Chapter 1

Introduction

South Africa's level of urbanization closely follows international trends in developed countries, with the highest level of economic activity focused in a few metropolitan areas; attracting both people and investment. The good functioning of these metropolitan areas is of strategic importance to the country, as these areas are the main focus for economic and social development. The level of transport services provided impacts directly on the efficiency and the quality of the development in the metropolitan areas. South African metropolitan areas are experiencing rapid growth, and are having difficulties in controlling the physical urban expansion. Both public and freight transport costs are negatively impacted by these phenomena. As demand for transport increases faster than the supply of these services, commuting and freight transportation costs increase at a higher than inflation rate. The community at large experiences the demands for more extensive infrastructure and services.

Customers, both businesses and private consumers, demands products and services at the point of utilization. The geographically dispersed point of supply and point of utilization are bridged through transport. The majority of urban freight is carried by means of road transport, and the definition of the *Organization for Economic Co-operation and Development* (OECD) for urban freight transport applies:

> *"The delivery of consumer goods (not only retail, but also by other sectors such as manufacturing) in city and suburban areas, including the reverse flow of used goods in terms of clean waste."* — OECD (2003)

Goods transport has a major impact on the economic power, quality of life, accessibility and attractiveness of local communities, especially in city and metropolitan areas, but receives much less attention in comparison to passenger movement. According to the first *State of*

Logistics Survey for South Africa prepared by CSIR Transportek (2004), 83% of the total tonnage transport bill of ZAR 134 billion is transported via road, while 22% of the total tonnage is transported within metropolitan areas. Freight transport within metropolitan and urban areas have different characteristics from long haulage, and the main attributes include (Taniguchi et al., 2004):

- Frequent deliveries of smaller quantities
- Low utilisation of the capacity of trucks
- Time windows

Efficiently transporting goods within urban areas facilitates the establishment of sustainable cities. OECD (2003) acknowledges the contribution that freight vehicles make to traffic congestion, energy consumption and negative environmental impacts. Yamada and Taniguchi (2005) conclude that the majority of benefits for freight carriers can be achieved by implementing advanced vehicle routing and scheduling systems, hence addressing congestion, energy consumption, and indirectly environmental impacts. The problem concerned with allocating customer deliveries (or collections) to vehicles, and determining the visiting order of those customers on each vehicle route, is classified as the Vehicle Routing Problem (VRP), and has as its main objective to minimize some measurable function, such as distance traveled, time traveled, or total fleet cost.

1.1 *Modeling* as research motivation

South Africa provides a fascinating interface between the developed and the developing world. In a critical review, Leinbach and Stansfield (2002) have emphasized that Industrial Engineers should re-adopt a systematic view. They argue that the perception of Industrial Engineers has been negatively impacted by their ability to model the obvious, and in the oversimplification of their models, to the extent that reality is not represented comprehensively. Industrial engineers should therefor appreciate the complex and intertwined relationships between social, political, and economic factors influencing urban freight transport systems.

A systematic approach in addressing a problem is illustrated in the lower cycle of Figure 1.1 where a problem is modeled, the model is solved, and the solution is interpreted so as to change the original problem through decisions (Rardin, 1998). Identifying and scoping a problem is not a trivial matter, and is important in ensuring that the final solution that a decision is based upon, will in fact represent, and ultimately address the core problem. Taha

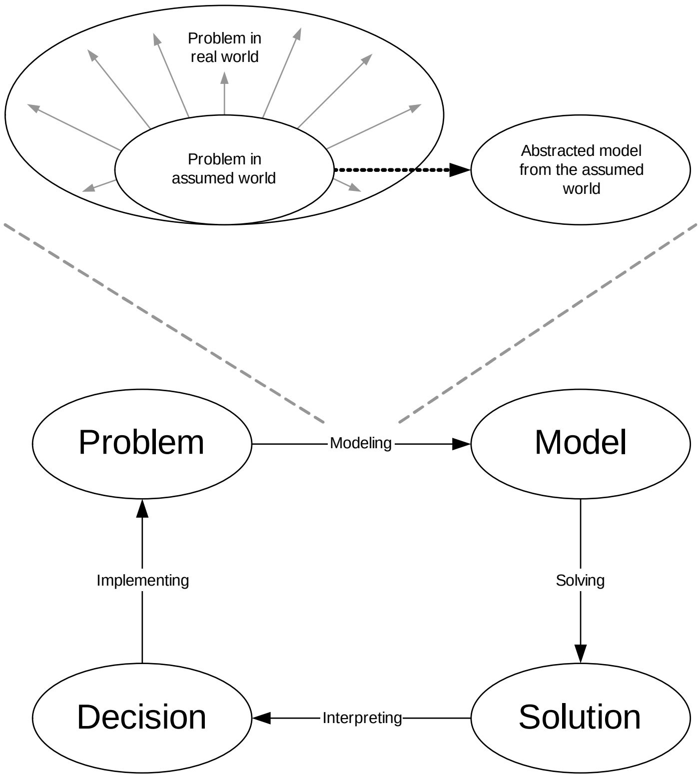

Figure 1.1: Operations Research cycle

(2003) expands the action of *modeling* in Figure 1.1 and illustrates how representations of the real world can easily be over-simplified. Interrelationships within the *real* world are so complex and abundant, that no one person can comprehend it in its entirety. We refer to the *problem in the real world* as the first level of abstraction. The human is a contextual being: the cultural, social and emotional context of an individual forms the individual's perception of the reality in which he or she exist. The second level of abstraction therefor represents the contextually sensitive view, referred here to as the *assumed reality*, that an individual has of the real problem. But even the abstract and fragmented view is often too complex to solve in its entirety. Through the actions of analyzing, and applying a methodology of *divide-and-conquer*, the individual scopes the problem in a structured way through simplifying assumptions. These assumptions may be justified in the absence of complete and accurate data about the assumed reality. The third level of abstraction is referred to as a *model*. The verb *modeling* therefor requires the problem solver to not only scope the problem, but also justify the endeavors to ensure that the assumed reality has been challenged to represent the real problem more comprehensively. This is illustrated through the arrows stretching the boundaries of the assumed reality towards the real world. Although the model can be any representation of a real problem, from scraps of paper with notes on them, a functional flow block diagram or process maps, in this thesis the term is used as a structured and mathematical model with an optimization intent.

Once the model is a true representation of the problem at hand, the decision maker can proceed to *solve* the model. It should be emphasized that only the model is solved, and not the problem itself. The availability and the ease of use of new generation optimization software have facilitated the process of solving models representing complex operational problems. The rapidly increasing processing power of computers brings the optimization opportunities right to the desk of the practitioner. The solution, however, is often but a list of numerical results.

The numerical solution, and its sensitivity to changes in parameters, requires careful consideration before recommendations and decisions are made, and is only considered as decision *support*. Implementation impacts, and possible change factors are considered before a final decision is made and implemented. The impact of the decision is then *assessed* so as to close the problem solving-cycle. Implemented changes may either address the original problem adequately, or may elicit new problems that require modeling, solving, and decision making.

1.2 *Intelligence* as the research driver

Freight carriers are sharing the road network with various modes of public transport. The use of private vehicles have rapidly increased. The increase can be attributed to both an increase in the number of trips undertaken, and increased journey lengths (Banister, 1995; Spence, 1998). Road network performance is negatively impacted by the higher usage of private vehicles and results in higher levels of congestion, and a significant reduction in operating speeds. Public transport performance is impacted negatively when operating speeds decrease, resulting in increased operating costs for the carriers, and thus impacting negatively on its attractiveness. As a result, the economically able part of the population turn to their private vehicles for a reliable source of transport, and unknowingly contributes to the hyper-congestion phenomenon.

Congestion does not only increase the stress levels of road users from a commuting point of view, but it also increases the complexity for vehicle and fleet managers overseeing the scheduling, routing, and optimization of their fleet concerns.

Carrier companies represent both public and private entities executing the logistic and distribution functions of freight. This thesis addresses the complexities of freight transport. Freight carriers are continuously expected to provide higher levels of service at lower rates, and therefor try to minimize their logistic costs, and maximize their profit. Sharing the road infrastructure with other vehicles such as private cars and public transport forces carriers to plan their freight routes more carefully. Enhanced vehicle routing and scheduling takes the congestion constraint into account and attempts to improve the vehicular utility through shorter routes and higher load factors. Software applications often do not provide adequate functionality by not being able to address complex business requirements such as companies having a fleet of vehicles that differ in capacity and/or running costs, and multiple scheduling where vehicles are allowed to complete a trip, return to the depot to renew it's capacity, i.e. offload goods collected, or loading goods to be delivered. The reason for software deficiencies are related to the extreme computational complexity when solving routing models. Human intervention is required to, for instance, split the fleet into vehicle categories that represent similar or the same capacity and/or costs. Each category is then solved independently, adjusting demand as customers are serviced by other categories. Human operators can also intervene by evaluating vehicular routes, and identifying vehicles that may be used for a second trip, and then schedule such vehicles accordingly. Although such interventions are mechanistic in nature, they require the time and effort of experienced individuals having a thorough understanding of vehicle routing so as to intervene wisely.

We refer to ourselves (in a more formal way) as *homo sapiens* — man the wise — and value our mental abilities to *think* and *reason* to assist us in improving our surroundings. We require our *thought processes* and *intelligence* to make decisions that will maximize the utility that we obtain from logistics — moving goods from points of manufacture to points of consumption that are geographically dispersed.

> *"What is mind? What is the relationship between mind and the brain? What is thought? What are the mechanisms that give rise to imagination? What is perception and how is it related to the object perceived? What are emotions and why do we have them? What is will and how do we choose what we intend to do? How do we convert intentions into action? How do we plan and how do we know what to expect from the future?"*—Albus (1999)

It seems clear from the quote by Albus (1999) that before one toss terms such as *thinking* and *planning* around, one should carefully consider how such actions take place, and how one intends to employ such actions to improve, for example, urban freight congestion.

1.2.1 Intelligence

In their leading text, Russell and Norvig (2003) introduces Artificial Intelligence (AI) as not only understanding the human intellect, but also building entities (or agents) that *are* intelligent. Although it encompasses a huge variety of subfields of study, with many varying definitions, the authors have categorized AI approaches in a two-dimensional framework represented in Figure 1.2.

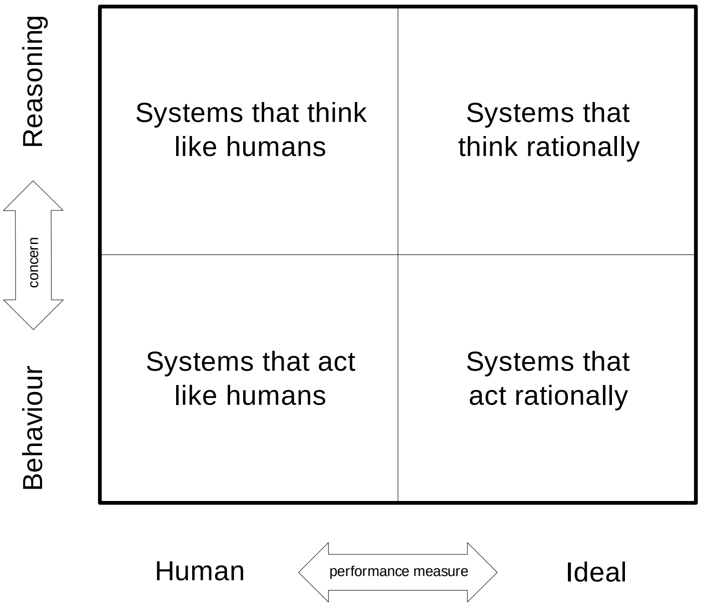

Figure 1.2: Categories of artificial intelligence (Adapted from **?**)

The top half of the framework is concerned with thought processes and reasoning, as opposed to the lower half that is concerned with the behavioral element of intelligence. The left side of the framework measures the success of an agent's intelligence against the fidelity of human performance. The right half establishes an *ideal* concept of intelligence as a benchmark, referred to as *rationality*. This is analogous to effectiveness — *doing the right things*. However, the *right* within *rationality* is only relative to what is known at the time of the *doing*.

An *agent* is something that acts. This thesis is concerned with the development of a computer agent that could intelligently intervene in the routing and scheduling of distribution vehicles. But how is it to be distinguished from mere *programming*? It should be able to operate autonomously, perceive the environment, persist over a period of time, and be able to adopt the goals and objectives of another entity. As an improvement on a basic agent, this thesis propose a *rational agent* that has a strategy to achieve the best possible outcome for a given objective, either known, or the expected outcome should some of the parameters be uncertain. The focus of the thesis is therefor not on understanding the human thought processes, but on creating a system that can think, and act rationally.

1.2.2 Complexity

Perfect rationality in modeling is often too difficult to attain due to too high computational demands when looking for exact solutions. Problems such as the routing and scheduling of vehicles can often not be solved exactly, and require the use of solution algorithms that provided approximate solutions where the optimality of the solution can neither be proved in advance, nor confirmed once a solution is found. The different opinions with regards to either finding an exact optimal solution versus settling for a *good enough* solution given a specific environment have led to the split that occurred between *Decision Theory* and *Artificial Intelligence* in the latter half of the twentieth century.

Decision Theory is the field of study where probability theory and utility theory are combined to present a formal framework for decision making under uncertainty. The field of operations research addresses complex management decisions rationally. The intention of the pure branch of decision theory is to obtain a rational decision, or a global optimum.

On the contrary, the complexity in finding a single optimum value led the pioneers of AI such as Herbert Simon (1916–2001) to prove that being able to find a *good enough* answer describes human behavior more accurately — and earned him the Nobel prize in economics in 1978. And although the computational ability of computers have increased dramatically over the past decade, the intention is still to assist mere mortal logistics decision makers to improve their ability to manage distribution fleets.

1.3 Formulating the research question

The primary research question that this thesis intends to answer is *whether it is feasible to develop a rational and intelligent agent to schedule a predefined variant of the VRP.* In order to answer the question, a number of secondary research questions will be stated in terms of the concept of an *intelligent agent.*

In his paper on the engineering of mind, Albus (1999) identifies four functional elements of an intelligent system.

Sensory perception — accepting input data from both outside and from within the system. The data is then transformed through classification and clustering into meaningful representations of the real world. The first secondary research question addresses the analysis of input data and is stated as follows:

> *How should customer parameters be clustered so that meaningful classification can be done prior to executing the solution process?*

Behavior generation — planning and controlling actions so that goals are achieved. An intelligent agent accepts task with goals, objects and priorities. The tasks are then broken up into jobs and, along with resources, are assigned to agents. Hypothetical plans are created and simulated to predict the outcome of the plans. The simulated results are evaluated, and the agent selects the best expected hypothesized plan. In terms of this thesis an agent refers to computational elements that plan and control the execution of a routing algorithm, correcting for errors and perturbations along the way. The planning processes of the agent are heuristics and metaheuristics that attempt to converge to optimal vehicle routes and schedules. This lead to another secondary research question:

> *How can heuristics and metaheuristics be used to establish vehicle routes and schedules in a complex and constrained environment?*

Value judgement — the computation of a predefined set of costs, risks, benefits, and or penalties related to the vehicle routes. In operations research terms these computational expressions are referred to as the objective function(s). The third secondary research question is derived from value judgement:

> *What should constitute the objective function of the model so that the real problem is adequately represented?*

World modeling — an overall strategy that uses input parameters and variables to update a knowledge database. Data is used to query the behavior generation of plans regarding current routes and schedules. The strategy further simulates possible results of future plans after analyzing the current plans. Simulated results are evaluated, using the value judgement, so the best expected plan for execution can be selected. After execution, the strategy allows for sensory expectations to be created regarding future actions — analogous to bumping your feet against an obstacle in the dark. After stumbling, and reacting to the pain, you lift your feet unnaturally high so as to avoid the next obstacle. The fourth and fifth, probably the most challenging secondary research questions addresses the agents ability to learn from the past and improve in future:

What critical parameters influence the agent's learning, and should therefor be included in creating future expectations?

How are future expectations created from the past performance?

1.4 Research design and methodology

The process diagram in Figure 1.3 provides an overview of an intelligent agent's decision process. The agent in this thesis will be a hybrid computerized solution algorithm that has

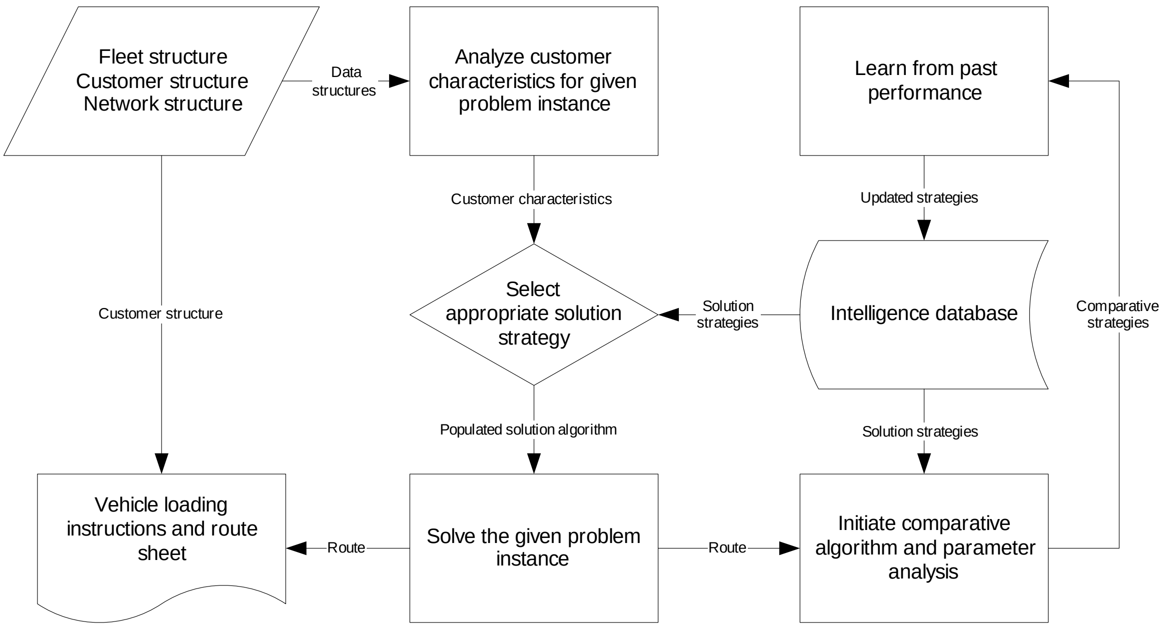

Figure 1.3: Overview of the intelligent agent's decision process

the following inputs:

- Fleet structure
- Customer structure, i.e. demand quantity, geographical location, time windows
- Network structure, derived from customers' geographical locations

The algorithm will analyze the clustering characteristics of the geographical distribution of customers. Based on the randomness (or clusteredness) of the distribution and the time window characteristics of the customers, the algorithm will select an appropriate solution strategy — a combination of a metaheuristic solution algorithm, along with its appropriate parameter values. The problem instance is solved, and the solution is interpreted and presented in a useful loading instruction and route sheet. Behind the scenes the algorithm will initiate comparative analysis of the proposed solution strategy by solving the provided problem instance with various metaheuristics and various parameter values for each metaheuristic. The algorithm will then learn from these analysis through a neural network, and update the intelligence database by recommending new solution strategies for the given problem instance, or reiterating current solution strategies.

The algorithm will be coded using the *MATLAB*® development environment. The analysis and solution components, and the comparative analysis components will run on separate computer processors to optimize for speed and in doing so, address the computational complexity of the hybrid algorithm.

1.5 The structure of the thesis

To elaborate on the exact nature of the research problem, Chapter 2 reviews literature on the VRP and its variants. The chapter concludes with the mathematical formulation of the *Capacitated Heterogeneous Fleet Vehicle Routing Problem with Multiple Soft Time Windows and Probabilistic Travel and Service Time* as addressed in this thesis. The review of solution algorithms, both exact and approximate, are conducted in Chapter 3, concluding with the recommendation of two metaheuristic solution algorithms, each covered in more detail in later chapters. The analysis of the customer structure is reviewed in Chapter 7, and the chapter proposes an algorithm to determine the level of clusteredness of a customer network. The algorithm is tested by analyzing benchmark data sets provided for pre-defined problem instances in literature.

Chapters 4 through 6 is dedicated to the development of various metaheuristic solution algorithms. Chapter 4 develops an improved initial solution algorithm to enhance the computational performance of the Tabu Search solution algorithm, developed in Chapter 5. The Genetic Algorithm is less sensitive to the quality of an initial solution, and is treated independently in Chapter 6. For each metaheuristic the various parameters are discussed, and default values proposed. The respective algorithms are discussed at high level, followed by detailed discussions of algorithmic particularities, and concluded by testing and validating the algorithm through benchmark data sets.

The integration of the algorithms, as well as the agent's ability to learn from repetitive decision making is covered in Chapter 8. The thesis is concluded in Chapter 9 with a critical analysis of the research contribution, and setting a research agenda.