of vehicles: 40-ton trucks, 10-ton trucks, and three types of helicopters with slightly different weight capacities. Helicopters have "infinite" volume capacity because loads can be slung beneath them. The vehicles start in different locations, and many do not arrive until several days into the planning period.

The fundamental structure of the problem was the same for all 10 teams in the LRT simulation, but the specific number of requests and number of vehicles available were changed occasionally over the course of the 5 trainings. In the earlier trainings, there were fewer vehicles, and in later trainings, there were more vehicles, so that a greater percentage of the cargo could be delivered. Here, we use two versions of the Snowland formulation, corresponding to these two cases. In both cases, there is plenty of 40-ton truck capacity, but the helicopter and 10-ton capacity is very limited.

This Snowland scenario contains many complexities and challenges. Most cargo must be moved by large truck, then by small truck, then by helicopter. The majority of the cargo can only be delivered by helicopter, and helicopters are in short supply. Cargo enters the country through two different entry points. Some areas are particularly remote, while others are easily accessible. Many of the cargo movement requests are too large to be transported in one load, and must be split across multiple vehicles and sent at different times. There are trade-offs between using helicopters only for deliveries to inaccessible areas, or allocating some to provide an airbridge to get cargo quickly from P, the largest entry point through C to the neediest areas around B. It is difficult to reach the helicopter base at B by road, so large quantities of cargo must be transported through the bad road from A to B to ensure that the helicopters are never idle waiting for cargo. According to the humanitarian logisticians who have designed, facilitated, and participated in the LRT training, these challenges and complexities make for a realistic transportation planning problem.

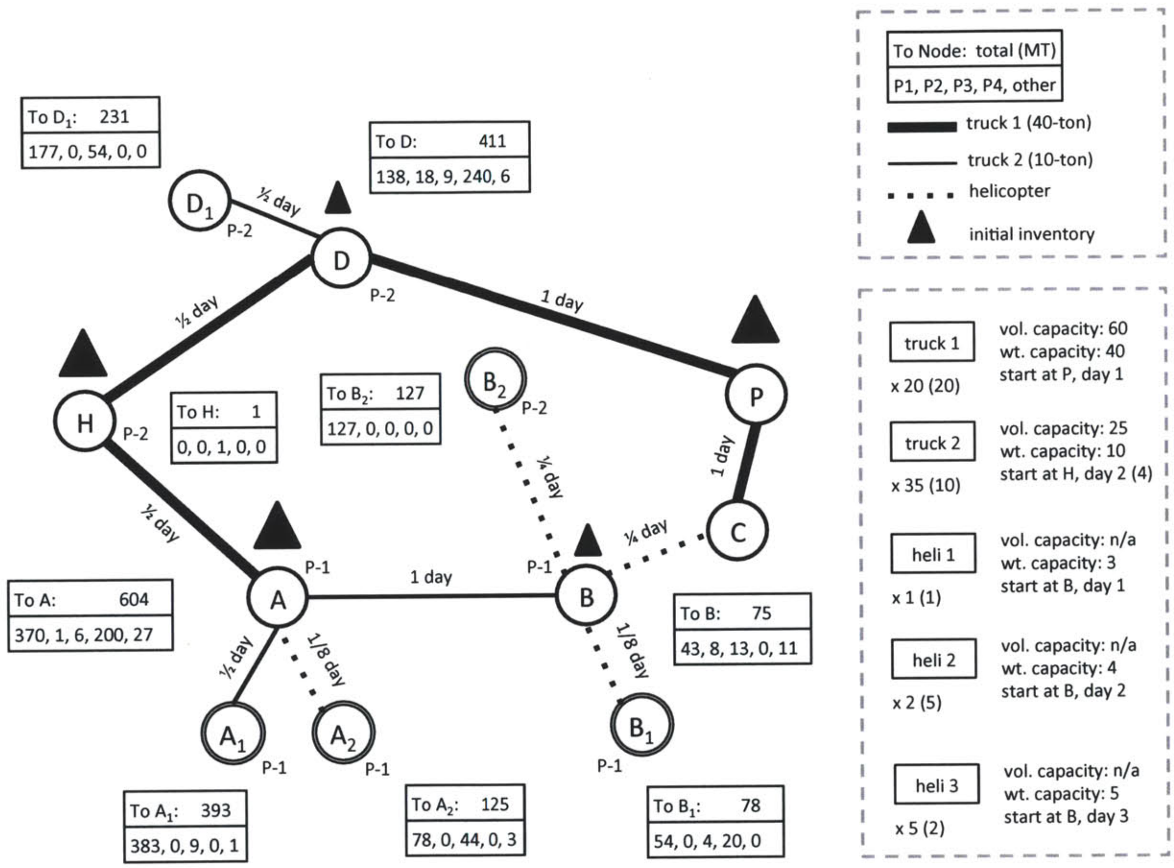


Figure 4-4: The "Snowland" problem: Cargo enters mainly through P, H, and A, to be transported to hubs and to destinations in the interior. There are two versions of this instance, with different numbers of vehicles; the mid-capacity version is shown in parentheses.

"Snowland Simplified" The Snowland problem described above leads to very large problems. Because it models every movement of every helicopter, a relatively large number of time periods per day are required to ensure each helicopter can make several flights per day. Therefore, it is difficult to solve problems with more than two-day time horizons, or 16 periods. However, cargo in this scenario can take more than three days to reach its final destination, so it was important to study problems with longer time horizons. Therefore, we developed a simplified version of the Snowland scenario that can be solved by an optimizer for longer time horizons.

With a few simplifications, we develop a formulation that approximates the optimal solution but enables larger problems to be solved. The key change is to reduce the number of time periods. In the Snowland scenario, and in many realistic scenarios, road trips take no less than half a day, so road vehicles alone should require no more than two periods per day. For helicopters, we simply lump all the flights (per day) into one flight, and multiply their capacity by the number of flights they can make per day. For example, if a helicopter can make four rotations per day, we multiply its capacity by four and modify its flight times so that it can make only one flight per day. (Costs are also updated accordingly.) In this manner, we can reduce the number of periods per day to two.

With this change, however, the helicopters can only reach one location rather than four different locations in a day. Therefore, we aggregate the demand to one node that stands in for an entire set of villages. The set can be defined by a region (e.g. a province in Snowland, such as Rocky), or by villages of a certain priority in a region. The resulting instance provides a conservative approximation of the optimal solution to the original problem, because it is easier to deliver more cargo when villages are aggregated into single nodes. The simplified problem instance is shown in Figure 4-5.

Testing on 2-day instances suggests that this formulation provides a reasonable, conservative approximation to the optimal solution. The original problem yielded an optimal solution (maximizing total deliveries) of 1174 MT, while the reformulated problem yielded a solution of 1185 MT.

4.5.3 Test Cases

In order to understand the performance of heuristics under different problem conditions, we test several cases of each stylized problem. As described earlier, the capacity (the number of vehicles) and the difficulty of delivering priority cargo may be varied. A third dimension of variation is whether an airbridge is used. The optimizer always has the option of employing an airbridge, but in the heuristic models, an airbridge must be explicitly included. Table 4.2 shows the combinations of these variations tested for each problem. In the "Feed Onward Transport" problem, the difficulty and capacity were varied, but the airbridge option was

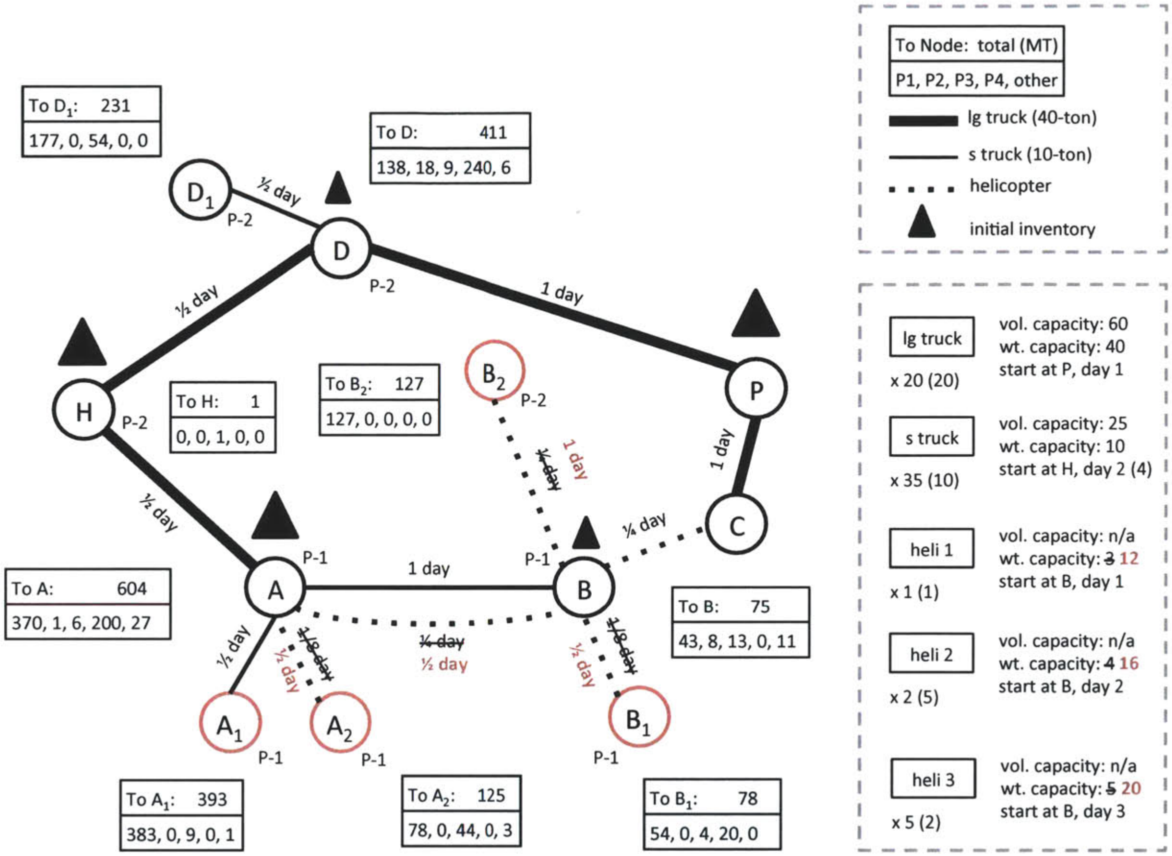


Figure 4-5: The "Snowland simplified" problem: Compared with the original, this problem has a reduced number of time periods and aggregated destination nodes. To reduce the number of time periods arc lengths were increased and vehicle capacities changed to compensate. Changes are highlighted in red.

not relevant. In the "Airbridge" problem, the high-capacity scenario was uninteresting, so the other two dimensions were varied while the capacity remained low. In the "Snowland" problem, the difficulty was irrelevant because the requests were defined in the scenario. The two capacity levels correspond to the vehicles provided to teams in the first few trainings (low capacity) and in the last few trainings (high capacity). Again, the airbridge is explicitly included for the heuristic models in two of the four cases.

A second element of the test cases is the set of specific heuristics and policies tested in each case. For simplicity, all problem instances were run with the same set of heuristics and policies. These sixteen combinations of decision process heuristics with various types of policies are provided in Table 4.3. The location-based decision process is run with each of the six cargo selection policies. Ten combinations of policies are used with the task-based

Difficulty of delivering priority cargo	Capacity (number of vehicles)
Easy	Low
Easy	High
Difficult	Low
Difficult	High

(a) "Feed Onward Transport"

Difficulty of delivering priority cargo	Airbridge in heuristic	Capacity (number of vehicles)
Easy	No	Low
Easy	Yes	Low
Difficult	No	Low
Difficult	Yes	Low

(b) "Airbridge" Problem

Capacity (number of vehicles)	Airbridge in heuristic
Low	No
Low	Yes
High	No
High	Yes

(c) "Snowland" Problem

Table 4.2: Problems and cases of each problem

decision process. Run 7 is a baseline set of item prioritization policies, and runs 8, 9, and 10 each vary one policy at a time from this baseline. Similarly, run 11 is a baseline set of destination prioritization policies, and runs 12, 13, and 14 each vary one policy at a time from this baseline. Finally, runs 15 and 16 represent "crosses", combinations of item prioritization and destination prioritization policies.

Run #	Decision Process	Cargo Selection Policy	Task Selection Policy	Vehicle Allocation Policy
1	Location-based	Item priority, lexicographic		
2	Location-based	Item priority, weighted mix		
3	Location-based	Destination priority, lexicographic		
4	Location-based	Destination priority, weighted mix		
5	Location-based	Shortest path to destination		
6	Location-based	Feed to bottleneck		
7	Task-based	Item priority, lexicographic	Most need for high-priority item types	Proportional
8	Task-based	Item priority, weighted mix	Most need for high-priority item types	Proportional
9	Task-based	Item priority, lexicographic	Most total need	Proportional
10	Task-based	Item priority, lexicographic	Most need for high-priority item types	First come first served
11	Task-based	Destination priority, lexicographic	Most need for high-priority destinations	Proportional
12	Task-based	Destination priority, weighted mix	Most need for high-priority destinations	Proportional
13	Task-based	Destination priority, lexicographic	Most total need	Proportional
14	Task-based	Destination priority, lexicographic	Most need for high-priority destinations	First come first served
15	Task-based	Item priority, lexicographic	Most need for high-priority destinations	Proportional
16	Task-based	Destination priority, lexicographic	Most need for high-priority item types	Proportional

Table 4.3: Combinations of decision process heuristics and policies run for each problem instance

4.6 Results and Discussion

In this section, we analyze the results for the test cases described above on all three problems. We first examine each case and each problem in detail, then discuss results across cases and across problems.

4.6.1 "Feed Onward Transport" Stylized Problem

A description of the "Feed Onward Transport" problem was given in Figure 4-2. We study four cases, shown in Table 4.2. Results are reported for all four cases and discussed below. For each case, sixteen combinations of heuristics and policies are tested; these sixteen "runs" are described in Table 4.3. There is one important point to note. The policies may specify that cargo be loaded in order of priority by type, for example, but they do not specify the order in which to load cargo of the same item type. In such cases, cargo is loaded in the order listed in the database, which might be the order in which requests were received. In this problem, cargo requests were listed in the database in the following order: all cargo for F was listed before cargo for E, and lower-priority item types were listed before higher-priority item types. For example, priority-4 items for F were first, followed by priority-3 items for F, etc. This order was chosen in order to emphasize the potential weaknesses in heuristics that do *not* pay attention to each kind of priority, because if item priority is not considered, the lowest-priority items are loaded first.

Low-Capacity, Easy Prioritization Case Figure 4-6 summarizes the results for this case. Recall that the low-capacity case has only two trucks and two helicopters, and requests delivery of equal amounts of all four types of items to each of two destination nodes. The closer destination node is high-priority, and all item types have the same volume (see Figure 4-2 for details). With these parameters, the maximum possible amount of cargo to deliver is 102.4 MT.

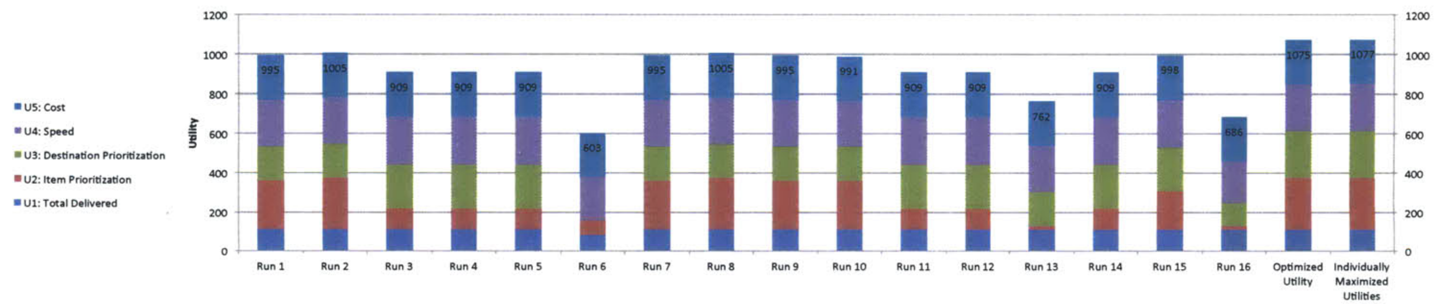
The first graph in Figure 4-6 shows the total utility of each "run" of the heuristic model (see Table 4.3 for descriptions of the specific policies in each run), along with the utility of the optimizer's solution (maximizing the sum of all component utilities) and the utilities obtained from the optimizer when maximizing each component of the utility function individually. The individual components of utility are highlighted in different colors within the stacked bar corresponding to each run, and the total utility value is shown. The next two graphs show the total metric tons delivered, broken down by item type priority and by destination type priority, respectively, and the final graph shows the average day of delivery. The table provides, for each run, the values for metric tons delivered in total and broken down by item type and destination priority. The table also provides the gap (difference) between the utility component value and the maximum value of that utility component, obtained by an optimizer maximizing that component individually. The gap in total utility is the gap between the total utility of the run and the utility obtained by an optimizer maximizing the sum of all five utility components. (These utilities have been scaled so that all utilities are positive numbers.) The top lines of the table indicate the policies used in each run, but refer to Table 4.3 for details.

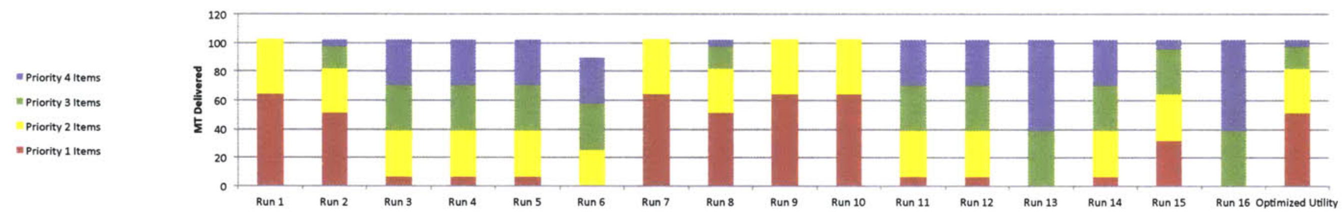
We first consider the optimized utility: it delivers an ideal mix of item types to an ideal mix of locations, and delivers the maximum possible amount of cargo (102.4 MT in this case). The gap shows that deliveries are only a little slower than the maximum possible value, but all other components are as high as they can be.

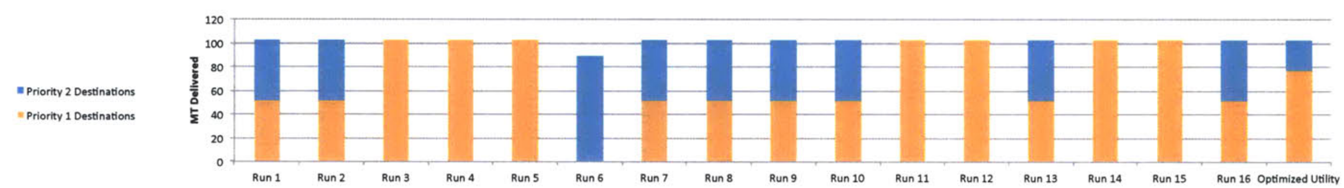
Next, we consider the location-based heuristics (runs 1-6). Nearly all policies deliver the maximum amount of cargo, so the difference in utility comes from the other components of the utility function. The best-performing policies are those that consider the item type priority, and the weighted mix policy is best of these, delivering the ideal mix of item types. The destination policies deliver only to priority-1 destinations, even the weighted mix policy, in this case because there are only two trucks: in each time period, the algorithm tries to allocate 70% of capacity to high-priority destinations, rounding up to the nearest whole truck, so they always go to high-priority destinations. These destination policies perform worse than the item type policies because, in this case, the additional utility from paying attention to item type prioritization outweighs the additional utility from paying attention to destination prioritization. In this case, the shortest path policy is equivalent to the destination prioritization policies, because cargo with the shortest path also happens to be going to the high-priority location. The feed policy performs poorly, in this case, because it forces cargo to go to the helicopter base, which in this case is a lower-priority destination. Clearly, a policy of maximizing the usage of the helicopters may not always lead to the best plans, especially when the helicopters are serving low-priority destinations. (Given their expense, however, it is unlikely they would be used in low-priority locations in practice.)

The task-based heuristics are divided into three sub-sets: those that prioritize by item (runs 7-10), those that prioritize by destination (runs 11-14), and "cross" policies that combine item and destination prioritization decisions (again, refer to Table 4.3 for details on the policies used in each run). Recall that in the former two sub-sets, the first run may

		Location-based						Task-based										Opt
		Item		Destination		Other		Item Prioritization				Destination Prioritization				Crossed		
		Lexic	Wtd	Lexic	Wtd	Short	Feed	Base	Wtd	TotNeed	FCFS	Base	Wtd	TotNeed	FCFS			
Run #:		**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**Opt**
Gap:	U1: Total Delivered	0.0	0.0	0.0	0.0	0.0	28.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	U2: Item Prioritization	10.2	-	156.3	156.3	156.3	189.3	10.2	-	10.2	10.2	156.3	156.3	242.9	156.3	67.0	242.9	-
	U3: Destination Prioritization	65.0	65.0	12.7	12.7	12.7	234.0	65.0	65.0	65.0	65.0	12.7	12.7	65.0	12.7	12.7	118.1	-
	U4: Speed	6.9	6.9	-	-	-	21.9	6.9	6.9	6.9	11.3	-	-	6.9	-	-	29.8	2.5
	U: Sum of utilities	79.8	69.7	166.3	166.3	166.3	471.8	79.8	69.7	79.8	84.2	166.3	166.3	312.5	166.3	77.0	388.6	-
Metric Tons Delivered:	Total delivered	102.4	102.4	102.4	102.4	102.4	89.6	102.4	102.4	102.4	102.4	102.4	102.4	102.4	102.4	102.4	102.4	102.4
	Priority 1 Items	64	51.2	6.4	6.4	6.4	0	64	51.2	64	64	6.4	6.4	0	6.4	32	0	51.2
	Priority 2 Items	38.4	30.72	32	32	32	25.6	38.4	30.72	38.4	38.4	32	32	0	32	32	0	30.72
	Priority 3 Items	0	15.36	32	32	32	32	0	15.36	0	0	32	32	38.4	32	32	38.4	15.36
	Priority 4 Items	0	5.12	32	32	32	32	0	5.12	0	0	32	32	64	32	6.4	64	5.12
	Priority 1 Destinations	51.2	51.2	102.4	102.4	102.4	0	51.2	51.2	51.2	51.2	102.4	102.4	51.2	102.4	102.4	51.2	76.8
	Priority 2 Destinations	51.2	51.2	0	0	0	89.6	51.2	51.2	51.2	51.2	0	0	51.2	0	0	51.2	25.6







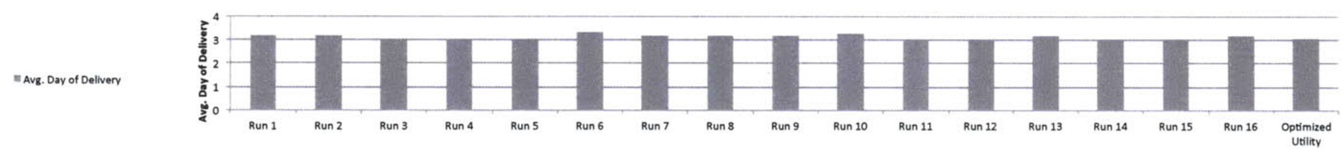


Figure 4-6: Results for "Feed Onward Transport" problem: low-capacity, easy case

be considered a "baseline", and the remaining three runs vary one factor at a time. As we might expect, the item prioritization policies perform well, with the weighted mix policy performing slightly better by delivering a better mix of item types (and in both cases, they are identical to their location-based counterparts). Selecting tasks based on the total need rather than the need for priority-1 items only (run 9) makes no difference in performance. The first-come-first-serve vehicle allocation policy makes deliveries slightly slower, probably because it allocates vehicles to the long-distance deliveries slightly more often. The task-based heuristics emphasizing destination prioritization perform slightly worse, overall, than the item prioritization policies, consistent with the results in the location-based heuristics. The only policy that makes a major difference is selecting tasks based on the most need overall rather than the most need for high-priority destinations: in this case, it delivers to both destinations because they have equivalent need, and delivers largely low-priority item types, resulting in much worse performance in utility. Of the two "cross" policies, the heuristic that selects policies based on item type and tasks based on destination performs well (run 15), because it pays attention to the destination priority in selecting the task to focus on, and to item priority in deciding which types of items to send. The opposite combination is not effective: it selects tasks based on item-priority needs, which are equivalent across both destinations, so it sends vehicles to both high- and low-priority destinations in equal proportions. In loading cargo, however, the destination is already set, so paying attention to the destination-priority in loading cargo makes no difference.

To summarize, across all runs, the best-performing policies are those that select cargo based on its item type priority, and attempt to send a weighted mix of cargo. The gain in utility from paying attention to item type (and not destination) outweighs the gain in utility from paying attention to destination (and not item type). We speculate that an even better policy would be a "cross" policy equivalent to that of run 15, except selecting cargo using a weighted mix item type prioritization policy; such a policy would pay attention to destination in deciding what task to focus on, then look at item type when loading cargo. While the use of different policies clearly affects the heuristic performance, there is little difference between the location-based and task-based processes. In this case, at least, the decision process has little effect on the performance.

Low-Capacity, Difficult Prioritization Case Figure 4-7 summarizes the results for this case. Recall that the "difficult prioritization" case is different from the easy case in that the high-priority node is the farther node, and the volume of higher-priority items is greater than that of lower-priority items, making them more difficult to deliver in large quantities. See Figure 4-2 for details. With these parameters, the maximum possible amount of cargo to deliver is 149.3 MT.

First, consider the optimized utility compared with the individually maximized utilities: none of the utility components are at their maximum possible values, suggesting that the optimizer made trade-offs between components in maximizing the overall utility. The gap in u_2, the item prioritization utility, is larger than the rest, suggesting this was one of the main sacrifices made in the optimal solution. In this "difficult prioritization" case, priority-1 items take up much more space in the vehicles (because they have larger volume), meaning that if many priority-1 items are delivered, the total weight delivered is much less overall. In this problem, it appears, the gain in u_2 from delivering a better mix of item types is outweighed by the gain in u_1 from delivering more weight overall. Indeed, the optimized solution delivers no priority-1 cargo at all. This begs a question: is this utility function truly representative of logisticians' preferences? It is unlikely that, when faced with this trade-off, planners would choose to deliver no priority-1 cargo. However, in this particular problem instance, it is much more difficult to deliver high-priority cargo, which is unlikely to be true in practice. Still, the case highlights a potential weakness in the utility function's performance in extreme cases.

Consider next the location-based heuristics (runs 1-6). The best-performing policies are those that deliver the most cargo overall. Again, because the item type prioritization policies deliver more high-priority items, which take up more space, less cargo can be delivered overall. The destination-prioritization policies perform better, because they do a better job prioritizing by destination and deliver fewer high-priority items (enabling more total deliveries). The shortest-path policy delivers a lot of cargo but only to the low-priority destination, resulting in a lower utility overall. The feed policy again performs relatively well, but only because it is equivalent to the destination policies.

Within the task-based heuristics, those that emphasize prioritization by item priority (runs 7-10) perform worse, again because they deliver less cargo overall. Those emphasizing destination priority (runs 11-14) perform better, again because they deliver more overall and

		Location-based						Task-based										Opt
		Item		Destination		Other		Item Prioritization				Destination Prioritization				Crossed		
		Lexic	Wtd	Lexic	Wtd	Short	Feed	Base	Wtd	TotNeed	FCFS	Base	Wtd	TotNeed	FCFS			
Run #:		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Opt
Gap:	U1: Total Delivered	190.3	133.2	101.1	101.1	83.3	101.1	190.3	133.2	190.3	190.3	101.1	101.1	0.0	101.1	190.3	0.0	19.0
	U2: Item Prioritization	-	24.2	70.2	70.2	51.0	70.2	-	24.2	-	-	70.2	70.2	104.6	70.2	30.6	104.6	80.0
	U3: Destination Prioritization	163.1	121.7	31.1	31.1	223.1	31.1	163.1	121.7	163.1	163.1	31.1	31.1	54.1	31.1	121.7	54.1	5.7
	U4: Speed	61.5	49.0	34.2	34.2	19.3	34.2	61.5	49.0	61.5	63.4	34.2	34.2	2.2	34.2	66.6	2.2	9.7
	U: Sum of utilities	300.3	213.5	122.3	122.3	261.6	122.3	300.3	213.5	300.3	302.1	122.3	122.3	46.2	122.3	294.9	46.2	-
Metric Tons Delivered:	Total delivered	64	89.6	104	104	112	104	64	89.6	64	64	104	104	149.3333	104	64	149.3333	140.83
	Priority 1 Items	64	32	8	8	16	8	64	32	64	64	8	8	0	8	32	0	0
	Priority 2 Items	0	25.6	32	32	32	32	0	25.6	0	0	32	32	21.33333	32	32	21.33333	38.46
	Priority 3 Items	0	19.2	32	32	32	32	0	19.2	0	0	32	32	64	32	0	64	38.37
	Priority 4 Items	0	12.8	32	32	32	32	0	12.8	0	0	32	32	64	32	0	64	64
	Priority 1 Destinations	32	44.8	104	104	0	104	32	44.8	32	32	104	104	74.66667	104	64	74.66667	96.04
	Priority 2 Destinations	32	44.8	0	0	112	0	32	44.8	32	32	0	0	74.66667	0	0	74.66667	44.79

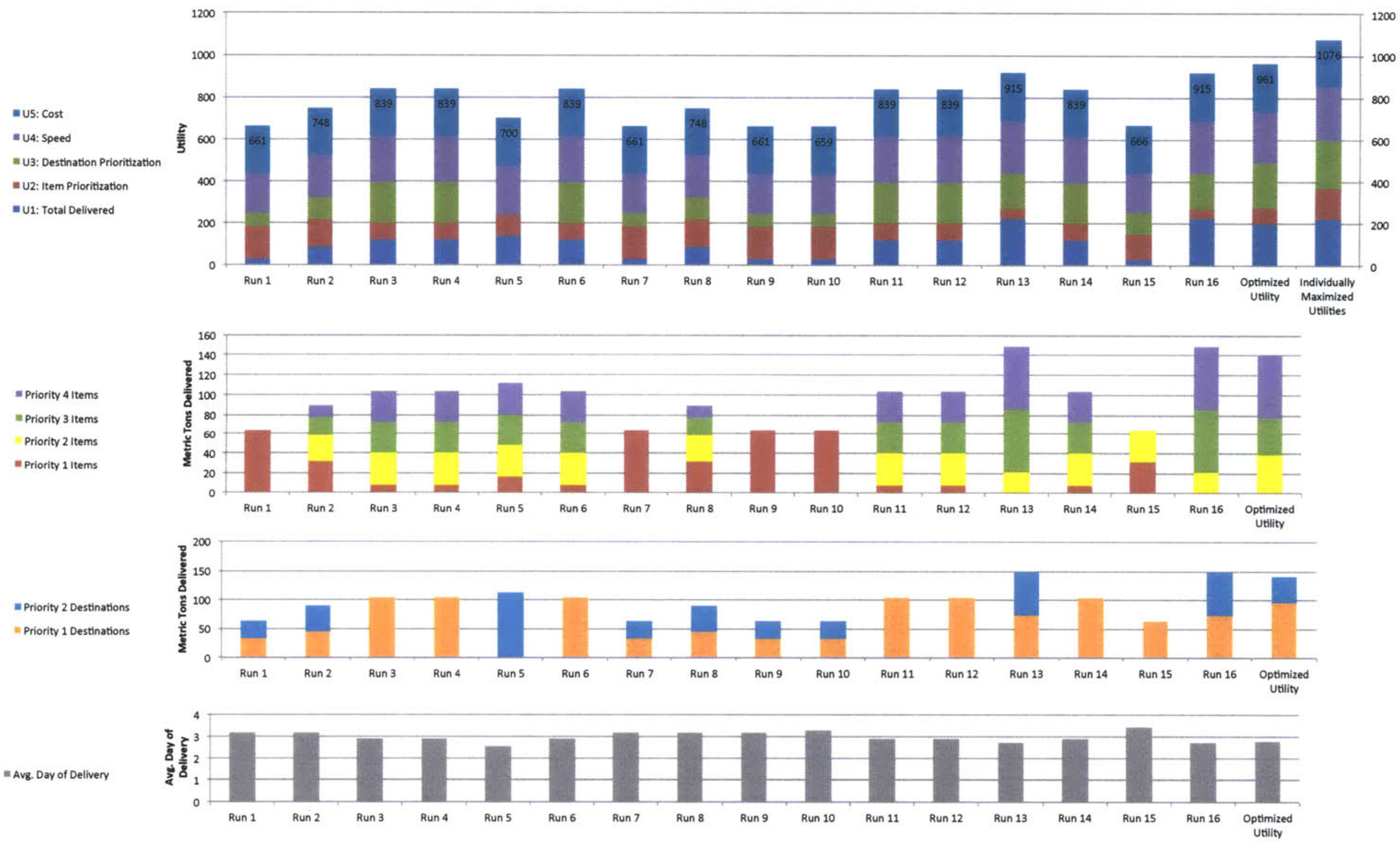


Figure 4-7: Results for "Feed Onward Transport" problem: low-capacity, difficult case

do a better job prioritizing by destination. The "cross" runs perform quite differently from one another. Selecting tasks based on destination and cargo by item type (run 15) results in fewer deliveries overall, like the item type prioritization policies. Selecting tasks based on item types and cargo by destination is equivalent to run 13. These two runs perform best out of all the heuristics because they allow the delivery of the lowest-priority items first, maximizing the total deliveries overall. Note, however, that the optimized solution does not deliver the maximum amount, but rather increases deliveries of priority-2 items, trading off some u_1 for u_2.

Overall, in this case, the best-performing policies are those that deliver the most cargo, in this case delivering very little of the high-priority item types because they take up the most capacity. The gain in utility from delivering more cargo overall outweighs the loss in utility from delivering lower-priority items. The policies that deliver the most cargo are those that prioritize by destination rather than item type. Recall that in the easy case described previously, the item type prioritization policies performed better than the destination prioritization policies. However, in the easy case, all heuristics delivered roughly the same amount of cargo overall, only trading off deliveries of various item types and to various destinations. In the difficult case, the trade-off is between delivering more high-priority items or delivering more cargo overall, and given the utility function, it is better to deliver more overall. Thus, destination policies perform better than item policies mainly because of poor performance of item policies rather than good performance of destination policies.

High-Capacity, Easy Prioritization Case Figure 4-8 summarizes the results for this case. Recall that the high-capacity case includes four trucks and four helicopters, twice as many as in the low-capacity case. See Figure 4-2 for details. With these parameters, the maximum possible amount of cargo to deliver is 204.8 MT.

Rather than going through these results in detail, we compare them with those of the low-capacity, easy prioritization case, to determine whether the increase in capacity changes the performance of the heuristics. One obvious difference is that more priority-1 cargo is delivered by all heuristics, simply because more cargo is delivered overall. The performance of each heuristic, however, is similar in character to that observed in the low-capacity case. The item prioritization policies perform slightly better than those based on destination, in

		Location-based						Task-based										Opt
		Item		Destination		Other		Item Prioritization				Destination Prioritization				Crossed		
		Lexic	Wtd	Lexic	Wtd	Short	Feed	Base	Wtd	TotNeed	FCFS	Base	Wtd	TotNeed	FCFS			
Run #:		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Opt
Gap:	U1: Total Delivered	0.0	0.0	26.4	13.2	26.4	0.0	0.0	0.0	0.0	0.0	52.8	52.8	0.0	52.8	52.8	0.0	0.0
	U2: Item Prioritization	13.1	13.1	94.7	94.7	94.7	82.5	13.1	13.1	13.1	13.1	95.5	95.5	94.7	95.5	54.6	89.7	0.0
	U3: Destination Prioritization	57.6	57.6	29.3	21.0	29.3	96.2	57.6	57.6	57.6	57.6	45.8	45.8	57.6	45.8	45.8	43.1	-
	U4: Speed	9.5	9.5	8.4	7.5	8.4	15.3	9.5	9.5	9.5	13.5	13.5	13.5	9.5	13.5	13.5	3.3	0.0
	U: Sum of utilities	80.4	80.4	158.5	136.3	158.5	194.3	80.4	80.4	80.4	84.4	207.2	207.2	162.0	207.2	166.4	136.3	-
Metric Tons Delivered:	Total delivered	204.8	204.8	179.2	192	179.2	204.8	204.8	204.8	204.8	204.8	153.6	153.6	204.8	153.6	153.6	204.8	204.8
	Priority 1 Items	64	64	32	32	32	32	64	64	64	64	32	32	12.8	32	57.6	12.8	63.99
	Priority 2 Items	64	64	32	32	32	44.8	64	64	64	64	32	32	64	32	32	64	64.01
	Priority 3 Items	64	64	51.2	64	51.2	64	64	64	64	64	32	32	64	32	32	64	44.81
	Priority 4 Items	12.8	12.8	64	64	64	64	12.8	12.8	12.8	12.8	57.6	57.6	64	57.6	32	64	31.99
	Priority 1 Destinations	102.4	102.4	128	128	128	76.8	102.4	102.4	102.4	102.4	128	128	102.4	128	128	102.4	128
	Priority 2 Destinations	102.4	102.4	51.2	64	51.2	128	102.4	102.4	102.4	102.4	25.6	25.6	102.4	25.6	25.6	102.4	76.8

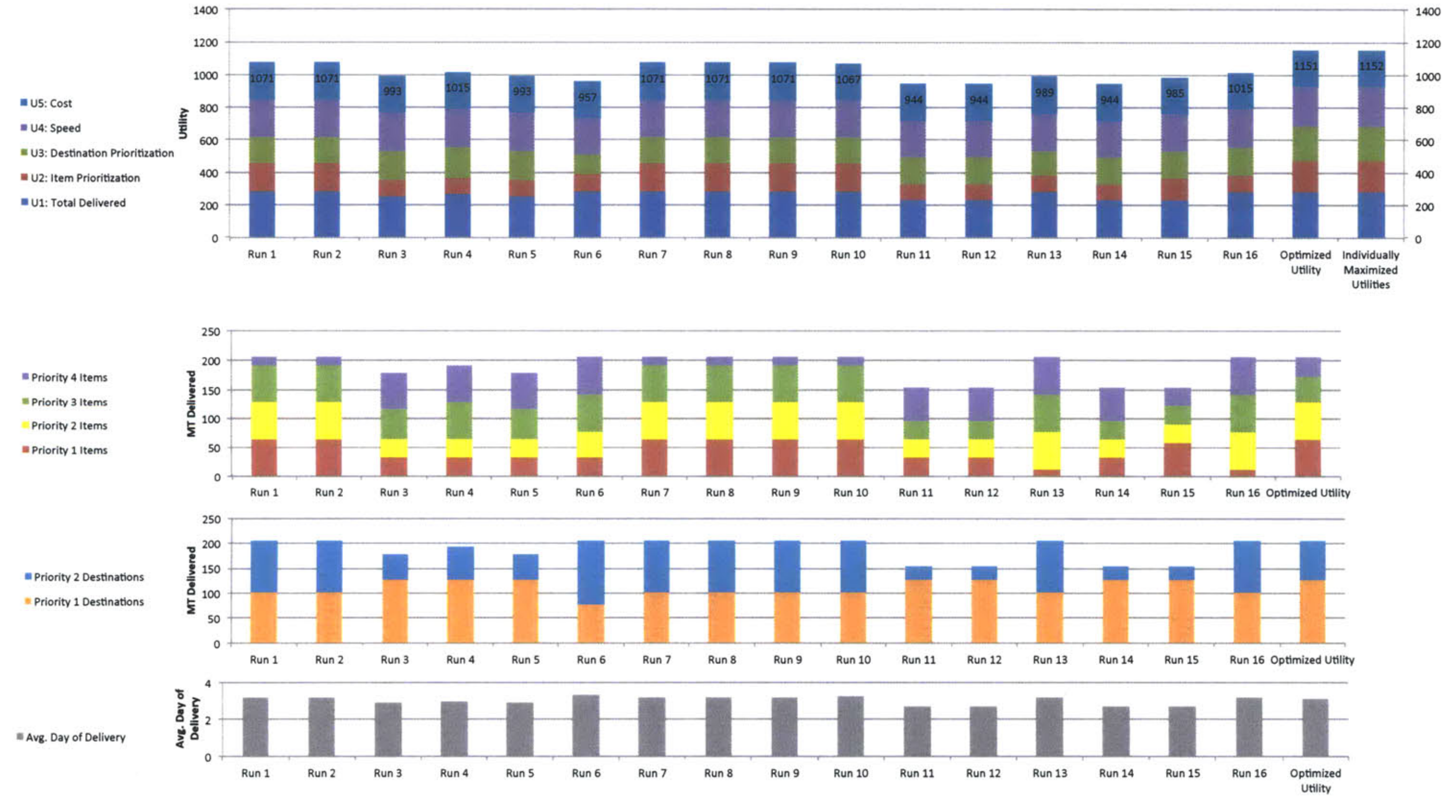


Figure 4-8: Results for "Feed Onward Transport" problem: high-capacity, easy case

both the location-based and task-based heuristics. It appears, based on these two cases, that the relative performance of the heuristics does not change when capacity is increased.

High-Capacity, Difficult Prioritization Case Figure 4-9 summarizes the results for this case. Recall that the high-capacity case includes four trucks and four helicopters, twice as many as in the low-capacity case, and that the difficult case makes it much harder to deliver higher-priority cargo, and assigns the farther node as a first-priority destination. See Figure 4-2 for details. With these parameters, the maximum possible amount of cargo to deliver is 224 MT.

Again, rather than going through these results in detail, they are compared with those of the low-capacity, difficult prioritization case, to determine whether the increase in capacity changes the performance of the heuristic. Thc conclusions are similar to those above for the easy prioritization case. The increase in capacity increases the amount of cargo delivered, enabling more delivery priority-1 items. However, the relative performance of the heuristics changes very little. The destination prioritization policies still perform better than the item prioritization policies, because they deliver more overall. The differences in performance between these heuristics are less pronounced, however, because all the heuristics are able to deliver at least three types of cargo, meaning the total weight delivered is more similar across heuristics. As we saw with the two easy cases, increasing the capacity from two trucks to four does not affect the relative performance of the heuristics in the difficult case.

Summary and Discussion The goal in studying these four cases of the "Feed Onward Transport" problem was to gain insights into how the set of heuristics and policies perform on this specific kind of network, when making high-priority deliveries is either harder (or not) than making low-priority deliveries, and when capacity is low or high.

None of the heuristics or policies perform as well as the optimization model, as might be expected, but in some cases the gap is small. However, in larger problems, the gap may increase. Interestingly, the 'feed' policy, which was designed based on this type of network, does not perform particularly well on this problem. However, the feed policy is more effective when the helicopters are a severely limited resource and a bottleneck to priority deliveries. In this simple problem, that was not the case.

There are no clear differences in performance between the location-based and task-based

		Location-based						Task-based										Opt
		Item		Destination		Other		Item Prioritization				Destination Prioritization				Crossed		
		Lexic	Wtd	Lexic	Wtd	Short	Feed	Base	Wtd	TotNeed	FCFS	Base	Wtd	TotNeed	FCFS			
Run #:		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Opt
Gap:	U1: Total Delivered	66.1	46.3	0.1	0.1	22.0	0.1	66.1	46.3	66.1	66.1	22.0	22.0	0.1	22.0	66.1	0.1	15.7
	U2: Item Prioritization	0.0	9.7	48.5	48.5	68.9	48.5	0.0	9.7	0.0	0.0	68.9	68.9	48.5	68.9	28.9	48.5	24.2
	U3: Destination Prioritization	72.5	58.0	0.1	0.1	80.5	0.1	72.5	58.0	72.5	72.5	0.1	0.1	24.2	0.1	22.8	24.2	-
	U4: Speed	48.9	35.8	19.7	7.5	14.9	11.4	48.9	35.8	48.9	51.4	30.2	30.2	2.1	30.2	49.3	2.1	15.8
	U: Sum of utilities	131.5	93.8	12.7	0.4	129.9	4.3	131.5	93.8	131.5	134.0	65.6	65.6	18.9	65.6	111.3	18.9	-
Metric Tons Delivered:	Total delivered	160	179.2	224	224	202.6667	224	160	179.2	160	160	202.6667	202.6667	224	202.6667	160	224	208.85
	Priority 1 Items	64	64	32	32	32	32	64	64	64	64	32	32	32	32	64	32	47.2
	Priority 2 Items	64	51.2	64	64	42.66667	64	64	51.2	64	64	42.66667	42.66667	64	42.66667	32	64	64.05
	Priority 3 Items	32	38.4	64	64	64	64	32	38.4	32	32	64	64	64	64	32	64	33.6
	Priority 4 Items	0	25.6	64	64	64	64	0	25.6	0	0	64	64	64	64	32	64	64
	Priority 1 Destinations	80	89.6	128	128	74.66667	128	80	89.6	80	80	128	128	112	128	128	112	128.04
	Priority 2 Destinations	80	89.6	96	96	128	96	80	89.6	80	80	74.66667	74.66667	112	74.66667	32	112	80.81

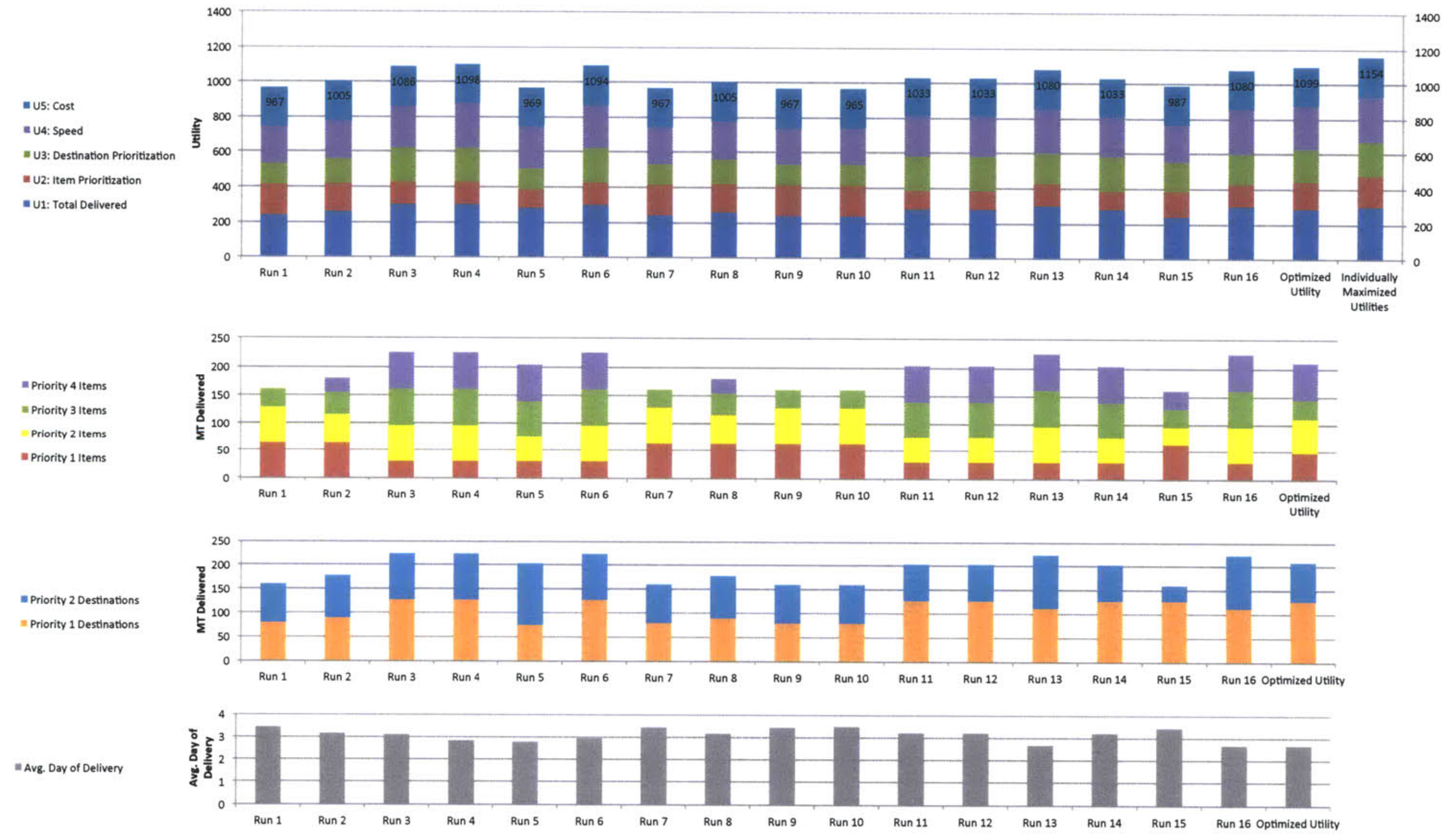


Figure 4-9: Results for "Feed Onward Transport" problem: high-capacity, difficult case

heuristic structure. The differences in performance across runs are due to the use of different policies for cargo selection, task selection, and vehicle allocation. There is a clear distinction between the performance of item type policies and destination policies. Item type prioritization policies performed better in the easy cases, because they enabled delivery of the right mix of items, without doing very poorly on prioritization by destination. In the difficult cases, the destination prioritization policies performed better, largely because the item type policies delivered large amounts of high-volume cargo that limited the total amount of cargo that could be delivered. There is also a distinction between the lexicographic and weighted mix policies, which is especially evident in the item prioritization policies: the weighted mix policies perform better than their lexicographic equivalents. Very little performance difference seems to be driven by the speed of delivery. In the easy cases, in particular, the speed of delivery is similar across all heuristics. In the difficult cases, the item prioritization policies are a little slower than the destination prioritization policies, adding to their relatively poor performance in these cases.

4.6.2 "Airbridge" Stylized Problem

A description of this problem was given in Figure 4-3. In this section, we look at four instances of this problem, shown in Table 4.2. For each case, sixteen combinations of heuristics and policies are tested; these sixteen "runs" were described in Table 4.3. As in the previous problem, note that cargo is loaded in the order in which it is listed in the database unless otherwise specified. As before, cargo was listed with all cargo for F before cargo for E, and with lower-priority item types before higher-priority item types, to emphasize the potential weaknesses in heuristics that do not pay attention to priority.

Easy Prioritization, No Airbridge Case Figure 4-10 summarizes the results for this case. This is a low-capacity case, with only two trucks and two helicopters, and equal amounts of all four types of items are requested to be delivered to each of the two destination nodes. It is 'easy' because all item types have the same volume and the closer destination node (F) is high priority. See Figure 4-3 for details. With these parameters, the maximum possible amount of cargo to deliver is 166.4 MT.

This case is labeled "No Airbridge" because the heuristics do not employ the airbridge option. Recall that the heuristics assign paths to cargo at the start of the algorithm, and an airbridge policy controls how much cargo (if any) is sent via shortest paths rather than least-cost paths. In this case, the policy was set so that no cargo was sent via shortest paths, meaning the airbridge from P to E was not used by any of the heuristics in this case.

The optimized solution has the maximum utility across all models, nearly as much as the utility obtained by maximizing each component individually. The solution provides a good mix of item types and delivers effectively to both destinations.

Among the location-based heuristics, those that prioritize by item type perform better than those that prioritize by destination, because they send a better mix of item types and are equally good at prioritizing by destination. However, all the heuristics do equally well at prioritizing by destination because there are no choices to be made: the helicopters always fly to F and the trucks always move to E because no cargo is routed along the airbridge. The same pattern appears in the task-based heuristics: those that select cargo by item type priority do better than those that do not. The first-come-first-served vehicle allocation policy (runs 10 and 14) makes no difference in performance, nor does selecting tasks based on the total need (runs 9 and 13). Interestingly, in the item type prioritization policies, using a weighted mix rather than lexicographic policy actually reduces the utility slightly; this is different from the behavior seen in the "Feed Onward Transport" problem. In general, however, these results show that item type prioritization policies perform better than destination prioritization policies, largely because the network structure forces the same destination prioritization on all of them. While none of the heuristic solutions reach the utility of the optimized solution, some are quite close, with gaps of only a few units of utility. However, the structure of this problem makes it easy for the heuristics to reach good solutions.

Easy Prioritization, Airbridge Case Figure 4-11 summarizes the results for this case. It is the same as the above case except that the heuristics may use the airbridge option. The amount of cargo using the airbridge must be pre-defined. Here, about 32 m^3 are routed through the airbrige. This amount was chosen because it is the equivalent of one day of work for one helicopter. Larger amounts could be tested as well, but this case is sufficient to see the effect of including the airbridge on the heuristics' performance. At the start of the algorithm, the cargo requests are sorted by their priority level (as specified by the prioritization policy), and the first 32 m^3 of cargo movement requests headed to E are

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Opt
		Location-based						Task-based										Opt
		Item		Destination		Other		Item Prioritization				Destination Prioritization				Crossed		
		Lexic	Wtd	Lexic	Wtd	Short	Feed	Base	Wtd	TotNeed	FCFS	Base	Wtd	TotNeed	FCFS			
Run #:		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Opt
Gap:	U1: Total Delivered	13.2	13.2	13.2	13.2	13.2	13.2	13.2	13.2	13.2	13.2	13.2	13.2	13.2	13.2	13.2	13.2	-
	U2: Item Prioritization	0.6	17.2	146.9	146.9	146.9	146.9	0.6	17.2	0.6	0.6	146.9	146.9	146.9	146.9	0.6	146.9	0.2
	U3: Destination Prioritization	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	U4: Speed	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.1
	U: Sum of utilities	10.6	27.2	156.9	156.9	156.9	156.9	10.6	27.2	10.6	10.6	156.9	156.9	156.9	156.9	10.6	156.9	-
Metric Tons Delivered:	Total delivered	153.6	153.6	153.6	153.6	153.6	153.6	153.6	153.6	153.6	153.6	153.6	153.6	153.6	153.6	153.6	153.6	166.4
	Priority 1 Items	64	57.6	6.4	6.4	6.4	6.4	64	57.6	64	64	6.4	6.4	6.4	6.4	64	6.4	64
	Priority 2 Items	51.2	47.36	32	32	32	32	51.2	47.36	51.2	51.2	32	32	32	32	51.2	32	57.6
	Priority 3 Items	32	39.68	51.2	51.2	51.2	51.2	32	39.68	32	32	51.2	51.2	51.2	51.2	32	51.2	25.6
	Priority 4 Items	6.4	8.96	64	64	64	64	6.4	8.96	6.4	6.4	64	64	64	64	6.4	64	19.2
	Priority 1 Destinations	102.4	102.4	102.4	102.4	102.4	102.4	102.4	102.4	102.4	102.4	102.4	102.4	102.4	102.4	102.4	102.4	102.4
	Priority 2 Destinations	51.2	51.2	51.2	51.2	51.2	51.2	51.2	51.2	51.2	51.2	51.2	51.2	51.2	51.2	51.2	51.2	64

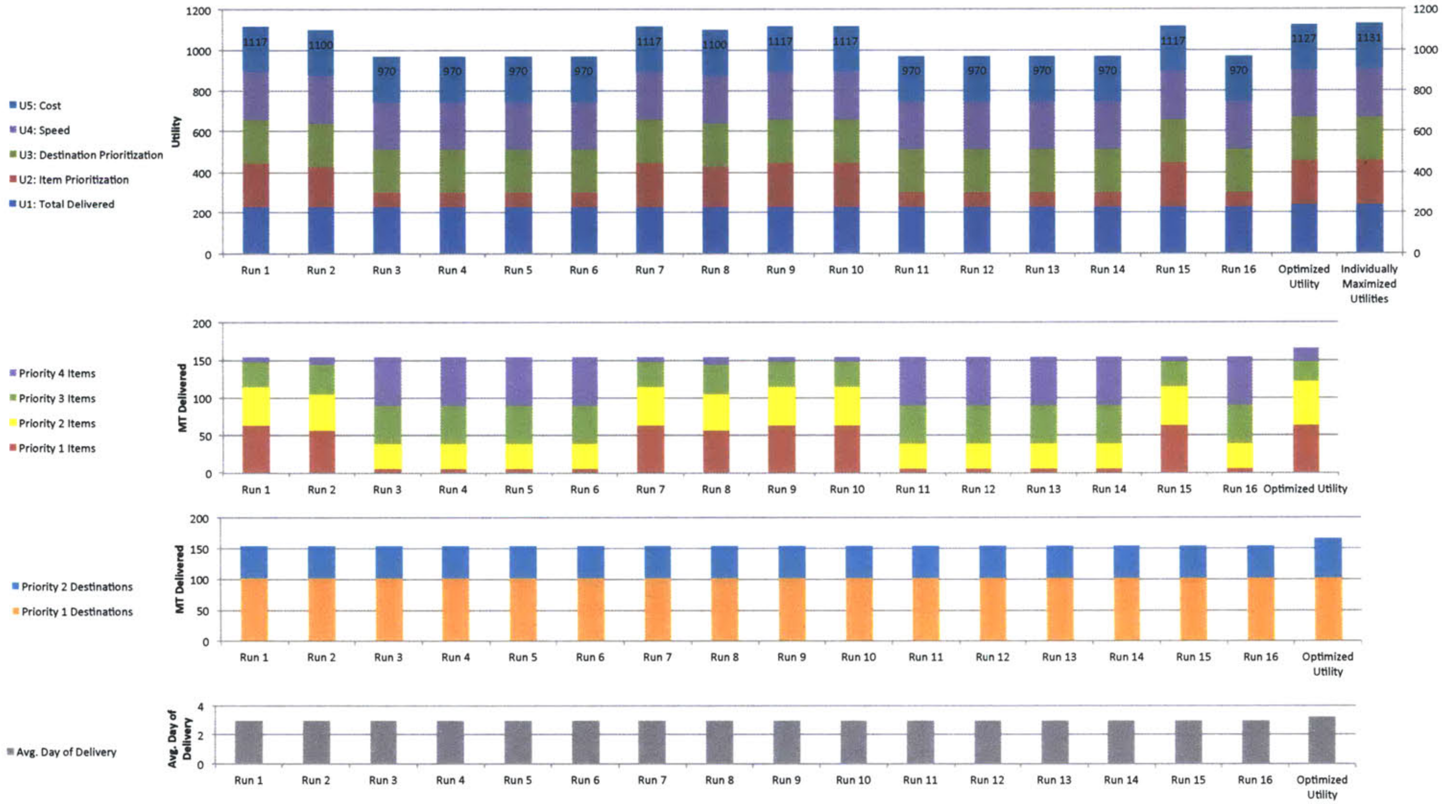


Figure 4-10: Results for "Airbridge" problem: easy case with no use of airbridge (in heuristics)

routed via the shortest path, which in this case is from P to E by helicopter. (The route applies to an entire request, so if the first request has more than 32 m^3 of cargo in it, the entire request is assigned a shortest path.) The remaining cargo movement requests are routed via the least-cost path, which is by road from P through D to E. The heuristic then proceeds in the normal manner, meaning it may send none or all of the airbridge cargo, so up to 32 m^3 may be sent via airbridge.

The optimized solution is the same in this case as in the previous case because the airbridge parameter applies only to the heuristics; the optimizer determines how much cargo to send on each route. In this problem, it does *not* send any cargo via the airbridge. There is no reason to use it because the high-priority destination is F, which is accessible only by helicopter. Use of the airbridge would take helicopter flights away from F, and the gain in utility from delivering more speedily to E is not worth the loss in utility from delivering less to F. With these results, we expect the heuristics to perform less well with the airbridge option, because it forces certain cargo movement requests to be moved via the airbridge.

Comparing the performance of the heuristics in this case to that in the previous case (with no airbridge) shows that the airbridge decreases the performance of some of the heuristics. In particular, the item type prioritization policies perform less well on prioritizing by destination. This is because they deliver cargo in order of item type priority, and many of those high-priority items were assigned to the airbridge path. They would otherwise be delivered by truck, but because they were assigned this path, they are delivered by air, taking helicopter capacity away from the high-priority node F that is accessible only via helicopter.

This case makes clear that an airbridge does not always enable better solutions; in this case, forcing the use of an airbridge actually makes them worse, because it takes resources away from more important uses of the helicopters. However, we might expect the airbridge to be useful in the difficult case, in which the priority of E, the node reached by the airbridge, is higher.

Difficult Prioritization, No Airbridge Case Figure 4-12 summarizes the results for this case. No airbridge routes are included in the heuristic models, and the problem parameters make it more difficult to deliver higher-priority cargo. Node E, the more difficult-

		Location-based						Task-based										Opt
		Item		Destination		Other		Item Prioritization				Destination Prioritization				Crossed		
		Lexic	Wtd	Lexic	Wtd	Short	Feed	Base	Wtd	TotNeed	FCFS	Base	Wtd	TotNeed	FCFS			
Run #:		**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**Opt**
Gap:	U1: Total Delivered	13.2	13.2	13.2	13.2	13.2	13.2	13.2	20.1	13.2	13.2	13.2	13.2	13.2	13.2	13.2	13.2	-
	U2: Item Prioritization	2.0	16.3	123.8	144.2	146.9	153.8	2.0	0.7	63.1	2.0	123.8	123.8	130.7	123.8	69.3	123.8	0.2
	U3: Destination Prioritization	65.0	91.0	0.0	26.0	0.0	6.5	65.0	71.5	6.5	65.0	0.0	0.0	6.5	0.0	0.0	0.0	0.0
	U4: Speed	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.1
	U: Sum of utilities	76.9	117.2	133.8	180.1	156.9	170.2	76.9	89.4	79.6	76.9	133.8	133.8	147.1	133.8	79.2	133.8	-
Metric Tons Delivered:	Total delivered	153.6	153.6	153.6	153.6	153.6	153.6	153.6	150.4	153.6	153.6	153.6	153.6	153.6	153.6	153.6	153.6	166.4
	Priority 1 Items	64	56.96	6.4	0	6.4	3.2	64	64	35.2	64	6.4	6.4	3.2	6.4	32	6.4	64
	Priority 2 Items	64	49.28	51.2	44.8	32	32	64	53.76	64	64	51.2	51.2	51.2	51.2	64	51.2	57.6
	Priority 3 Items	25.6	39.68	64	64	51.2	54.4	25.6	26.72	51.2	25.6	64	64	64	64	51.2	64	25.6
	Priority 4 Items	0	7.68	32	44.8	64	64	0	5.92	3.2	0	32	32	35.2	32	6.4	32	19.2
	Priority 1 Destinations	70.4	57.6	102.4	89.6	102.4	99.2	70.4	67.2	99.2	70.4	102.4	102.4	99.2	102.4	102.4	102.4	102.4
	Priority 2 Destinations	83.2	96	51.2	64	51.2	54.4	83.2	83.2	54.4	83.2	51.2	51.2	54.4	51.2	51.2	51.2	64

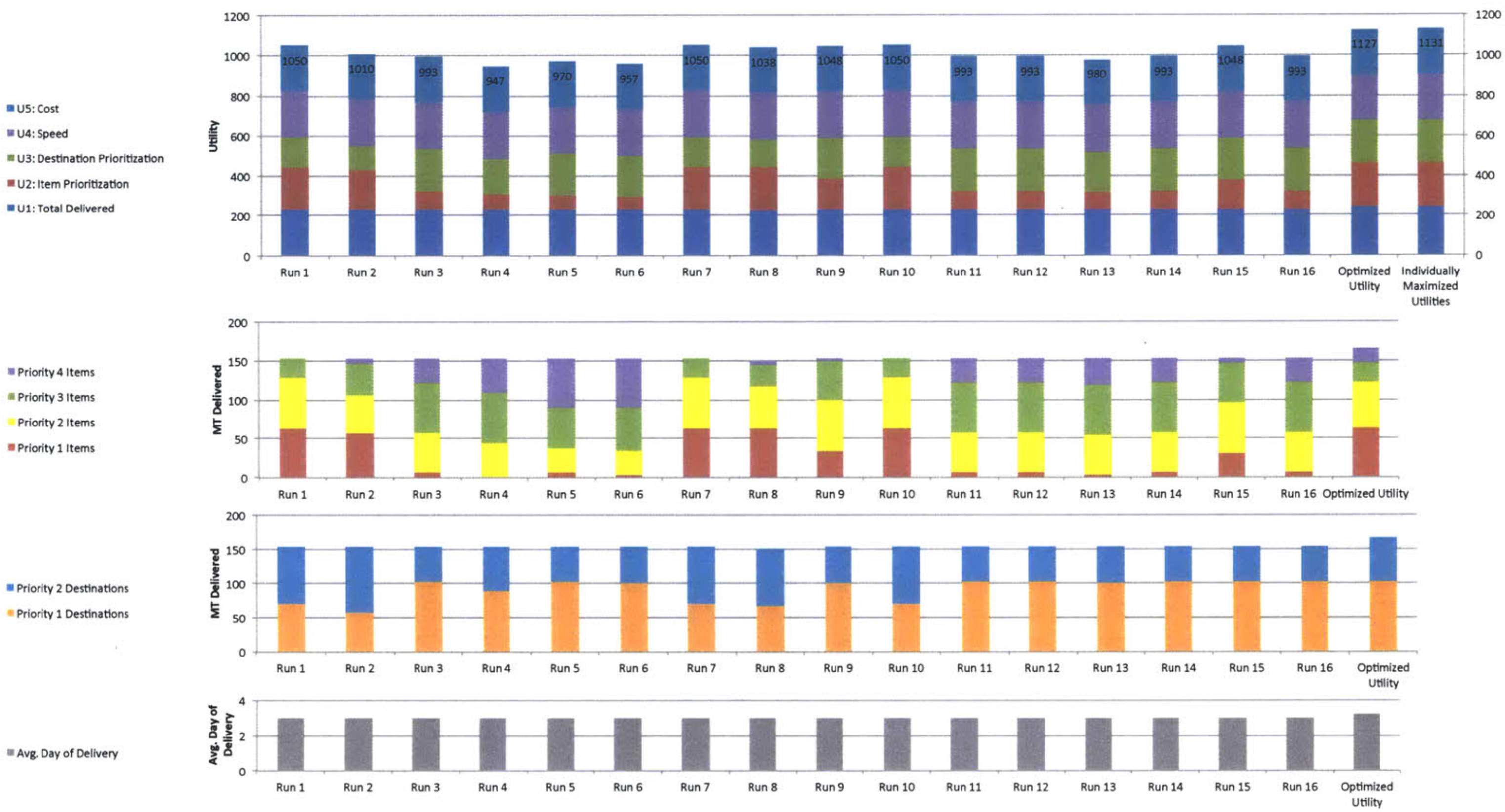


Figure 4-11: Results for "Airbridge" problem: easy case with 32 m^3 via airbridge (in heuristics)

to-reach node, is the high-priority node. In this case, we expect the airbridge to be more useful. In this difficult case, the maximum amount of cargo that can be delivered is 200 MT.

The optimized solution is not as good as the sum of the individually maximized component utilities, mainly because it sacrifices prioritizing by item type in order to deliver more cargo overall. In this difficult prioritization case, delivering higher-priority items takes up more space (higher-priority items have larger volume), thus reducing the total capacity for delivery. As a consequence, the optimized solution delivers less high-priority items because the utility function values overall deliveries more than good prioritization. A second consequence is that the policies that prioritize by item type deliver less overall and have lower utilities than the other policies. The same dynamics were found in the "Feed Onward Transport" difficult case.

An additional weakness of the heuristics is in prioritizing by destination. The optimized utility delivers to a good mix of destinations, with more cargo going to the high-priority node E than to the low-priority node F. However, without the use of the airbridge, the mix of destinations to which cargo is delivered is pre-determined because the helicopters can only fly to F and the trucks can only reach E. The destination prioritization policies do better than the item prioritization policies only because they are not forced to deliver large amounts of high-priority cargo, not because they make better destination decisions. The airbridge, however, should alleviate this effect somewhat.

Difficult Prioritization, Airbridge Case Figure 4-13 summarizes the results for this case. The highest-priority (defined according to the invoked policy) 32 m^3 of cargo are assigned to move via the airbridge, and the remaining cargo are assigned to lowest-cost paths.

The optimized solution is the same as that of the previous case, because the enabling of the airbridge applies only to the heuristics. The optimizer chooses exactly how much cargo to send via the airbridge. The optimized solution includes 20 flights between P and E, transporting 160 m^3 of cargo, much more than the 32 m^3 used here. Clearly, using the airbridge only until cargo can reach the remote destination, as the teams did, is not the best solution. Given the inadequacy of the amount of cargo routed via airbridge, we might expect only minor improvements in the heuristic solution. Indeed, examining the results of

		Location-based						Task-based										Opt
		Item		Destination		Other		Item Prioritization				Destination Prioritization				Crossed		
		Lexic	Wtd	Lexic	Wtd	Short	Feed	Base	Wtd	TotNeed	FCFS	Base	Wtd	TotNeed	FCFS			
Run #:		**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**Opt**
Gap:	U1: Total Delivered	140.4	90.5	13.7	13.7	13.7	13.7	140.4	90.5	140.4	140.4	13.7	13.7	13.7	13.7	140.4	13.7	12.3
	U2: Item Prioritization	14.5	27.5	79.5	79.5	79.5	79.5	14.5	27.5	14.5	14.5	79.5	79.5	79.5	79.5	14.5	79.5	53.7
	U3: Destination Prioritization	162.3	140.6	90.2	90.2	90.2	90.2	162.3	140.6	162.3	162.3	90.2	90.2	90.2	90.2	162.3	90.2	0.0
	U4: Speed	51.8	38.8	2.5	2.5	2.5	2.5	51.8	38.8	51.8	51.8	2.5	2.5	2.5	2.5	51.8	2.5	4.2
	U: Sum of utilities	298.7	227.2	115.6	115.6	115.6	115.6	298.7	227.2	298.7	298.7	115.6	115.6	115.6	115.6	298.7	115.6	-
Metric Tons Delivered:	Total delivered	112	134.4	186.6667	186.6667	186.6667	186.6667	112	134.4	112	112	186.6667	186.6667	186.6667	186.6667	112	186.6667	188.03
	Priority 1 Items	64	48	16	16	16	16	64	48	64	64	16	16	16	16	64	16	32
	Priority 2 Items	32	38.4	42.66667	42.66667	42.66667	42.66667	32	38.4	32	32	42.66667	42.66667	42.66667	42.66667	32	42.66667	40.03
	Priority 3 Items	16	28.8	64	64	64	64	16	28.8	16	16	64	64	64	64	16	64	52
	Priority 4 Items	0	19.2	64	64	64	64	0	19.2	0	0	64	64	64	64	0	64	64
	Priority 1 Destinations	32	44.8	74.66667	74.66667	74.66667	74.66667	32	44.8	32	32	74.66667	74.66667	74.66667	74.66667	32	74.66667	128.03
	Priority 2 Destinations	80	89.6	112	112	112	112	80	89.6	80	80	112	112	112	112	80	112	60

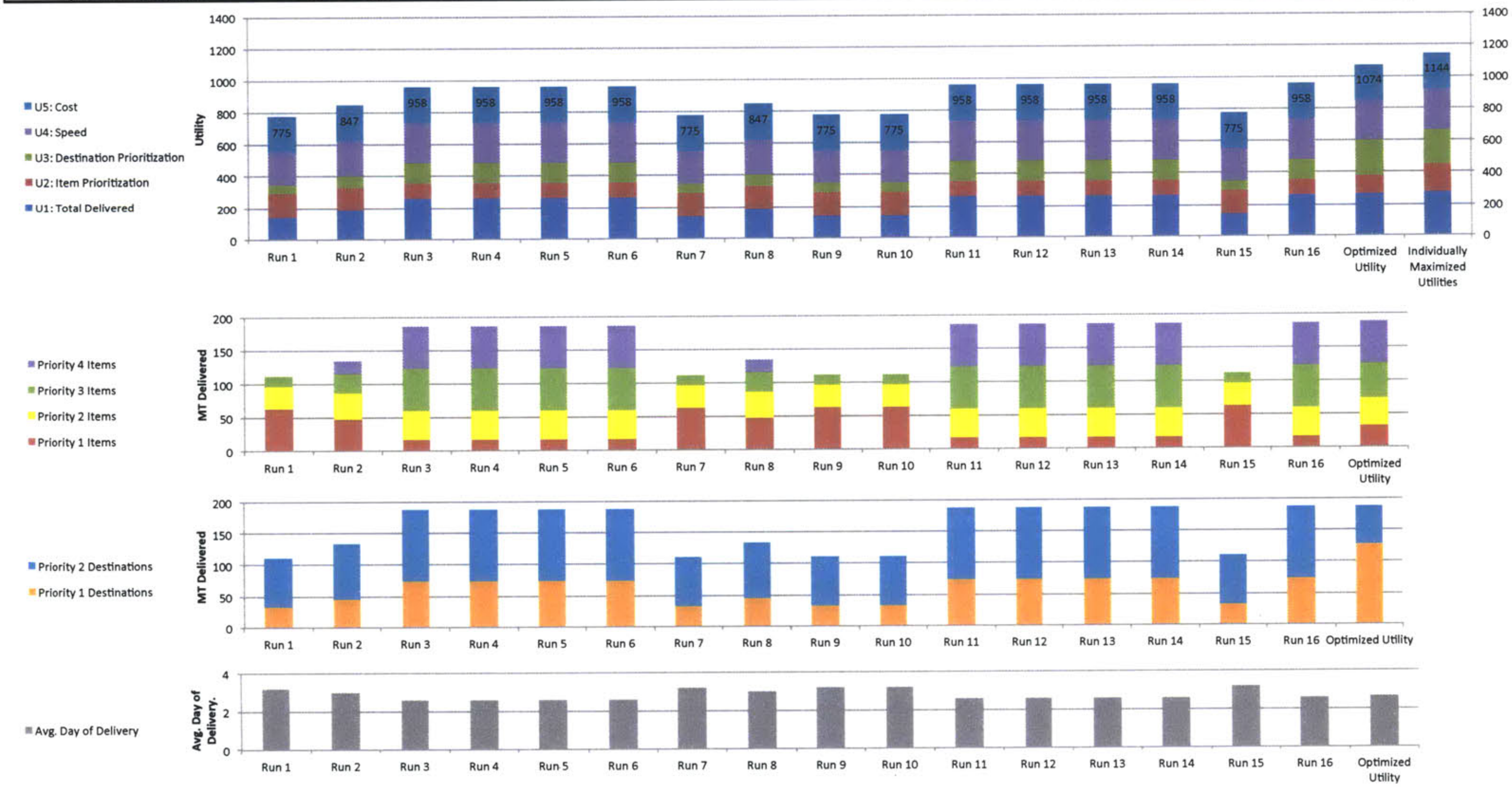


Figure 4-12: Results for "Airbridge" problem: difficult case with no airbridge (in heuristics)

		Location-based						Task-based										Opt
		Item		Destination		Other		Item Prioritization				Destination Prioritization				Crossed		
		Lexic	Wtd	Lexic	Wtd	Short	Feed	Base	Wtd	TotNeed	FCFS	Base	Wtd	TotNeed	FCFS			
Run #:		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Opt
Gap:	U1: Total Delivered	140.4	71.7	11.0	11.0	13.7	8.9	140.4	95.8	104.8	140.4	11.0	11.0	17.2	11.0	140.4	35.8	12.3
	U2: Item Prioritization	14.5	40.5	82.3	82.3	79.5	85.9	14.5	27.2	34.1	14.5	82.3	82.3	78.8	82.3	14.5	68.0	53.7
	U3: Destination Prioritization	101.4	67.2	72.1	72.1	90.2	58.6	101.4	93.5	135.2	101.4	72.1	72.1	85.7	72.1	101.4	126.2	0.0
	U4: Speed	48.9	36.2	0.0	0.0	2.5	2.1	48.9	40.2	46.5	48.9	0.0	0.0	11.9	0.0	48.9	13.2	4.3
	U: Sum of utilities	234.9	145.2	95.1	95.1	115.6	85.2	234.9	186.3	250.3	234.9	95.1	95.1	123.1	95.1	234.9	172.9	-
Metric Tons Delivered:	Total delivered	112.0	142.8	189.3	189.3	186.7	191.3	112.0	132.0	128.0	112.0	189.3	189.3	183.3	189.3	112.0	165.3	188.0
	Priority 1 Items	64.0	35.6	8.0	8.0	16.0	2.0	64.0	43.0	32.0	64.0	8.0	8.0	10.0	8.0	64.0	16.0	32.0
	Priority 2 Items	32.0	46.4	53.3	53.3	42.7	61.3	32.0	48.8	64.0	32.0	53.3	53.3	53.3	53.3	32.0	53.3	40.0
	Priority 3 Items	16.0	41.6	64.0	64.0	64.0	64.0	16.0	25.4	32.0	16.0	64.0	64.0	64.0	64.0	16.0	64.0	52.0
	Priority 4 Items	-	19.2	64.0	64.0	64.0	64.0	-	14.8	-	-	64.0	64.0	56.0	64.0	-	32.0	64.0
	Priority 1 Destinations	80.0	92.4	85.3	85.3	74.7	93.3	80.0	73.2	48.0	80.0	85.3	85.3	77.3	85.3	80.0	53.3	128.0
	Priority 2 Destinations	32.0	50.4	104.0	104.0	112.0	98.0	32.0	58.8	80.0	32.0	104.0	104.0	106.0	104.0	32.0	112.0	60.0

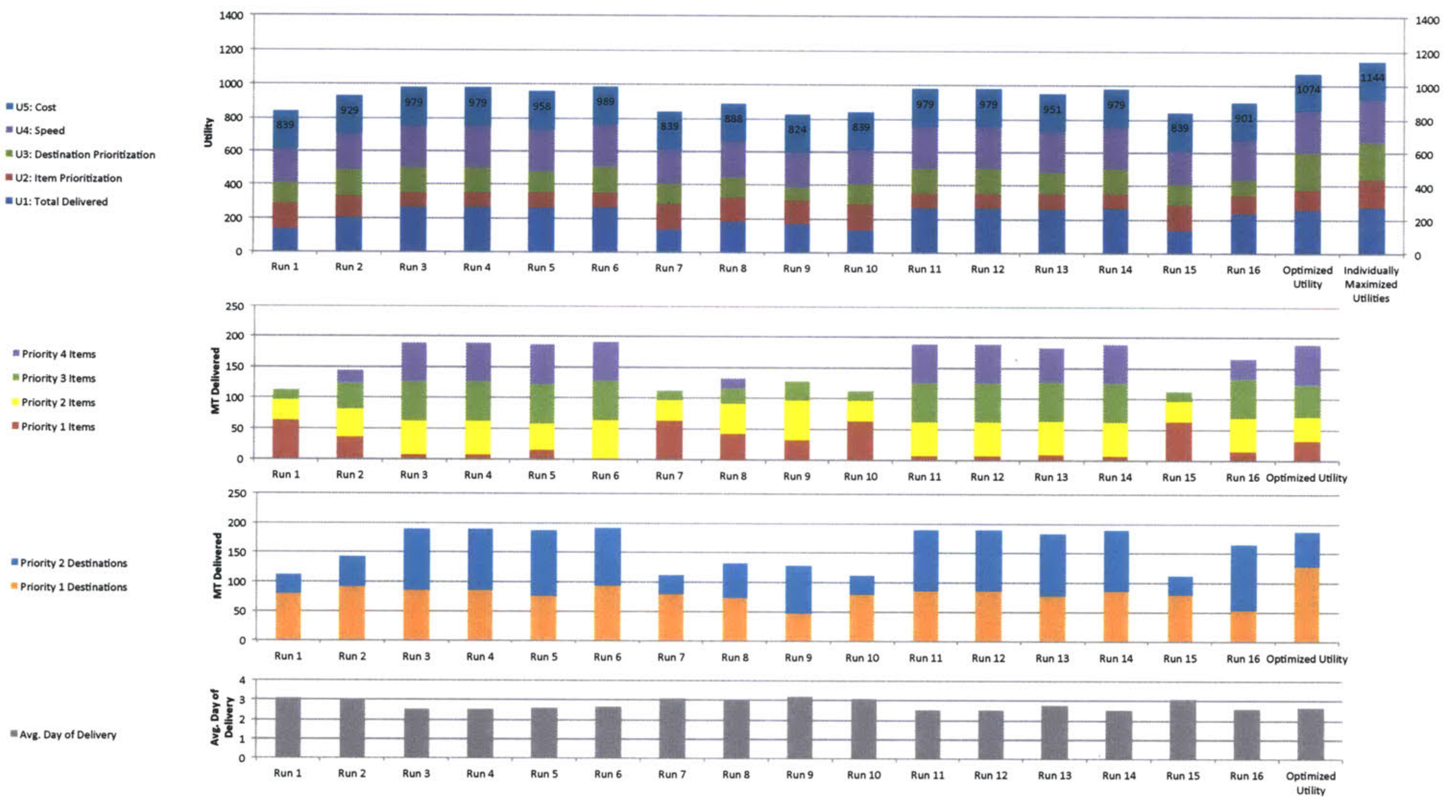


Figure 4-13: Results for "Airbridge" problem: difficult case with 32 m^3 via airbridge (in heuristics)

this case in comparison with the previous case (same difficulty but with no airbridge) shows that there were minor improvements in performance on prioritization by destination, but few other changes.

Summary and Discussion The main conclusion from the "Airbridge" case is that choosing whether or not to utilize an airbridge, and deciding how much cargo to send along it, can have a major influence on the performance of the heuristics. The airbridge was not useful in the easy case, but was quite useful in the difficult case. The optimizer was able to discover this and find the right amount of cargo to send via the airbridge, but the heuristics do not have this capability. The heuristics require the pre-definition of paths for cargo, which models the behavior of humanitarian logisticians, so the amount of cargo to send via airbridge is determined ahead of time. In these cases, 32 m^3 of cargo were assigned to go via airbridge. The amount was chosen based on the observed behavior of humanitarian logisticians: they often chose to use an airbridge only until cargo started arriving by road, and also chose not to use all their vehicles for the airbridge. The 32 m^3 policy was chosen because it is the the capacity of one of the two helicopters in one day of work (and cargo started arriving by truck on the second day). The results show that this amount was too high in the easy case and too low in the difficult case. Performance changed in the expected direction, but only slight increases (or decreases) in performance were observed. We conclude that the choice of whether and how much to utilize an airbridge is important, and should not be left to simple heuristics like that used by the humanitarian logisticians (and modeled here in a deliberately naive manner). When an airbridge opportunity is present in a network, the optimization model is able to utilize it effectively, but the heuristics (as defined here) are unable to do so without an additional heuristic to determine the best way to use it.

4.6.3 "Snowland" Problem

The previous two stylized problems were intended to provide some intuition on how well different combinations of heuristics and policies performed under different circumstances. With this problem, it will be more difficult to understand the reasons heuristics perform well or poorly, but it will provide a better indication of the capabilities of heuristics on a problem of realistic size and complexity. It is particularly appropriate to test heuristics on this "Snowland" problem because it is based on the problem the teams solved in the LRT training, on whose behavior these heuristics are modeled. If they perform well, it would suggest that humans may be reasonably good at solving realistic problems. If the heuristics perform poorly, identifying their weaknesses will enable the development of training and tools to enhance humans' abilities to solve this type of problem.

Two versions of the "Snowland" problem were described earlier: the exact problem solved by the LRT teams, and a slightly simplified version that can be solved by the optimizer for 6-day instances. We study the latter version here, because we wish to compare the heuristic performance to that of the optimizer on a 6-day plan. Planning only 1 or 2 days ahead is a much simpler problem, and our goal is to test the performance of heuristics in a more complex environment. A summary of the problem was provided in Figure 4-5.

Mid-Capacity Case Figure 4-14 summarizes the results for the mid-capacity case. The maximum amount that can be delivered is 1,886 MT, which is about 93% of the total. (Most of the heuristics, and all of the human teams, were unable to deliver this amount of cargo.)

The optimized solution is quite close to the individually maximized utilities, suggesting it does about as well as possible for this problem. The mix of destinations is about equal to the ideal mix (70% of deliveries to high-priority destinations). The mix of item types does not appear to be close to the ideal mix (50% to priority-1 items, 30% to priority-2, 15% to priority-3, and 5% to priority-4). However, the cargo movement requests in this problem were not evenly divided among item types, but contained many more requests for delivery of priority-1 and priority-4 items than the other two types; this mix of cargo types is the best that can be achieved in this problem (and is therefore the right mix in this particular problem context).

Comparing the optimized solution to solutions from all the heuristics shows the superiority of the optimization approach. In total utility scores, some of the heuristics are close to the optimal solution, but closer inspection reveals weaknesses in their solutions. The optimized solution delivers more cargo overall, more priority-1 items, at least as many priority-2 items, more priority-3 items, more priority-4 items, and more cargo to priority-1 destinations than any of the heuristics, so it is better on nearly all dimensions of the utility function.

		Location-based						Task-based										Opt
		Item		Destination		Other		Item Prioritization				Destination Prioritization				Crossed		
		Lexic	Wtd	Lexic	Wtd	Short	Feed	Base	Wtd	TotNeed	FCFS	Base	Wtd	TotNeed	FCFS			
Run #:		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Opt
Gap:	U1: Total Delivered	42.3	37.3	11.6	41.0	19.0	46.5	49.6	54.7	49.3	57.0	57.0	47.7	54.9	57.0	50.9	54.9	1.1
	U2: Item Prioritization	7.7	4.8	2.0	6.8	2.1	14.5	7.9	5.2	7.9	7.9	31.2	17.5	26.4	31.2	7.9	26.4	0.1
	U3: Destination Prioritization	49.2	45.7	27.1	62.3	37.4	54.5	60.6	68.7	62.1	64.7	60.6	48.6	63.1	60.6	56.2	63.1	3.4
	U4: Speed	32.5	24.7	25.4	12.5	3.2	12.7	63.6	62.7	63.5	63.1	58.4	47.9	59.2	58.1	63.1	59.2	0.3
	U: Sum of utilities	126.7	107.6	61.2	117.5	56.9	123.2	176.8	186.5	177.8	187.8	202.4	156.7	198.6	202.0	173.2	198.6	-
Metric Tons Delivered:	Total delivered	1,484.3	1,522.8	1,721.8	1,494.6	1,664.5	1,451.5	1,427.4	1,387.9	1,430.4	1,370.2	1,370.0	1,442.5	1,387.0	1,370.0	1,417.4	1,387.0	1,803.3
	Priority 1 Items	1,005.6	1,000.6	1,106.5	971.6	1,051.6	894.2	1,043.1	997.9	1,053.1	979.0	799.8	872.3	831.6	799.8	1,026.2	831.6	1,149.5
	Priority 2 Items	18.8	26.7	26.7	23.8	18.8	26.7	18.8	26.7	18.8	18.8	25.5	25.5	18.8	25.5	18.8	18.8	26.7
	Priority 3 Items	19.9	55.4	104.9	26.2	118.1	47.0	16.1	47.8	16.1	16.1	69.2	69.2	61.9	69.2	16.1	61.9	136.9
	Priority 4 Items	440.0	440.0	440.0	440.0	440.0	440.0	317.5	283.7	310.6	324.5	440.0	440.0	440.0	440.0	324.5	440.0	443.1
	Priority 1 Destinations	985.3	990.4	1,079.9	883.8	1,022.6	958.0	916.7	866.9	899.7	919.6	959.4	1,022.7	922.3	959.4	966.8	922.3	1,212.7
	Priority 2 Destinations	499.1	532.4	641.9	610.7	641.9	493.5	510.7	521.0	530.7	450.6	410.6	419.8	464.7	410.6	450.6	464.7	590.6

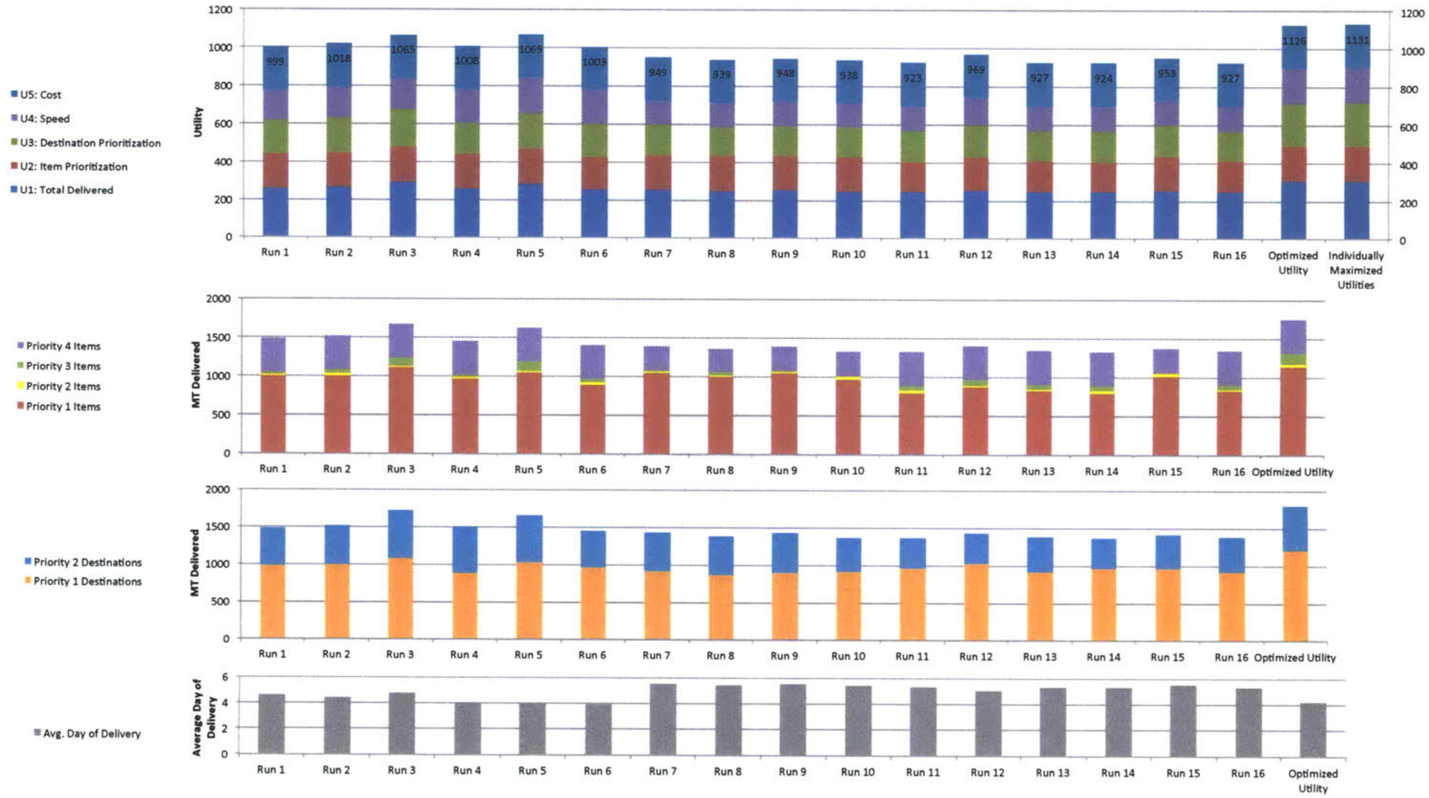


Figure 4-14: Results for "Snowland" problem: mid-capacity case

Across the location-based heuristics, the two best are the shortest-path policy (run 5), which prioritizes cargo with the shortest distance to its final destination, and the lexicographic destination-prioritization policy. The former, the shortest-path policy, is weak because it tends to ignore certain high-priority destinations (this is much more obvious on close inspection of the results than in the summaries provided here). The latter, the lexicographic destination policy, is more promising because it performs reasonably well on most dimensions, with the largest utility gaps in destination-prioritization and in speed. The location-based heuristics with item type prioritization policies generate mediocre solutions; they are particularly weak in the total amount delivered and in destination prioritization.

While the lexicographic destination policy is one of the best, the weighted mix destination policy does not perform very well, delivering less cargo to priority destinations and fewer priority items. Probably, the weighted mix policy does less well because it sacrifices some deliveries to high-priority destinations for low-priority destinations, and the lexicographic policy does well because there are enough easy-to-reach low-priority destinations (such as D) to satisfy the 30% target in the utility function.

The shortest-path policy (run 5) performs quite well, but closer inspection of the solution shows that it delivers almost nothing to the areas around B. By delivering large amounts of cargo to the areas around A, it reaches a high total for cargo delivered to high-priority locations, but this uneven distribution of deliveries is a weakness. The "feed" policy (run 6) is the worst of the location-based heuristics. This policy aimed to feed cargo to A to ensure that enough cargo was available to be shipped onward to B, keeping the helicopters busy. Many of the teams in the LRT training were very concerned with sending enough cargo to the helicopters, but this heuristic does not appear to be very effective. Either it was not actually important to keep the helicopters busy, or the other heuristics happen to keep them busy anyway; the latter is more likely.

The task-based heuristics generate worse solutions than any of the location-based heuristics. Solutions 7 through 16 deliver much less cargo overall, and exhibit worse mixes of item types and locations. The tasks defined in the algorithm were based on the tasks formulated by teams observed in the LRT training: delivering from A to B, delivering from P, delivering from A, delivering from B, and delivering from D. The heuristics look at the needs across these tasks and allocate vehicles proportionally to the needs. In this problem, it seems, this general strategy fails. One possible reason is that changing a vehicle from one task to another requires it to make a journey from the first task to the second. For example, if a vehicle was working on the task delivering from A, and then was reallocated to delivering from D, it would be required to move from A to D. The heuristic is reasonably intelligent about these decisions: it selects the nearest vehicle to allocate to each task, and it tries to load cargo onto vehicles even when they are transitioning from one task to another. Nevertheless, such journeys may be less productive in terms of utility. A second possible reason is that allocating vehicles in proportion to task needs may not be the right strategy. Some tasks are perhaps less important than others, despite having high needs. However, this possibility is rendered less likely because the first-come-first-served policy (runs 10 and 14) shows no improvement over the proportional policy. It is much more likely that the task-based policies are simply less efficient in making use of the available vehicles than the location-based policies.

Mid-Capacity Case, with Airbridge The same Snowland problem was solved again with the airbridge policy in place for all heuristics: approximately 28 MT of high-priority cargo was routed from P to B via C, rather than all the way around through D, H, and A. Again, the amount of cargo to route through the airbridge was chosen based on the ways in which the observed teams made this decision: they allocated half of their helicopter capacity for days 1 and 2 to fly between B and C, while the other half made deliveries from B. After days 1 and 2, the teams never used the helicopters as an airbridge, reasoning that by then cargo had started to arrive by road. The results are not shown because they show little difference from those without an airbridge. However, the optimized solution does make use of the airbridge option, sending four large truckloads for cargo (about 52 tons) to C for onward transport by helicopter. These deliveries occurred *after* the first two days, suggesting they were not intended to increase the speed of delivery but rather to increase capacity. The optimizer was probably attempting to quickly increase the total amount of cargo delivered in the last few days of the scenario. As we found in the stylized "Airbridge" problem, it appears that people do not make good decisions about the utilization of airbridges.

High-Capacity Case A second Snowland instance was solved with a different configuration of vehicles that resulted in a higher capacity. The main difference is that a much larger number of small (10-ton) trucks were made a available, and they were made available several days earlier. In theory, this should result in a large increase in capacity. However, the capacity of the scenario previously discussed is already fairly high, and the additional trucks appear to enable only a small number of additional deliveries, probably to low-priority but hard-to-reach locations.

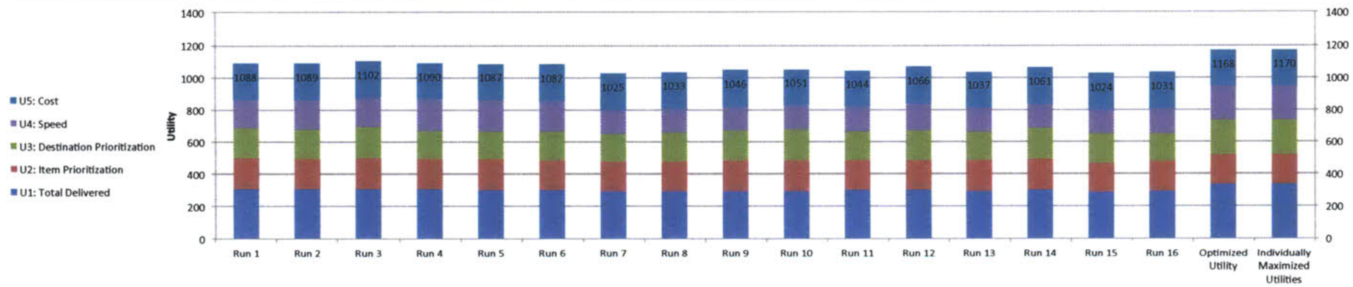
Figure 4-15 summarizes the results for this case. The optimizer is able to deliver almost all the cargo. In some ways, this case is less interesting because prioritization decisions are not important when all the cargo can be delivered. However, none of the heuristics are able to deliver all the cargo. They are able to deliver more cargo than in the low-capacity case, and as a result the performance is more similar across heuristics and policies. The difference between the location-based and task-based heuristics are less clear, as are the differences between policies, though they are not qualitatively different from what was found in the lower-capacity case. The difference between the optimized solution and those of the heuristics, especially in the total amount of cargo delivered, is clear.

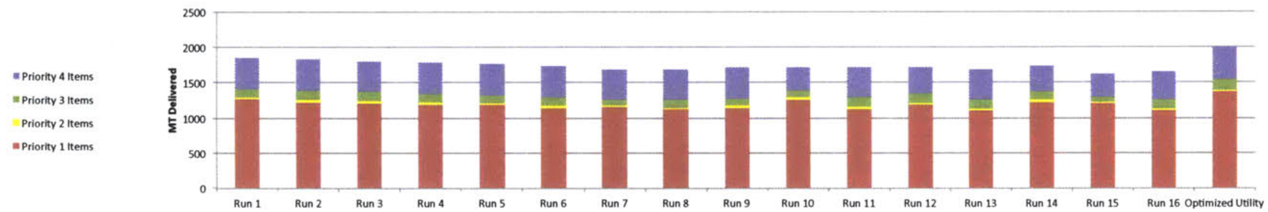
Summary and Discussion The goal in testing heuristics on the Snowland problem was to understand how well they perform on a particular problem of realistic size and complexity. This particular problem should show the heuristics to best advantage, because the heuristics are based on the behavior of people solving exactly this problem.

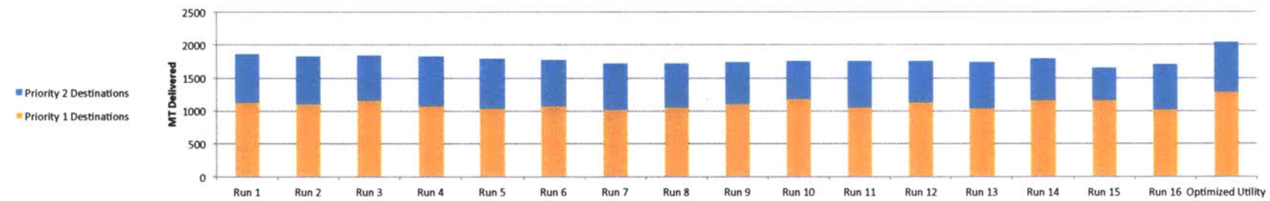
The results suggest that location-based heuristics are superior to task-based heuristics, because they use vehicles more efficiently. We originally expected that task-based heuristics would be superior because they consider the need for cargo movement "globally", across the entire network, rather than "locally", only at the individual nodes. We reasoned that they should be able to get closer to a global optimum. However, the low-capacity case showed that the task-based heuristics were universally weaker than the location-based heuristics. Location-based heuristics are very good at making efficient use of vehicles, because as soon as they are available, they are tasked again, doing something useful (i.e. policy-directed) wherever they happen to be. We can speculate that efficient use of vehicles is more important than allocating vehicles in line with global needs.

There is less distinction between the performance of policies that prioritize by item type and those that prioritize by destination. In this problem, prioritizing by destination led to slightly superior performance, but this is probably because there was more "room for

		Location-based						Task-based										Opt
		Item		Destination		Other		Item Prioritization				Destination Prioritization				Crossed		
		Lexic	Wtd	Lexic	Wtd	Short	Feed	Base	Wtd	TotNeed	FCFS	Base	Wtd	TotNeed	FCFS			
Run #:		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Opt
Gap:	U1: Total Delivered	22.9	25.9	23.8	27.3	30.3	33.0	40.1	40.1	37.2	37.0	35.8	35.5	38.9	32.2	49.8	42.8	0.0
	U2: Item Prioritization	2.2	0.3	0.7	1.0	2.0	1.0	3.4	1.6	1.7	2.3	0.7	0.7	0.7	0.7	4.5	0.7	0.0
	U3: Destination Prioritization	24.5	28.3	19.9	36.1	40.5	34.3	42.2	37.6	27.8	19.3	36.4	27.0	41.3	19.8	28.3	43.0	0.2
	U4: Speed	31.8	26.3	22.8	15.0	10.0	19.4	58.4	57.6	56.5	60.5	53.1	40.6	51.8	56.2	62.9	52.3	1.5
	U: Sum of utilities	79.8	79.2	65.6	77.6	81.1	86.0	142.5	135.3	121.6	117.4	124.1	102.0	131.1	107.0	143.9	136.9	-
Metric Tons Delivered:	Total delivered	1,858.1	1,834.9	1,851.2	1,824.3	1,800.6	1,780.0	1,724.7	1,724.6	1,747.5	1,748.8	1,758.5	1,760.6	1,733.8	1,786.2	1,649.6	1,703.8	2,040.0
	Priority 1 Items	1,277.6	1,227.3	1,212.7	1,193.3	1,187.8	1,149.1	1,153.0	1,120.6	1,147.4	1,259.0	1,131.0	1,187.5	1,106.3	1,227.6	1,202.3	1,106.3	1,366.7
	Priority 2 Items	18.8	26.7	26.7	26.7	18.8	26.7	26.2	26.7	26.2	26.2	26.7	26.7	26.7	26.7	26.2	26.7	26.7
	Priority 3 Items	113.9	133.2	126.9	120.6	118.1	120.6	75.6	109.9	109.0	97.4	126.9	126.9	126.9	126.5	55.0	126.9	139.6
	Priority 4 Items	440.0	440.0	440.0	440.0	440.0	440.0	438.0	435.5	430.1	332.8	428.5	374.0	428.5	360.0	332.8	398.5	460.0
	Priority 1 Destinations	1,128.1	1,105.4	1,155.9	1,058.3	1,031.8	1,069.2	1,021.4	1,049.3	1,108.2	1,177.3	1,056.7	1,113.1	1,026.7	1,156.7	1,156.2	1,016.7	1,275.1
	Priority 2 Destinations	729.9	729.5	695.3	765.9	768.9	710.9	703.3	675.3	639.3	571.4	701.7	647.4	707.1	629.5	493.4	687.1	764.9







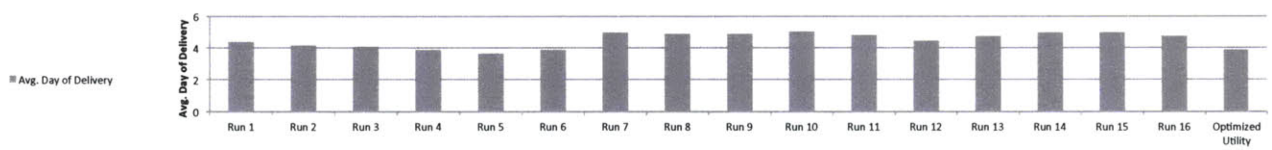


Figure 4-15: Results for "Snowland" problem: high-capacity case

improvement" by paying attention to destination priority than by paying attention to item type priority; the mix of item types was already naturally close to the ideal mix. This is not a particularly satisfying conclusion, because it suggests that there is no clear conclusion to be drawn: the performance of various policies depends on the specific structure of the problem. We will come back to this observation later.

Comparing the heuristics to the optimal solution highlights the superiority of the optimized solution: it performs better than the heuristics on nearly all dimensions. It is tempting to conclude that optimization is simply a superior method and, despite the complexity, should be adopted in humanitarian transportation planning. Still, it is worth translating those gaps into real terms. In the more interesting low-capacity case, the optimized solution delivers on average 256 MT more than the heuristics, but the best heuristic comes within 81 MT of the optimized total deliveries. In this problem, that 81 MT might be 2,000 rolls of plastic sheeting for shelter, or 10,000 hygiene kits and 2,500 squatting plates for latrines, or 15,000 high energy biscuits. Comparing in terms of utility, the best heuristic solutions achieve only 57 utils less than the optimized solution, equivalent to delivering 10% more of the requested cargo, or delivering 1.2 days faster on average. It is worth the effort to develop better planning methods even to reduce small optimality gaps, when delivering essential cargo. On the other hand, the heuristics are not as bad as might be expected from short-sighted, greedy-style search. The prioritization policies are reasonably good at steering the greedy search toward good solutions, as measured by the utility function. Combining policies and tweaking the heuristics is likely to yield an even better set of policies that could be implemented by human planners or simple decision support tools. Both of these solutions are more likely to be adopted and have an immediate impact on practice.

4.6.4 Discussion

By studying the three problems discussed in this section – two stylized problems and a realistic problem – we hoped to gain intuition about the strengths and weaknesses of the heuristics and policies described in this chapter, in comparison with a mixed-integer optimization approach. Across cases, there is no clear best-performing heuristic or policy. The optimized solution is always superior (as we would expect), but the gaps between optimized solutions and those of the better-performing heuristics are not very large in any of the cases.

Across three problems and several instances of each, no clear "winner" emerged among the combinations of heuristics and policies tested in this chapter. We attempted to distinguish between prioritizing by item type or destination, between policies that make lexicographic prioritization choices or attempt to send a weighted mix, and between the structure of location-based and task-based decision processes. On the latter point, the low-capacity Snowland problem suggests that location-based processes are superior, probably because they make more efficient use of available vehicles than do the task-based processes, despite the more global perspective inherent in the task-based processes. Distinguishing between item type and destination prioritization policies is difficult: destination policies performed better in the Snowland case, but the item type policies performed better in the easy stylized problems. Probably, destination prioritization is more important in complex networks in which there are many choices about where to send trucks, and in which later ability to deliver depends on earlier choices. The stylized networks were too simple to exhibit this dynamic. In reality, though, the distinction is not worth making, and policies should utilize *both* prioritization criteria. It was difficult for human decision-makers to manage so many kinds of criteria in their decision-making, but decision support tools could assist. Between lexicographic and weighted mix modes of prioritization, the results conflict. While the lexicographic policies actually performed somewhat better in the Snowland case than their weighted mix counterparts, we speculate that this is an artifact of the difficulty in delivering higher-priority cargo in this case: taking any capacity from high-priority locations decreased the utility of the resulting solution. In general, though, and as we saw in the stylized problems, a weighted mix mode of prioritization seems more likely to lead to a better mix of priority deliveries in the final solution.

The results thus suggest that location-based heuristics are superior to task-based heuristics, that both item type and destination prioritization are important, and that weighted mix prioritization is slightly better than lexicographic. However, when one policy performed better than another, the difference could often be traced to some specific element of the problem structure. For example, in the airbridge problem, prioritizing by destination was unimportant because each mode of transport could only serve one destination. In the difficult cases, item prioritization policies delivered less overall cargo because high-priority items took up more space. Feed policies may only be useful when there is a very important bottleneck in the network. While these issues are specific to the small set of problems discussed in this chapter, such issues are part of reality as well. It would be best to find a heuristic that is robust to these types of issues, performing relatively well in a variety of situations. The evidence from this study suggests that building on these "human heuristics" in ways that combine policies of prioritization by item type and destination might lead to more robust heuristics.

The success of each of the policies is driven in part by the utility function with which we evaluated them. For example, the item type prioritization policy performed poorly in the difficult cases because the utility function valued the total amount of cargo delivered more than the delivery of high-priority items. We developed this utility function based on the preferences of experienced humanitarian logisticians, as described in Chapter 3, but it is worth remembering that different conclusions might be reached with different utility functions. In addition, the utility function measures performance in the aggregate, looking at the total amount of cargo delivered to priority-1 locations, for example. However, it does not measure how evenly deliveries were spread among all the priority-1 locations, so high-utility solutions could be obtained by greedily serving the closest places first. Future work should look at modifications to this utility function to counteract this effect.

We have thus far said little about speed of delivery and operations cost; most of the discussion has centered around the other three elements of the utility function (total cargo delivered, prioritization by item type, and prioritization by destination). The operations cost is nearly constant across all solutions, so it does not contribute to differences in performance. Cost is the same across solutions because all solutions use all available vehicles each day. The speed of delivery, on the other hand, does vary across solutions. In some cases, those that deliver very slowly also seem to be weak in other areas. However, speed is an important contributor to the success of the shortest-path heuristic, which delivers more and faster by greedily serving closer destinations first. Still, the three other components of the utility function seem to drive performance more strongly, especially the total amount of cargo delivered.

We also tested a decision-making heuristic outside of the formal models developed in this chapter: the use of an airbridge. Teams decided up-front how much cargo to send via an airbridge, usually allocating half their helicopter capacity for the first few days until cargo could arrive by other modes. Both the stylized airbridge cases and the realistic Snowland cases showed they need to pay more attention to this decision.

A general weakness of the heuristics, compared with the optimization model, is their inability to make trade-offs between the multiple goals in our utility function. In the difficult cases of the stylized problems, for example, the optimized solution sacrifices deliveries of priority-1 items in order to deliver more cargo overall. In both the Snowland problem and the stylized airbridge problem, the optimized solution moves capacity from one set of arcs to another (the airbridge), sacrificing one set of deliveries to gain capacity in another part of the network. The optimizer is able to "see" that the trade-off is worth it in terms of utility. The heuristics, on the other hand, can only work toward the goals defined by its prioritization policies.

Nevertheless, despite this key weakness, the heuristics are able to create plans that deliver a reasonable amount of cargo. In the Snowland low-capacity case, for example, they deliver on average 75% of the cargo delivered by the optimizer, and the best heuristic delivers 95% of that delivered by the optimizer. This performance is surprisingly good, considering the simplicity of the greedy search strategies employed by humans.

4.7 Conclusions

In this chapter, we attempted to understand the effectiveness of the human decision-making approaches observed in our ethnographic study of humanitarian transportation planners. We created heuristics based on their approaches, modeling two decision processes (location-based and task-based), along with various combinations of policies or decision rules dictating how to prioritize the routing and loading of vehicles. In testing these heuristics against an optimization approach, we found that optimized solutions generally perform better (as we expected). However, many of the heuristics find solutions that deliver only a little less cargo, especially those that utilize the location-based decision process. No overwhelming weaknesses were identified in the policies; instead, we found that policies prioritizing by item type or destination both work well in different cases, suggesting that some combination of these policies might be more robust across problems.

The goal in evaluating heuristics was to find ways to improve transportation planning in practice, so we should consider how these insights from the model world translate to reality. The heuristic models developed in this chapter were not intended to realistically represent the disaster situation nor human behavior. We modeled a deterministic, unchanging transportation planning problem (since our field study showed humans considering such a problem), exploring the performance of heuristics in the absence of the dynamics and uncertainty that characterize disaster response. In a more realistic problem setting, re-planning would be required to account for changes in the situation. We modeled pure versions of decision processes observed within the messy complexity of human behavior, in order to understand the effectiveness of their processes and policies. In reality, humans generally perform worse than pure implementations of their own decision processes (Bowman, 1963), in part because humans do not implement them consistently. The same is true in this scenario: the teams in our ethnographic study came up with much worse plans than those created by the heuristics developed in this chapter (we could not evaluate their solutions against one another, but we did obtain some data on their solutions which support this statement). The teams probably struggled not only with implementing a decision process, but with creating it, determining priorities, managing information, and simply understanding the problem.

This chapter has shown that human decision processes can be reasonably effective if implemented consistently. Therefore, training that reinforces planners' existing intuitive approaches could help to improve transportation planning in practice, especially if combined with the development of more robust decision policies. On the other hand, a clear weakness of the model heuristics was their lack of ability to make trade-offs between goals. Policies direct search 'blindly' toward one goal (or multiple goals) without, for example, considering the sacrifice of priority items to reach priority destinations. In problems where such trade-offs enable much better solutions, heuristics show a serious weakness.

Improving transportation planning in practice may require a balance between strategies that find the very best plans, such as optimization, and those that are practical to implement in the field. Optimization approaches can be complex to set up and to implement, and re-deploying such tools in each emergency might be cumbersome. Moreover, it is unclear that optimized solutions continue to perform well as the environment (or problem) changes in an emergency response (this question is left to future research). Our results show that optimized solutions do perform better than heuristics based on human behavior, so there may be justification for creating optimization-based tools.

On the other hand, our results show that planners' intuitive approaches also generate pretty good transportation plans, when implemented consistently. Their success suggests a more straightforward and easily implemented approach to improving transportation planning in practice. Training and decision support tools could be developed to reinforce planners' intuitive approaches, help them to manage information, and suggest policies or rules to assist in each decision. The location-based decision process, which was more effective than the task-based process, would make it easier to implement such a system, because it requires no central planner. The same process and decision rules could be taught or provided to planners in each city within a large transportation network. Future work should seek the best set of decision rules, either robust across problems or specific to easily identifiable features of problems, to be used in such a system. We are currently building on this chapter's results to develop tools and systems for the Logistics Cluster, and we hope that our work will enable better transportation planning in future emergencies.

References

Akkihal, A. (2006). Inventory pre-positioning for humanitarian operations. Master's thesis, MIT.

Altay, N. and Green, W. G. I. (2006). Or/ms research in disaster operations management. *European Journal of Operational Research*, 175:475–493.

Apte, A. (2009). Humanitarian logistics: A new field of research and action. In *Foundations and Trends in Technology, Information and OM*, volume 3, pages 1–100.

Armacost, A. P., Barnhart, C., and Ware, K. A. (2002). Composite variable formulations for express shipment service network design. *Transportation Science*, 36(1):1–20.

Balcik, B., Beamon, B. M., and Smilowitz, K. (2008). Last mile distribution in humanitarian relief. *Journal of Intelligent Transportation Systems*, 12(2):51–63.

Barbarosoglu, G., Ozdamar, L., and Cevik, A. (2002). An interactive approach for hierarchical analysis of helicopter logistics in disaster relief operations. *European Journal of Operational Research*, 140:118–133.

Beamon, B. M. (2004). Humanitarian relief chains: issues and challenges. In *34th International Conference on Computers and Industrial Engineering, San Francisco, CA, November 14-16.*

Beamon, B. M. and Kotleba, S. A. (2006a). Inventory management support systems for emergency humanitarian relief operations in south sudan. *International Journal of Logistics Management*, 17(2):187–212.

Beamon, B. M. and Kotleba, S. A. (2006b). Inventory modelling for complex emergencies in humanitarian relief operations. *International Journal of Logistics: Research and Applications*, 9(1):1–18.

Ben-Zion, U., Cohen, Y., Peled, R., and Shavit, T. (2008). Decision-making and the newsvendor problem: an experimental study. *Journal of the Operational Research Society*, 59(9):1281–1287.

Bendoly, E., Donohue, K., and Schultz, K. L. (2006). Behavior in operations management: Assessing recent findings and revisiting old assumptions. *Journal of Operations Management*, 24:737–752.

Bolton, G. E. and Katok, E. (2008). Learning by doing in the newsvendor problem: a laboratory investigation of the role of experience and feedback. *Manufacturing and Service Operations Management*, 10(3):519–538.

Bowman, E. H. (1963). Consistence and optimality in managerial decision making. *Management Science*, 9(2):310–321.

Campbell, A., Vandenbussche, D., and Hermann, W. (2008). Routing for relief efforts. *Transportation Science*, 42(2):127–145.

Caunhye, A., Nie, X., and Pokharel, S. (2012). Optimization models in emergency logistics: A literature review. *Socio-Economic Planning Sciences*, 46:4–13.

Chaikin, D. (2003). Towards improved logistics: challenges and questions for logisticians and managers. *Forced Migration Review*, 18:10.

Chomilier, B., Samii, R., and Van Wassenhove, L. N. (2003). The central role of supply chain management at ifrc. *Forced Migration Review*, 18:15–16.

Clark, A., Culkin, B., and Worcestershire, U. (2007). A network transshipment model for planning humanitarian relief operations after a natural disaster. In *22nd European Conference on Operational Research*, volume 44, pages 1–34.

Crainic, T. G. (2000). Service network design in freight transportation. *European Journal of Operational Research*, 122:272–288.

Crainic, T. G. and Kim, K. H. (2007). Intermodal transportation. In Barnhart, C. and Laporte, G., editors, *Handbooks in operational research and management science: Transportation, volume 14*, chapter 8, pages 467–537. North-Holland.

Croson, R. and Donohue, K. (2002). Experimental economics and supply-chain management. *Interfaces*, 32(5):74–82.

De Angelis, V., Mecoli, M., Nikoi, C., and Storchi, G. (2007). Multiperiod integrated routing and scheduling of world food programme cargo planes in angola. *Computers and Operations Research*, 34:1601–1615.

de la Torre, L., Dolinskaya, I. S., and Smilowitz, K. (2011). Disaster relief routing: Integrating research and practice. *Socio-Economic Planning Sciences*, (in press).

Doerner, K., Focke, A., and Gutjahr, W. (2007). Multicriteria tour planning for mobile healthcare facilities in a developing country. *European Journal of Operational Research*, 179(3):1078–1096.

Duran, S., Gutierrez, M., and Keskinocak, P. (2011). Pre-positioning of emergency items for care international. *Interfaces*, 41(3):223–237.

Ergun, O., Karakus, G., Keskinocak, P., Swann, J., and Villarreal, M. (2010). Operations research to improve disaster supply chain management. *Encyclopedia of Operations Research and Management Science*, 6:3802–3810.

Farvolden, J. M. and Powell, W. B. (1994). Subgradient methods for the service network design problem. *Transportation Science*, 28(3):256–272.

Fenton, G. (2003). Coordination in the great lakes. *Forced Migration Review*, 18:23–24.

Grünert, T. and Sebastian, H.-J. (2000). Planning models for long-haul operations of postal and express shipment companies. *European Journal of Operational Research*, 122:289–309.

Gustavsson, L. (2003). Humanitarian logistics: context and challenges. *Forced Migration Review*, 18:6–8.

Haghani, A. and Oh, S.-C. (1996). Formulation and solution of a multi-commodity, multi-modal network flow model for disaster relief operations. *Transportation Research Part A: Policy and Practice*, 30(3):231–250.

Hastie, R. and Dawes, R. (2001). *Rational choice in an uncertain world: The psychology of judgment and decision making.* Sage Publications, Inc.

Hill, A. V. (1982). An experimental comparison of human schedulers and heuristic algorithms for the traveling salesman problem. *Journal of Operations Management*, 2(4):215–223.

Holguin-Veras, J., Perez, N., Jaller, M., Destro, L., and Wachtendorf, T. (2010). On the need to reformulate humanitarian logistics modeling: Deprivation costs, and material convergence (draft version). In *Proceedings of the 89th Transportation Research Board Annual Meeting.*

Holguin-Veras, J., Perez, N., Ukkusuri, S., Wachtendor, T., and Brown, B. (2007). Emergency logistics issues impacting the response to katrina: a synthesis and preliminary suggestions for improvement. In *to be presented at the Transportation Research Board Conference, Washington, DC.*

Huang, M., Smilowitz, K., and Balcik, B. (2012). Models for relief routing: Equity, efficiency and efficacy. *Transportation Research Part E: Logistics and Transportation Review*, 48:2 – 18.

Jaller, M. and Holguin-Veras, J. (2011). Locating points of distribution in disasters with social costs considerations. In *Proceedings of the 90th Transportation Research Board Annual Meeting (CD-ROM).*

Kaatrud, D. B., Samii, R., and Van Wassenhove, L. N. (2003). Un joint logistics centre: a coordinated response to common humanitarian logistics concerns. *Forced Migration Review*, 18:11–14.

Kahneman, D., Slovic, P., and Tversky, A., editors (1982). *Judgment under uncertainty: heuristics and biases.* Cambridge University Press.

Kalkanci, B., Chen, K.-Y., and Erhun, F. (2011). Contract complexity and performance under asymmetric demand information: an experimental evaluation. *Management Science*, 57(4):689–704.

Kim, D., Barnhart, C., Ware, K., and Reinhardt, G. (1999). Multimodal express package delivery: A service network design application. *Transportation Science*, 33(4):391–407.

Kovacs, G. and Spens, K. M. (2007). Humanitarian logistics in disaster relief operations. *International Journal of Physical Distribution and Logistics Management*, 37(2):99–114.

Krolak, P., Felts, W., and Nelson, J. (1972). A man-machine approach toward solving the generalized truck-dispatching problem. *Transportation Science*, 6(2):149–170.

Larson, R. C., Metzger, M. D., and Cahn, M. F. (2006). Responding to emergencies: lessons learned and the need for analysis. *Interfaces*, 36(6):486–501.

Lin, Y., Batta, R., Rogerson, P., Blatt, A., and Flanigan, M. (2011). A logistics model for emergency supply of critical items in the aftermath of a disaster. *Socio-Economic Planning Sciences*, 45(4):132–145.

Long, D. C. and Wood, D. F. (1995). The logistics of famine relief. *Journal of Business Logistics*, 16(1):213–229.

MacGregor, J. N. and Chu, Y. (2011). Human performance on the traveling salesman and related problems: A review. *The Journal of Problem Solving*, 3(2).

Mete, H. and Zabinsky, Z. (2010). Stochastic optimization of medical supply location and distribution in disaster management. *International Journal of Production Economics*, 126(1):76–84.

Molinaro, P. and Blanchet, S. (2003). Logistics under pressure: Unicef's back to school programme in afghanistan. *Forced Migration Review*, 18:20–22.

Moxnes, E. (1998). Not only the tragedy of the commons: misperceptions of bioeconomics. *Management Science*, 44(9):1234–1248.

Ozdamar, L., Ekinci, E., and Kucukyazici, B. (2004). Emergency logistics planning in natural disasters. *Annals of Operations Research*, 129(1-4):217–245.

Pedraza Martinez, A. J., Hasija, S., and Van Wassenhove, L. N. (2010). An operational mechanism design for fleet management coordination in humanitarian operations. INSEAD Working Paper 2010/87/TOM/ISIC.

Perez, N., Holguin-Veras, J., Mitchell, J. E., and Sharkey, T. C. (2010). Integrated vehicle routing problem with explicit consideration of social costs in humanitarian logistics. Submitted August 2010. http://transp.rpi.edu/ HUM-LOG/Doc/IRP.pdf.

Powell, W. and Sheffi, Y. (1983). The load planning problem of motor carriers: Problem description and a proposed solution approach. *Transportation Research Part A: General*, 17(6):471–480.

Powell, W. B. (1986). A local improvement heuristic for the design of less-than-truckload motor carrier networks. *Transportation Science*, 20(4):246–257.

Powell, W. B. and Sheffi, Y. (1989). Design and implementation of an interactive optimization system for network design in the motor carrier industry. *Operations Research*, 37(1):12–29.

Rawls, C. and Turnquist, M. (2010). Pre-positioning of emergency supplies for disaster response. *Transportation Research Part B: Methodological*, 44(4):521–534.

Robinson, Jr., E. P. and Swink, M. L. (1995). A comparative model of facility network design methodologies. *Journal of Operations Management*, 13(3):169–181.

Russell, T. (2005). The humanitarian relief supply chain: analysis of the 2004 south east asia earthquake and tsunami. Master's thesis, MIT.

Salmerón, J. and Apte, A. (2010). Stochastic optimization for natural disaster asset prepositioning. *Production and Operations Management*, 19(5):561–574.

Schweitzer, M. E. and Cachon, G. P. (2000). Decision bias in the newsvendor problem with a known demand distribution: experimental evidence. *Management Science*, 46(3):404–420.

Shen, Z., Dessouky, M., and Ordóñez, F. (2009a). A two-stage vehicle routing model for large-scale bioterrorism emergencies. *Networks*, 54(4):255–269.

Shen, Z., Ordóñez, F., and Dessouky, M. (2009b). The stochastic vehicle routing problem for minimum unmet demand. In *Optimization and Logistics Challenges in the Enterprise*, pages 349–371. Springer.

Simon, H. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, 69(1):99–118.

Sterman, J. D. (1989). Modeling managerial behavior: misperceptions of feedback in a dynamic decision making experiment. *Management Science*, 35(3):321–339.

Swink, M. L. and Robinson, Jr., E. P. (1997). Complexity factors and intuition-based methods for facility network design. *Decision Sciences*, 28(3):583–614.

Tatham, P. and Kovacs, G. (2007). The humanitarian supply network in rapid onset disasters. In *Proceedings of NOFOMA, 7-8 June, Reykjavik, Iceland.*

Taylor, P. B. and Iwanek, R. (1980). Intuitive versus optimal solutions to some problems in distribution. *Omega*, 8(2):183–192.

Thomas, A. and Kopczak, L. (2006). Life-saving supply chains: challenges and the path forward. In Lee, H. L. and Lee, C.-Y., editors, *Building supply chain excellence in emerging economies*, chapter 4. Springer.

Tomasini, R. M. and Van Wassenhove, L. N. (2009). *Humanitarian Logistics.* INSEAD Business Press.

Tversky, A. and Kahneman, D. (1974). Judgment under uncertainty: heuristics and biases. *Science*, 185(4157):1124–1131.

Tzeng, G., Cheng, H., and Huang, T. (2007). Multi-objective optimal planning for designing relief delivery systems. *Transportation Research Part E: Logistics and Transportation Review*, 43(6):673–686.

Ukkusuri, S. V. and Yushimito, W. F. (2008). A location-routing approach for the humanitarian pre-positioning problem. *Transportation Research Record*, 2089:18–25.

Van Hentenryck, P., Bent, R., and Coffrin, C. (2010). Strategic planning for disaster recovery with stochastic last mile distribution. In Lodi, A., Milano, M., and Toth, P., editors, *Integration of AI and OR Techniques in Constraint Programming for Combinatorial Optimization Problems*, volume 6140 of *Lecture Notes in Computer Science*, pages 318–333. Springer Berlin Heidelberg.

Van Wassenhove, L. N. (2006). Humanitarian aid logistics: supply chain management in high gear. *Journal of the Operational Research Society*, 57:475–489.

Wieberneit, N. (2008). Service network design for freight transportation: a review. *OR Spectrum*, 30:77–112.

Yi, W. and Kumar, A. (2007). Ant colony optimization for disaster relief operations. *Transportation Research Part E: Logistics and Transportation Review*, 43(6):660–672.

Yi, W. and Ozdamar, L. (2007). A dynamic logistics coordination model for evacuation and support in disaster response activities. *European Journal of Operational Research*, 179(3):1177–1193.

Zhu, J., Huang, J., Liu, D., and Han, J. (2008). Resources allocation problem for local reserve depots in disaster management based on scenario analysis. In *The 7th International Symposium on Operations Research and its Applications. Lijiang, China*, pages 395–407.

Chapter 5

Conclusions

One of the goals of this thesis was to develop ways to improve transportation planning in humanitarian supply chains. A common route to improving supply chain planning is to develop mathematical models that solve for the best possible plans, but the humanitarian context presents particular challenges for this approach. In emergency response, information is scarce and unreliable, the situation changes rapidly, and there are multiple, conflicting goals. In short, the problem formulation itself is not clear, making it challenging to build models that create useful plans. Nevertheless, modeling approaches may enable the generation of better transportation plans, given appropriate formulations. This thesis aimed to understand how human decision-makers formulated and solved humanitarian planning problems, and to identify the strengths and weaknesses of their problem-solving approaches. With this knowledge, improvements to transportation planning could be developed based on the best mix of modeling and human approaches.

Each of the three main chapters in the thesis looked at human approaches to transportation planning from a different perspective. Chapter 2 described an ethnographic study of humanitarian transportation planners in a realistic field setting, which focused on the human processes of sensemaking and solving. I found that sensemaking leads to an understanding of the problem similar to a formulation, and that solving resembles a greedy search process guided by decision rules or policies that direct the search toward goals. Chapter 3 used a conjoint analysis survey to quantify the goals of human planners in a utility function that can be used as a performance measure or objective function in mathematical models. The results show that humanitarian logisticians are most concerned with delivering as much cargo as possible and prioritizing deliveries of priority items to priority destinations. Chapter 4 used heuristics and optimization to investigate the effectiveness of the human decision processes found in Chapter 2, with respect to the performance measure developed in Chapter 3. The results show that optimization approaches lead to superior solutions, in part because they can balance multiple objectives simultaneously, but that human decision processes could be nearly as good if implemented consistently with appropriate decision rules. Each of these chapters contains its own discussion of conclusions, which are not repeated here.

Instead, I step back to consider the implications of the three chapters taken together. First, what are the strengths and weaknesses of human planning approaches? One human strength is their ability to work without a pre-defined problem formulation, making sense of the problem as they begin to solve it. This process of sensemaking may also be a weakness: the process may not always lead to the best formulation in which to search for transportation plans. Another weakness is the greedy search process, which does not necessarily find the best possible transportation plan. However, the location-based decision process described in Chapter 2 (in which dispatch decisions are made at each location independently), if implemented consistently and with good decision rules, can achieve good transportation plans in many types of problems.

These findings suggest that transportation planning could be improved through two possible routes. Developing optimization-based decision support tools would lead to the best possible plans, if models could be formulated that capture enough of the context to create plans that work well in reality. A simpler approach is to develop training and basic decision support tools that enable humans to consistently follow their own intuitive heuristics. (An example of such an approach was provided at the conclusion of Chapter 4; it involves implementing consistent decision rules across multiple locations making independent dispatch decisions.) The former approach might lead to the best transportation plans, but the latter approach is faster and easier to implement, so both are worth pursuing.

The thesis also suggests better ways to model humanitarian transportation planning problems. What should be included in a problem formulation, if a model is to generate plans that will be useful in reality? My key finding is that the objective is essential. Humanitarian logisticians spent much of their time deciding how to prioritize aid deliveries, essentially sorting out the right objectives. Models that simply maximize deliveries or minimize cost do not create plans in line with the goals of humanitarian logisticians. By maximizing a sum of utilities over several objectives, plans can be created that meet the most important objectives of humanitarian supply chains. This type of model is an important step toward optimization-based decision support tools that could be useful in the field.

Beyond improvements to humanitarian transportation planning, this research contributes to the broader discussion about the benefits of rational and non-rational problem-solving. The decision-making literature includes many studies that show both the drawbacks and benefits of human reasoning shortcuts. Similarly, the field of operations management provides many examples of rational mathematical models that have led to improved performance, and many that have failed. There has been little investigation of the conditions in which nor the reasons why such models succeed or fail. This research begins to address this gap, by highlighting the differences in the ways humans and models perceive problems.

Humans perceive problems through sensemaking: exploring the problem to determine constraints, identifying challenges to address and tasks to complete, and teasing out objectives by confronting dilemmas. Model formulations are developed by humans, of course, but usually they are not the same people who operate the systems. If a model formulation is missing important elements of the problem, the solutions may not make sense to the system operators, and the model will fail to be useful in reality. This thesis found, for example, that the objective function was critically important. A model that minimized cost or maximized overall deliveries would not produce plans acceptable to humanitarian logisticians concerned with prioritization. Proving that a model finds optimal solutions, within its own "model world," is not sufficient to ensure its value in practice unless the formulation is consistent with reality. On the other hand, human problem solvers are myopic, considering one decision at a time without searching for trade-offs that might lead to much better transportation plans. Humans may fail to appreciate the potential of mathematical models if their solutions do not make sense within their understanding of the problem. In this thesis, for example, planners did not consider using helicopters as an airbridge after the first few days, because they saw airbridges only as a method of speeding up deliveries to locations that could not yet be reached by road. The optimization models used airbridges to increase capacity in later planning days, a solution the humanitarian logisticians might have rejected as wasting expensive helicopter capacity. Here, again, a model would fail to be useful, but in this case the failure may lie with the humans rather than the model.

These observations suggest that if human and modeling approaches do not align, it may be for one of two reasons: model formulations may be missing important elements, or models may find non-intuitive solutions. Both issues lead to model solutions that do not make sense to humans, and thus to 'rational' models that are not implemented in reality. Because of the different ways in which models and humans perceive and solve problems, what is 'rational' in one approach may seem irrational in the other. This thesis, by showing when and why human and modeling approaches conflict, suggests that there is not a clear division between rational and irrational problem-solving in ill-defined, real-world problems.

This thesis makes several contributions, many of which were detailed within the three main chapters and are not repeated here. Looking across chapters, the thesis adds to our understanding of how and how well humans solve operational problems in ill-defined, real-world environments, based on the extreme case of humanitarian transportation planning. It probes the usefulness of mathematical models in this challenging setting and shows when and why optimization models are and are not useful in this context. By investigating human and modeling approaches to transportation planning, this thesis paves the way for building better models and supporting better human decision-making. In support of better models, I developed a performance metric based on expert preferences over multiple objectives (and a process for creating additional such performance metrics for other situations), and I showed how its use as an objective function leads to better transportation plans. In support of better human decision-making, I evaluated the strengths and weaknesses of human reasoning approaches, and I identified specific areas where training and decision support could enable humans to make better decisions. The insights from this thesis not only enable better models and better human decision-making, but also suggest ways to combine the strengths of models and humans to improve transportation planning for humanitarian response.

Future Work

This thesis explored the strengths and weaknesses of human decision-making in humanitarian transportation planning. Future work should build upon this knowledge to support better human decision-making and build better models for humanitarian transportation planning. In addition, future work should investigate whether, in other ill-defined, complex problems, humans understand and solve problems the same way they do in humanitarian transportation planning.The following paragraphs provide specific suggestions for future work that builds on each of the main chapters of this thesis.

In Chapter 2, I developed models of human sensemaking and solving processes based on ten similar cases of humanitarian transportation planning. Future work should explore how these decision-making processes vary in different scenarios and different kinds of problems. One of the findings was that people simplify problems by understanding them in terms of intuitive tasks. Future studies could investigate whether transportation planners at commercial logistics companies also think of their work in such terms, to determine whether the models transfer to similar problem structures in different contexts. To investigate whether the models are applicable to different problem structures, future studies could investigate sensemaking and solving in engineering design problems or software debugging, for example. Qualitative studies could be supplemented by experimental studies in more controlled environments, to test the effects found in the qualitative studies, or to investigate whether specific interventions can alleviate some of the weaknesses in human decision-making processes.

A second direction for future work is to use the same methods employed in Chapter 1 to study another important dynamic in human decision-making. In humanitarian transportation planning, people focused on the problem's objectives, and their choice of objectives (and their implementation of objectives in decision policies) was the main driver in their decision-making process. I plan to study another problem in which I expect uncertainty, rather than objectives, to be the main driver of decision-making: a supply chain design problem in the humanitarian context. The investigation should show how humans incorporate uncertainty in their decisions.

In Chapter 3, I developed a performance measure for humanitarian aid delivery plans, based on the preferences of a group of expert humanitarian logisticians. This utility function represents a first step in developing empirically-based objectives for aid delivery models. The relatively small group of respondents represent the real decision-makers in the Logistics Cluster, so their utility functions are a good place to start, but future work could examine the preferences of decision-makers in other organizations, or people who fill roles other than logistics. In addition, utility functions could be developed for a wider variety of disaster scenarios. This work focused on the first week after a major emergency. Priorities would likely change as the urgency of life-saving aid diminished.

The survey in this thesis measured the utility of five objectives identified in the ethnographic study, but different sets of objectives could be explored. In particular, it would be useful to explore the relative utility of measures like the effectiveness, equity, and efficiency objectives utilized in previous humanitarian aid delivery models. Also, this survey showed that people prefer a weighted mix mode of prioritization, but did not explore what kind of mix people prefer the most. Future work should measure the utility of multiple mixes of deliveries in order to determine the right 'ideal point'. Finally, the same objective measurement methodology could be extended to other fields, looking at how people trade off objectives of profit and sustainability, for example.

In Chapter 4, I tested heuristics derived from human decision-making processes against an optimization model, on a small set of test problems. Drawing general conclusions is difficult in part because there is no set of standard test problems for aid delivery transportation planning. I plan to develop a set of test problems, including several stylized instances representing prototypical challenges in disaster response transportation and a few large instances based on real past disasters. It would also be useful to the heuristics and optimization models on problems that change over time, to determine how robust they are to specific kinds of disruptions and changes.

The next step is to develop decision support tools that build on the insights from this thesis. Future work should explore the best combination of decision policies, so that people can be trained to use them in dispatch decision-making. Additional kinds of training and decision support should be developed, spanning the range from optimization models that come up with the best plans to simple decision support tools that can be used easily in practice. Combining such tools with the strengths of human intuition should lead to better humanitarian aid delivery in future emergencies.

Appendix A

Conjoint analysis survey attributes and questions

A.1 Detailed Attribute Descriptions

In each survey question, the respondent can click the "[?]" to access a more detailed definition of each attribute-level. This appendix presents the text of these definitions.

1. Total Deliveries

    (a) Deliver 80% of cargo requested for this week

    It is impossible to deliver all of the cargo requested for delivery this week: only 80% of it is delivered by the end of the week.

    (b) Deliver 60% of cargo requested for this week

    It is impossible to deliver all of the cargo requested for delivery this week: only 60% of it is delivered by the end of the week.

    (c) Deliver 40% of cargo requested for this week

    It is impossible to deliver all of the cargo requested for delivery this week: only 40% of it is delivered by the end of the week.

2. Item Type Prioritization

    (a) Priority-1 (shelter) items first: Load vehicles with Priority-1 items before any other cargo.

Priority-1 (shelter) cargo is always loaded onto vehicles before Priority-2 cargo. Similarly, Priority-2 (health) is loaded before Priority-3 (water and sanitation), which is loaded before Priority-4 (food).

As a result, at the beginning of the response, most people will receive only Priority-1 (shelter) items, then a little later they will receive Priority-2 (health) items, then Priority-3 (water and sanitation) items, and finally Priority-4 (food) items.

(b) More high-priority items: Load vehicles with a mix of items, but more of the higher-priority types.

Each truck carries a mix of items, but more of those types with higher priorities. As long as there is enough of each item type available to be shipped, each truck is loaded with about:

- *50% Priority-1 (shelter) cargo*
- *30% Priority-2 (health) cargo*
- *15% Priority-3 (water and sanitation) cargo*
- *5% Priority-4 (food) cargo*

As a result, at the beginning of the response, most communities will receive some of the higher-priority item types, and little of the lower-priority item types.

(c) Even mix of items: Load vehicles with an even mix of item types.

Each vehicle is loaded with all four types of items, in equal proportions. As long as there is enough of each item type available to be shipped, each truck carries about:

- *25% Priority-1 (shelter) cargo*
- *25% Priority-2 (health) cargo*
- *25% Priority-3 (water and sanitation) cargo*
- *25% Priority-4 (food) cargo*

As a result, most communities will receive about the same amount of each type of item.

3. Location Prioritization

(a) Priority-1 locations first: Send vehicles to high-priority locations before serving other locations.

Cargo is sent to the high-priority locations before any other areas, until most of their needs have been served. As long as there is enough cargo for the high-priority areas, the available vehicles are divided as follows:

- *90% of the available vehicles serve high-priority locations*
- *10% of the available vehicles serve low-priority locations*

As a result, at the beginning of the response, high-priority locations will receive much more of their needed cargo, while low-priority locations will receive only a little.

(b) More high-priority locations: Send vehicles to all locations but send more to the high-priority locations.

More cargo goes to the high-priority locations than the other locations, until most of their needs have been served. As long as there is enough cargo for the high-priority areas, the available vehicles are divided as follows:

- *70% of the available vehicles serve high-priority locations*
- *30% of the available vehicles serve low-priority locations*

As a result, at the beginning of the response, high-priority locations will receive more of their needed cargo, while low-priority locations will receive less.

(c) Even mix of locations: Send vehicles to all locations in equal proportions.

Cargo is delivered to both high-priority and low-priority locations in equal proportions. As long as there is enough cargo for each location, the available vehicles are divided as follows:

- *50% of the available vehicles serve high-priority locations*
- *50% of the available vehicles serve low-priority locations*

As a result, at the beginning of the response, high-priority locations and low-priority locations receive about the same amount of cargo.

4. Speed of Delivery

(a) 1-3 days: Complete most deliveries in 1-3 days.

Most (80%) of the deliveries this week are completed by day 3. The rest (20%) take between 4 and 7 days to deliver.

(b) 2-6 days: Complete most deliveries in 2-6 days.

Most (90%) of the deliveries this week are completed in days 2 through 6. No deliveries are made on day 1, and only a few (10%) are made on day 7.

(c) 4-7 days: Complete most deliveries in 4-7 days.

Most (80%) of the deliveries this week are completed on days 4 through 7. Only a few (20%) are delivered in the first 3 days.

5. Cost

(a) $0.5 million cost for this week

The operation costs $0.5 million for the first week, as part of a three-month operation that costs $6 million.

(b) $2.0 million cost for this week

The operation costs $2.0 million for the first week, as part of a three-month operation that costs $24 million.

(c) $3.5 million cost for this week

The operation costs $3.5 million for the first week, as part of a three-month operation that costs $42 million.

A.2 Survey text and questions

In this appendix, the full text of the survey is presented. Table 3.3 presented a general overview of the survey structure. This appendix shows details on those questions not addressed in the paper. Text in *italics* is survey text.

Survey introduction

This survey will explore what aspects of aid delivery plans are most important to you. We will ask you to evaluate different aid delivery plans, based on several characteristics.

Here's an example. [Figure A-1 shows the example question.]

Example: Which plan do you prefer?

Total Deliveries	**Deliver 60%** of cargo requested for this week [?]	**Deliver 60%** of cargo requested for this week [?]
Item Type Prioritization	**Priority-1 (shelter) items first** Load vehicles with Priority-1 items before any other cargo [?]	**Even mix of items** Load vehicles with an even mix of item types [?]
Location Prioritization	**Priority-1 locations first** Send vehicles to high-priority locations before serving other locations [?]	**Even mix of locations** Send vehicles to all locations in equal proportions [?]
Speed of Delivery	**1-3 days** Complete most deliveries in 1-3 days [?]	**2-6 days** Complete most deliveries in 2-6 days [?]
Cost	**$3.5 million** cost for this week [?]	**$2.0 million** cost for this week [?]

[Show introduction]

Figure A-1: Example question; forms part of survey introduction.

All the plans in the survey describe the first week of a major emergency response. An earthquake occurred in a cold, mountainous area, destroying much of the existing infrastructure. We will ask you to evaluate aid delivery plans in this context. Normally, you might have more information about needs, but please try to evaluate the plans based on the limited information provided here.

Each plan you will see in this survey is described like those above, with the same set of five characteristics:

1. *Total Deliveries: Operational constraints limit the total amount of cargo that can be delivered. Each plan delivers some percentage of the total cargo that was requested for delivery during this week (the first week of the response).*

2. *Item Type Prioritization: There is an order of priority for item types. Some types of aid are more urgently needed than others, and the following prioritization applies:*

   - *Priority-1: shelter (because the climate is harsh and shelters are destroyed)*
   - *Priority-2: health (because existing systems are crippled)*

- *Priority-3: water and sanitation (because many systems are damaged and IDP camps are forming)*
- *Priority-4: food (because livelihoods are damaged and food stocks limited)*

The plan's item type prioritization determines which types of items reach communities this week, and the amount they receive of each type.

3. *Location Prioritization: There is an order of priority for locations. Some locations (or communities) are in more urgent need than others. There are two levels of prioritization:*

    - *Priority-1 locations have larger affected populations and more damage*
    - *Priority-2 locations have smaller affected populations but are still damaged*

    The plan's location prioritization determines which communities receive aid this week, and how much cargo they receive.

4. *Speed of Delivery: The speed of cargo deliveries can vary, depending on the challenges of access.*

5. *Cost: Each plan includes a cost, representing the cost of operations for the first week of the response. (You will be able to see this text at any time during the survey by clicking the "[Show introduction]" link below each question.)*

Which plan do you prefer? Please answer the example question above.
Click the right arrow to continue.

Build-your-own plan question

Figure A-2 shows the build-your-own plan question, which appears at the start of the survey.

Introduction to screening questions

Thank you for indicating your ideal aid delivery plan. Unfortunately, it is not always possible to implement the ideal plan. We want to understand the kinds of trade-offs you are willing to make, and which aspects of performance are most and least important to you.

To begin, please tell us about your ideal aid delivery plan. For each characteristic, select your preferred level.

If you want more information about each level of performance, click the [?] for a definition (it will open a new window)

Characteristic	Select Your Preference
Total Deliveries	**Deliver 40%** of cargo requested for this week [?] **Deliver 60%** of cargo requested for this week [?] **Deliver 80%** of cargo requested for this week [?]
Item Type Prioritization	**Priority-1 (shelter) items first**: Load vehicles with Priority-1 items before any other cargo. [?] **More high-priority items**: Load vehicles with a mix of items, but more of the higher-priority types. [?] **Even mix of items**: Load vehicles with an even mix of item types. [?]
Location Prioritization	**Priority-1 locations first**: Send vehicles to high-priority locations before serving other locations. [?] **More high-priority locations**: Send vehicles to all locations but send more to the high-priority locations. [?] **Even mix of locations**: Send vehicles to all locations in equal proportions. [?]
Speed of Delivery	**1-3 days**: Complete most deliveries in 1-3 days. [?] **2-6 days**: Complete most deliveries in 2-6 days. [?] **4-7 days**: Complete most deliveries in 4-7 days. [?]
Cost	**$0.5 million** cost for this week. [?] **$2.0 million** cost for this week. [?] **$3.5 million** cost for this week. [?]

[Show introduction]

Figure A-2: Build-your-own ideal plan question

You will be shown a series of different plans, and asked whether or not they are acceptable to you. Of course, better plans are preferable, but not always feasible. Here we want to know which plans are truly unacceptable – meaning that you would not defend these plans to donors.

Sample screening question

Figure A-3 shows a sample screening question. Six such questions were asked in the survey.

Here are a few aid delivery plans. For each one, indicate whether it is a possibility or not.

(1 of 6)

Total Deliveries	**Deliver 60%** of cargo requested for this week [?]	**Deliver 40%** of cargo requested for this week [?]	**Deliver 80%** of cargo requested for this week [?]
Item Type Prioritization	**Even mix of items**: Load vehicles with an even mix of item types. [?]	**Priority-1 (shelter) items first**: Load vehicles with Priority-1 items before any other cargo. [?]	**Priority-1 (shelter) items first**: Load vehicles with Priority-1 items before any other cargo. [?]
Location Prioritization	**Even mix of locations**: Send vehicles to all locations in equal proportions. [?]	**Priority-1 locations first**: Send vehicles to high-priority locations before serving other locations. [?]	**Priority-1 locations first**: Send vehicles to high-priority locations before serving other locations. [?]
Speed of Delivery	**1-3 days**: Complete most deliveries in 1-3 days. [?]	**2-6 days**: Complete most deliveries in 2-6 days. [?]	**4-7 days**: Complete most deliveries in 4-7 days. [?]
Cost	**$0.5 million** cost for this week. [?]	**$2.0 million** cost for this week. [?]	**$0.5 million** cost for this week. [?]
	◯ Acceptable ◯ Unacceptable: would not defend to a donor	◯ Acceptable ◯ Unacceptable: would not defend to a donor	◯ Acceptable ◯ Unacceptable: would not defend to a donor

[Show introduction]

Figure A-3: Sample screening question

Unacceptable rule question

Figure A-4 shows the "unacceptable rule" question. This question is shown to respondents after four screening question, only if their answer patterns suggest that they consider one or more attribute-levels unacceptable (and only those attribute-levels are provided as options within the question).

It seems that you have avoided plans with certain characteristics shown below. Would any of these be **totally unacceptable**? If so, mark the **one feature that is most unacceptable**.

◯ **$3.5 million** cost for this week. [?]

◯ **Priority-1 (shelter) items first**: Load vehicles with Priority-1 items before any other cargo. [?]

◯ **Deliver 40%** of cargo requested for this week [?]

◯ **4-7 days**: Complete most deliveries in 4-7 days. [?]

◯ **Priority-1 locations first**: Send vehicles to high-priority locations before serving other locations. [?]

◯ None of these is totally unacceptable.

Figure A-4: Unacceptable rule question

Introduction to choice tasks

Thank you for indicating what kinds of plans you would find acceptable. In the next section, we'll examine these acceptable plans again. In order to understand the trade-offs between different plans, we'll ask you to choose which of several plans you prefer.

Holdout questions

As an example of choice tasks, both holdout questions are provided here (they are identical to the other choice tasks). Figure A-5 shows the first holdout question and Figure A-6 the second.

Among these three, which is the best option? (We've grayed out any features that are the same, so you can just focus on the differences.)

Total Deliveries	**Deliver 60%** of cargo requested for this week [?]	**Deliver 60%** of cargo requested for this week [?]	**Deliver 40%** of cargo requested for this week [?]
Item Type Prioritization	**Priority-1 (shelter) items first** Load vehicles with Priority-1 items before any other cargo. [?]	**Even mix of items**. Load vehicles with an even mix of item types. [?]	**Priority-1 (shelter) items first**: Load vehicles with Priority-1 items before any other cargo. [?]
Location Prioritization	**Priority-1 locations first**. Send vehicles to high-priority locations before serving other locations. [?]	**Even mix of locations**. Send vehicles to all locations in equal proportions. [?]	**More high-priority locations**: Send vehicles to all locations but sends more to the high-priority locations. [?]
Speed of Delivery	**1-3 days**. Complete most deliveries in 1-3 days. [?]	**2-6 days**. Complete most deliveries in 2-6 days. [?]	**1-3 days**. Complete most deliveries in 1-3 days. [?]
Cost	**$3.5 million** cost for this week [?]	**$2.0 million** cost for this week [?]	**$3.5 million** cost for this week [?]
	○	○	○

[Show introduction]

Figure A-5: Holdout question 1 (the third choice is designed to be dominated by the other two)

Demographic questions

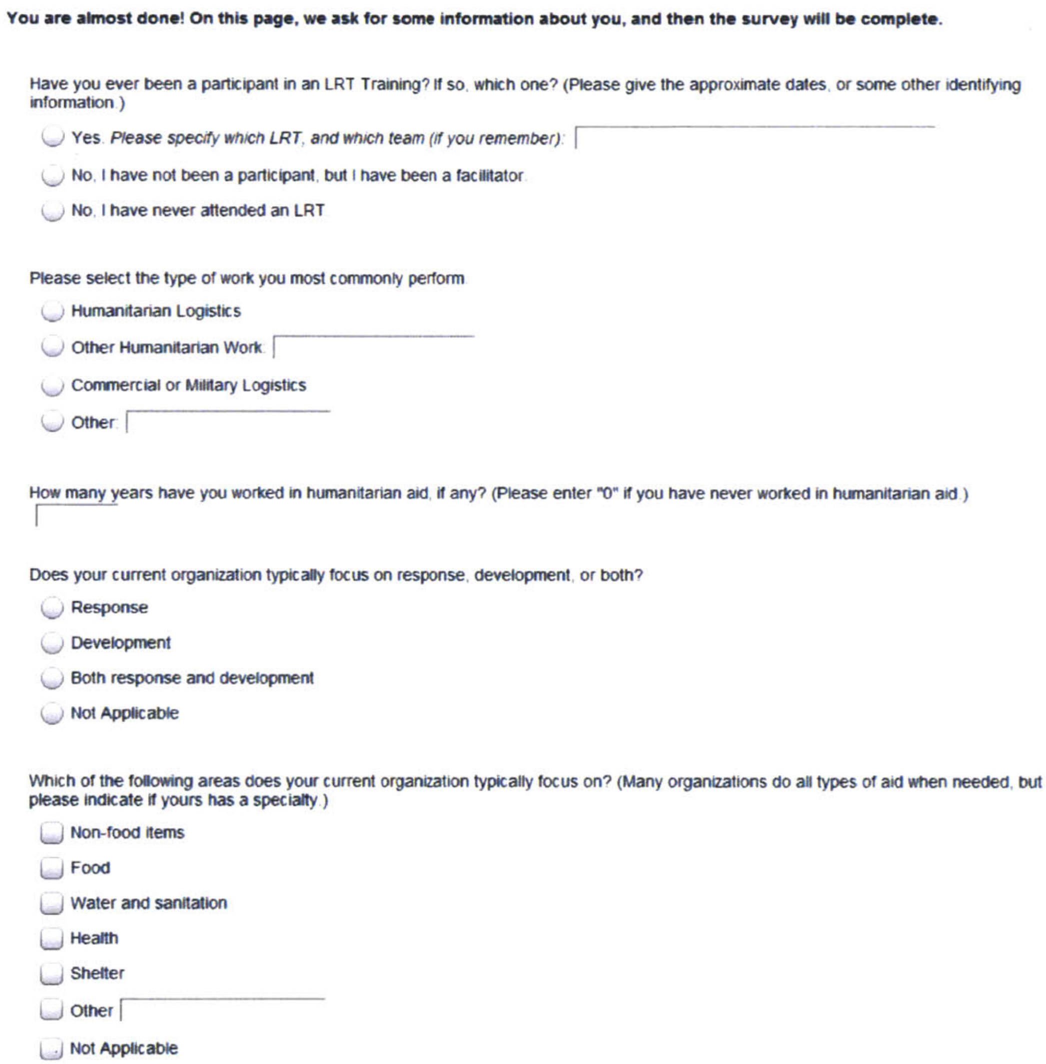
Figure A-7 shows the demographic questions asked at the end of the survey.

Among these three, which is the best option? (We've grayed out any features that are the same, so you can just focus on the differences.)

Total Deliveries	**Deliver 80%** of cargo requested for this week [?]	**Deliver 80%** of cargo requested for this week [?]	**Deliver 60%** of cargo requested for this week [?]
Item Type Prioritization	**More high-priority items**: Load vehicles with a mix of items, but more of the higher-priority types. [?]	**Even mix of items**: Load vehicles with an even mix of item types. [?]	**More high-priority items**: Load vehicles with a mix of items, but more of the higher-priority types. [?]
Location Prioritization	**More high-priority locations**: Send vehicles to all locations but send more to the high-priority locations. [?]	**Even mix of locations**: Send vehicles to all locations in equal proportions. [?]	**More high-priority locations**: Send vehicles to all locations but send more to the high-priority locations. [?]
Speed of Delivery	**2-6 days**: Complete most deliveries in 2-6 days. [?]	**1-3 days**: Complete most deliveries in 1-3 days. [?]	**1-3 days**: Complete most deliveries in 1-3 days. [?]
Cost	**$2.0 million** cost for this week. [?]	**$3.5 million** cost for this week. [?]	**$0.5 million** cost for this week. [?]
	○	○	○

[Show introduction]

Figure A-6: Holdout question 2



You are almost done! On this page, we ask for some information about you, and then the survey will be complete.

Have you ever been a participant in an LRT Training? If so, which one? (Please give the approximate dates, or some other identifying information.)

- Yes. *Please specify which LRT, and which team (if you remember):*
- No, I have not been a participant, but I have been a facilitator.
- No, I have never attended an LRT

Please select the type of work you most commonly perform

- Humanitarian Logistics
- Other Humanitarian Work:
- Commercial or Military Logistics
- Other:

How many years have you worked in humanitarian aid, if any? (Please enter "0" if you have never worked in humanitarian aid.)

Does your current organization typically focus on response, development, or both?

- Response
- Development
- Both response and development
- Not Applicable

Which of the following areas does your current organization typically focus on? (Many organizations do all types of aid when needed, but please indicate if yours has a specialty.)

- Non-food items
- Food
- Water and sanitation
- Health
- Shelter
- Other
- Not Applicable

Figure A-7: Demographics questions

(This page intentionally left blank.)

Appendix B

Pseudocode for heuristic models of human decision-making processes

Pseudocode is presented on the following pages for both the location-based and task-based heuristics, with all policy variations.

Pseudocode for location-based heuristic, page 1

- Begin location-based heuristic (*Heuristic_local)*
  - Set the paths for each CMR: (*SetCmrPaths)*
    - Sort the list of CMRs by:
      - item priority, if an item-priority policy
      - destination priority, if a destination-priority policy
      - item or destination priority if indicated in the airbridge policy
      - database order otherwise
    - For each CMR in the sorted list
      - If we have not yet sent enough cargo via shortest paths (as defined in the path policy), assign the CMR a shortest path; otherwise, assign a cheapest path. To do so,
      - Find the shortest path on a network whose arc lengths are either the distances or the estimated cost to travel, where the estimated cost is the distance multiplied by the average cost for all transport assets (defined in the database) that have access to this arc. (*GetDijkstraShortestPath)*
  - For each time step
    - Find the location, at the start of this period, of all CMRs and all vehicles
    - Remove from the list any CMRs that have already reached their destinations
    - Process all assignments *(ProcessAssignments)*
      - For each assignment defined by the user that is starting at this time period, and for each vehicle that needs to be assigned to it,
        - Find the vehicle, of the right type and not already assigned, whose last planned location is closest to any of the nodes in the assignment network
        - Assign the asset to the assignment: (*AssignAsset)*
          - Find the shortest path from the asset's last location to any of the nodes in the assignment
          - Route the asset (assign legs) along the shortest path
          - Assign the asset to only move within the subnetwork defined by the assignment (this does not affect the movement of the asset to the assignment network, defined previously)
        - If there are no vehicles of the right type that are not already assigned, do not assign any vehicles (and do not throw an error)
      - For each assignment defined by the user that is ending in this time period,
        - Unassign all vehicles: remove the restriction that they travel within the subnetwork defined by the assignment
    - If this time period is the beginning of a new day, reset the running (daily) total sent towards the feed-to-node, for the feed-to policy
    - For each node that has transport assets
      - *SendCargoForwardLocal*
        - For the destination weighted mix policy, find the total capacity of all assets at this node
        - For the organization policy, create two parallel arrays to note the organization names and how much cargo has been sent toward each, and initialize the latter to zeros
        - For each asset at this node,
          - For the destination weighted policy, update a running total of the cumulative capacity of the assets looped through with the capacity of this asset
          - If there is cargo waiting at this node,
          - Sort the list of loads by:
            - item type priority, if an item type policy
            - destination priority, if a destination policy
            - shortest remaining path for each load, if a shortest policy
            - bottleneck node, if a bottleneck policy
            - FIFO, if a FIFO or organization policy

Pseudocode for location-based heuristic, page 2

- For the destination-weighted policy, determine which destination-priority to look for with this asset:
  - Assign it to p1 destination cargo if the cumulative capacity (its capacity plus that of the others looped before this) falls in the range defined for p1 cargo, rounded up to the nearest multiple of this asset's capacity (to ensure that if there is only one asset, it goes to p1). The same for the other destination priorities.
- If the asset is not already routed beyond the current time (which may happen with assignments or re-routings),
- *SetVehicleDestinationBasedOnCmr*
  - For each waiting load,
    - If the asset can travel to the load's next stop (in its accessible or its assigned network),
    - If FIFO or shortest policy, select this load and exit the loop (this is the first in the sorted list that meets requirements)
    - If item type or destination or feed-to policy, add this CMR to a running total of the amount waiting to go to each adjacent node, of each priority, and select the load if it represents the largest such running total (but keep looping through the loads)
    - If organization policy, look up how much cargo has been sent for this load's organization, and select this load if it is the least-served organization found so far (but keep looping through the loads)
  - If FIFO, item type, destination-lexic, organization, or shortest, set the destination based on the selected load:
  - If a destination weighted policy, try to set the destination to that with the most cargo of the destination-priority this asset is looking for, but if there isn't any, set it to the one selected above:
  - If a feed-to policy, try to set the destination from a load going to the top-ranked feed-to-node that can be reached within the time horizon, but if we've already sent enough cargo to it, then the next-ranked feed-to-node, etc. If all feed-to-nodes are adequately served, then to the next load without a feed-to-node, and if there are none of those, then back to the top-ranked feed-to-node.
  - Set the vehicle destination by adding a leg to the asset going from the current node to the adjacent node that the selected load must reach next, and return the leg.
  - If there were no loads at this node, return Nothing.
- If the destination could not be set based on a CMR, i.e. if there were no loads waiting there, *SetDestinationForEmptyAsset* (defined below) and skip to next asset
- If item type weighted policy, call *LoadVehicleWithCmrs* repeatedly, looking for the right amount of each priority-level
- For all other policies,
- *LoadVehicleWithCmrs*
  - If item type weighted policy, initialize a running total of the volume or weight we want to fill this time
  - While there are loads waiting here,
    - If item type weighted policy, call *SelectWaitingCmrLoadWithNextNode* looking for the right type of priority (and if it isn't found, call it again later just looking for the highest-priority cargo)
    - If destination weighted policy, call *SelectWaitingCmrLoadWithNextNode* looking for the right type of destination-priority (and if it isn't found, call it again later just looking for the highest-priority cargo)

- If organization policy, *SelectWaitingCMRLoadWithLeastServedOrganization*
  - Select the waiting load with the least-served organization that is going in the same direction as the asset
- If feed-to policy, *SelectWaitingCmrLoadWithUnfilledFeedToNode*
  - Select the waiting load for the highest-ranked feed-to-node that has not yet been fed enough and that is going in the same direction as the asset.
  - If there are no CMRs that meet this criteria, then call *SelectWaitingCmrLoadWithNextNode* to see if there are any CMRs at all going in the same direction as the asset.
- For all policies, if we did not already select a load, *SelectWaitingCmrLoadWithNextNode:*
  - Select the first in the ordered list of CMRs that is going in the same direction as the asset and, if passed an item priority, that has the item priority we're looking for.
- If item type weighted policy, calculate the number of items to fill the desired weight or volume, and update the remaining volume or weight to fill
- For all other policies, *GetNumItemsThatFitInAsset:* Calculate the number of items from the selected load that will fit in the remaining capacity of the asset, and limit it by the amount available here
- If the truck is full, so that no items can be loaded, exit the asset loop.
- If organization policy, update the amount sent to each organization with the amount to be loaded here
- Add a new load to the asset's leg, and add the amount of items calculated above.
- Subtract the amount loaded from the amount waiting
- Update the running total sent to each feed-to node, for the feed-to policy
- If there are no items left from this load, remove it from the list of waiting loads

- (Loop to the next waiting load)

- If there was no cargo waiting at this node, and if the asset has not already been routed somewhere else,
- *SetDestinationForEmptyAsset*
  - Get the updated locations of all loads, and remove those that are delivered
  - If item type priority, destination priority, or feed-to policy, find the node with the largest amount of highest-priority cargo
  - If FIFO, organization, or shortest policy, find the node with the largest amount of cargo
  - Set the asset's destination to the adjacent node along the path toward the selected node: add a leg to the asset going to that node.
  - If no loads were waiting, don't set any destination.
  - Skip to the next asset

- (Loop to next asset)

- (Loop to next node)

- Write results to sheets

Pseudocode for task-based heuristic, page 1

- *Heuristic_task*
  - *SetCmrPaths* [same as in location-based algorithm]
    - Sort the list of CMRs by:
      - item priority, if an item-priority policy
      - destination priority, if a dest-priority policy
      - item or destination priority if indicated in the path policy
      - FIFO order otherwise
    - For each CMR in the sorted list
      - If we have not yet sent enough cargo via shortest paths (as defined in the path policy), assign the CMR a shortest path; otherwise, assign a cheapest path. To do so,
      - *GetDijkstraShortestPath* on a network whose arc lengths are either the distances or the estimated cost to travel, where the estimated cost is the distance multiplied by the average cost for all transport assets (defined in the database) that have access to this arc.
  - *BuildTasks*
    - Define a collection of tasks, including first the set of user-defined tasks and then a set of auto-generated tasks for every origin and every vehicle type. The user-defined tasks are always chosen first, but the presence of the others ensures that we can always deliver everything.
  - For each time period,
    - Find the location, at the start of this period, of all CMRs and all vehicles, but subtract any inventory already slated to leave in the future (and not added to its arrival location if it has not yet arrived there)
    - Remove from the list any CMRs that have already reached their destinations
    - *ProcessAssignments*
      - For each assignment defined by the user that is starting at this time period, and for each vehicle that needs to be assigned to it,
        - Find the vehicle, of the right type and not already assigned, whose last planned location is closest to any of the nodes in the assignment network
        - *AssignAsset* to the assignment
          - Find the shortest path from the asset's last location to any of the nodes in the assignment
          - Route the asset (assign legs) along the shortest path
          - Assign the asset to only move within the subnetwork defined by the assignment (this does not affect the movement of the asset to the assignment network, defined previously)
        - If there are no vehicles of the right type that are not already assigned, do not assign any vehicles (and do not throw an error)
      - For each assignment defined by the user that is ending in this time period,
        - Unassign all vehicles: remove the restriction that they travel within the subnetwork defined by the assignment
    - If this time period is the beginning of a new day, reset the running (daily) total sent towards the feed-to-node, for the feed-to policy
    - Make a local copy of the tasks collection to modify in this time period
    - Do while there are still tasks in this period's list of tasks
      - Make a list of all the assets that are currently available for tasking, meaning moves are not planned for them beyond the current time
      - If there are no assets available, go to the next time period
      - For each task in this period's list of tasks, calculate the needs...
        - Clear the needs stored in this task
        - For each load at the task's origin node
          - If the load is waiting for this task,
            - Add the load amount to the total needs for this task

        - If item priority or destination priority policy, add the load amount to the needs for this task separated by priority level
        - If feed-to policy, and if the load is for a feed-to-node, and if we have not yet sent enough towards its feed-to-node, add the load amount to the needs for this task separated by rank of feed-to-node
        - If shortest path policy, add this load amount to the sum of path lengths (numerator of average) and add 1 to the total number of loads (denominator of average)
    - Next load
- Next task
- Remove any tasks that have zero needs from the list of tasks for this period
- If there are no tasks remaining in the list of tasks for this period, go to the next time period
- Select a task by some policy...
    - If most-need policy, select the task with the largest total amount of cargo waiting for it
    - If item priority, destination priority, or feed-to-node policy, select the task with the highest priority level, then within that level, the task with the greatest need (ties go to the task higher in the list)
    - If shortest path policy, select the task with the shortest average path length across loads
    - For defined task order or if we did not select a task earlier, select the first task in the list.
- Remove from the list of assets any assets that are the wrong type for the selected task
- If there are no appropriate assets for this task, remove this task from the list and continue to the next task
- Sort the collection of assets by the projected arrival time at the task's origin node, from shortest to longest
- Sort the list of loads at the task's origin node based on policies...
    - If item type policy, sort loads by item type priority
    - If destination policy, sort loads by destination priority, and also find the total capacity of all the assets here (to be used later)
    - If organization policy, set up arrays to track for each organization the number of CMRs that were sent (note that this is inside the task loop, so it resets for the next task. Could potentially put it outside to even out across time step or entire thing)
    - If shortest delivery policy, sort loads by length of path remaining for each load, shortest to longest
    - If feed-to-node policy, sort in order of rank of feed-to-node
- Decide how many vehicles to allocate to this task:
    - If proportional policy, find the total need for tasks with this vehicle type among all user-defined tasks (or auto-defined, if the selected task is auto-defined), and find the fraction of the total need made up by this task's need. Allocate the same fraction of the available vehicles to this task, rounded up to the nearest whole vehicle.
    - If first-come-first-served policy, allocate just enough vehicles to satisfy all the need (not just the priority need) of this task, rounded down so we don't allocate non-full trucks (except in the case of auto-defined tasks, allocate all available trucks).
- For each vehicle appropriate for the task, until the max number of vehicle assigned to this task,
    - Route the asset on a shortest path to the task's origin, by adding legs from its last location to the origin node, starting at the current time
    - Look for loads going along the path just added, from the asset's location to the task's origin, and *LoadVehicleWithCmrs*

- If destination-weighted policy, determine which destination-priority to look for with this asset:
  - Assign it to p1 destination cargo if the cumulative capacity (its capacity plus that of the others looped before this) falls in the range defined for p1 cargo, rounded up to the nearest multiple of this asset's capacity (to ensure that if there is only one asset, it goes to p1). The same for the other destination priorities.
- Route the vehicle to a destination:
  - If it's a "Path" type task, then assign legs for a path from the task origin along the task path to the task destination
  - If it's an "Origin" type task, then assign legs for a path the same way as in the location-based algorithm, except this time consider not just adjacent nodes but all nodes accessible to this asset.
  - *SetVehicleDestinationBasedOnCmr* [same as in location-based algorithm except we look at all accessible destinations not just adjacent nodes]
    - For each waiting load,
      - If the asset can travel to the load's farthest stop in its accessible or its assigned network,
      - If FIFO or shortest policy, select this load and exit the loop (this is the first in the sorted list that meets requirements)
      - If item type or destination or feed-to policy, add this CMR to a running total of the amount waiting to go to each adjacent node, of each priority, and select the load if it represents the largest such running total (but keep looping through the loads)
      - If organization policy, look up how much cargo has been sent for this load's organization, and select this load if it is the least-served organization found so far (but keep looping through the loads)
    - If FIFO, item type, destination-lexic, organization, or shortest, set the destination based on the selected load:
    - If a destination weighted policy, try to set the destination to that with the most cargo of the destination-priority this asset is looking for, but if there isn't any, set it to the one selected above:
    - If a feed-to policy, try to set the destination from a load going to the top-ranked feed-to-node, but if we've already sent enough cargo to it, then the next-ranked feed-to-node, etc. If all feed-to-nodes are adequately served, then to the next load without a feed-to-node, and if there are none of those, then back to the top-ranked feed-to-node:
    - Set the vehicle destination by adding a leg to the asset going from the current node to the adjacent node that the selected load must reach next, and return the leg.
    - If there were no loads at this node, return Nothing.
  - Load the first leg of the vehicle's new path with cargo...
  - If item type weighted policy, call *LoadVehicleWithCmrs* repeatedly, looking for the right amount of each priority-level
  - For all other policies,
  - *LoadVehicleWithCmrs*
    - If item type weighted policy, initialize a running total of the volume or weight we want to fill this time
    - While there are loads waiting here,
      - If item type weighted policy, call *SelectWaitingCmrLoadWithNextNode* looking for the right type of priority (and if it isn't found, call it again later just looking for the highest-priority cargo)

Pseudocode for task-based heuristic, page 4

        - If destination weighted policy, call *SelectWaitingCmrLoadWithNextNode* looking for the right type of destination-priority (and if it isn't found, call it again later just looking for the highest-priority cargo)
        - If organization policy, *SelectWaitingCMRLoadWithLeastServedOrganization*
            - Select the waiting load with the least-served organization that is going in the same direction as the asset
        - If feed-to policy, *SelectWaitingCmrLoadWithUnfilledFeedToNode*
            - Select the waiting load for the highest-ranked feed-to-node that has not yet been fed enough and that is going in the same direction as the asset.
            - If there are no CMRs that meet this criteria, then call *SelectWaitingCmrLoadWithNextNode* to see if there are any CMRs at all going in the same direction as the asset.
        - For all policies, if we did not already select a load, *SelectWaitingCmrLoadWithNextNode:*
            - Select the first in the ordered list of CMRs that is going in the same direction as the asset and, if passed an item priority, that has the item priority we're looking for.
        - If item type weighted policy, calculate the number of items to fill the desired weight or volume, and update the remaining volume or weight to fill
        - For all other policies, *GetNumItemsThatFitInAsset:* Calculate the number of items from the selected load that will fit in the remaining capacity of the asset, and limit it by the amount available here
        - If organization policy, update the amount sent to each organization with the amount to be loaded here
        - Add a new load to the asset's leg, and add the amount of items calculated above.
        - Subtract the amount loaded from the amount waiting
        - Update the running total sent to each feed-to node, for the feed-to policy
        - If there are no items left from this load, remove it from the list of waiting loads
      - (Loop to the next waiting load)
    - Copy the loads from the first leg, just loaded, to any subsequent legs in the asset's new path
  - (Loop to the next vehicle for this task)
  - Remove this task (because we've now gotten through all the assets allowed for it, or there is no remaining cargo for it)
- (Loop to the next task)
- (Loop to the next time period)