

VICTORIA UNIVERSITY OF WELLINGTON
Te Whare Wānanga o te Ūpoko o te Ika a Māui



School of Engineering and Computer Science
Te Kura Mātai Pūkaha, Pūrorohiko

PO Box 600
Wellington
New Zealand

Tel: +64 4 463 5341
Fax: +64 4 463 5045
Internet: office@ecs.vuw.ac.nz

Multifaceted Web Service Composition

Alexandre Sawczuk da Silva

Supervisors: Dr. Hui Ma, Prof. Mengjie Zhang

May 14, 2015

Submitted in partial fulfilment of the requirements for
PhD.

Abstract

A short description of the project goes here.

Contents

1	Introduction	1
1.1	Problem Statement	1
1.2	Motivations	2
1.3	Research Goals	4
1.4	Published Papers	8
1.5	Organisation of Proposal	8
2	Literature Review	9
2.1	An Overview of Web Service Composition	9
2.1.1	Web Service Composition in Practice	11
2.2	Single-Objective QoS-Aware Composition Approaches	12
2.2.1	Biologically-Inspired Approaches	12
2.2.2	Other Optimised Composition Approaches	16
2.3	Multi-objective, Top-K, and Many-objective Composition Approaches	17
2.4	AI Planning-based Composition Approaches	20
2.4.1	Hybrid Approaches	22
2.5	Semantic Composition Approaches	23
2.6	Dynamic, Distributed, and Cloud Composition	28

Figures

1.1	Research questions and their corresponding objectives.	7
2.1	Typical steps in a workflow-based automated Web service composition solution [61].	10
2.2	The different problem representations employed by biologically-inspired Web service composition approaches.	13
2.3	Basic example of Web service composition using the Graphplan algorithm. . .	20

Chapter 1

Introduction

1.1 Problem Statement

As applications increasingly interact with the Web, the concept of Service-Oriented Architecture (SOA) [66] emerges as a popular solution. The key components of SOA are often Web services, which are functionality modules that provide operations accessible over the network via a standard communication protocol [29]. One of the greatest strengths of Web services is their modularity, because it allows the reuse of independent services that provide a desired operation as opposed to having to re-implement that functionality. The combination of multiple modular Web services to achieve a single, more complex task is known as *Web service composition*. At a basic level, services are combined according to the functionality they provide, i.e. the inputs required by their operations and the outputs produced after execution. Several services may be connected to each other, with the outputs of a service satisfying the inputs of the next, starting from the composition's given overall input and finally leading to the composition's required output.

An example of an application of Web service composition is ENVISION, an EU-funded project whose aim is to provide a framework for discovering and composing Web services that perform geospatial analysis on data, thus enabling environmental information to be more easily processed for research and decision-making purposes [57]. ENVISION is meant to be used by scientists who are experienced with geographic models but do not have a technical computing background, therefore the motivation of this work was to create a solution that is as simple to use as possible. The system offers a way to search for (discover) services that provide environmental data as well as processing services by using a word-based and coordinate-based search. Users create compositions by manually selecting and assembling a Business Process Model (BPM), which is later transformed into an engine-runnable representation. The environmental data and processing applications are packaged as services, meaning that they are easily available for reuse and thus contribute to the faster identification of important environmental trends.

Despite the usefulness of Web service composition, conducting such a process manually is fraught with difficulties. In order to illustrate these problems, an experiment was performed in which students with a good knowledge of programming and of Web services were asked to manually create Web service compositions to address a range of well defined real-world problems [52]. Results show that they faced a number of difficulties at different composition stages, from the discovery of potential services to their combination. During the discovery phase, the most popular search tools used by students were Web service portals and generic search engines. Authors highlighted that not a single discovery tool from the literature was used, because no single solution offers a large variety of services. This suggests that larger standardised service portals should be created, provided that the qual-

ity of the services offered is maintained. During the combination phase, students faced discrepancies between the concepts used by the interfaces of services: the terms used in the interfaces of any two services are often different from each other, even though those services may handle the exact same domain. Another problem is that the services used in a composition may produce too much data, or incur too much latency, meaning that the final composition is slowed down as a result. Standardising and automating the service composition process would eliminate the need for dealing with these difficulties manually, thus developing a system capable of creating compositions in a fully automated manner is one of the holy grails of the field [58].

Automated Web service composition is complex, and in fact it is considered an *NP-hard problem*, meaning that the solution to large tasks is not likely to be found with reasonable computation times [61]. Due to this complexity, a variety of strategies have been investigated in the literature, focusing on two fundamental composition approaches: *workflow-based approaches*, where the central idea is that the user provides an abstract business process which is to be completed using concrete services, and *AI planning-based approaches*, where this abstract process is not typically required and instead the focus is on discovering the connections between services that lead from the task's provided output to its desired [61]. In workflow-based approaches, the abstract functionality steps necessary to accomplish the task requested by the user have already been provided, so the objective is to select the concrete services with the best possible quality to fulfil each workflow step. The advantage of such approaches is that the selection of services may easily be translated into an optimisation problem where the objective is to achieve the best overall composition quality. This optimisation is often performed using *Evolutionary Computation* techniques. However, the disadvantage is that a workflow must have been already defined, which likely means that it has to be manually designed. Planning-based approaches, on the other hand, have the advantage of both determining the workflow to be used in the composition and selecting services to be used at each step, abiding by user constraints. The disadvantage of such approaches is that they are unable to also perform quality-based optimisation on the selected services.

The overall goal of this thesis is to propose a Web service composition approach that hybridises elements of AI planning-based approaches and Evolutionary Computation workflow-based approaches, enabling the construction of a workflow according to user constraints as well as the optimisation of that workflow according to the quality of its composing services. This new approach tackles several aspects of Web service composition, such as the use of multi-objective optimisation, semantic Web service selection, and dynamic Web service composition.

1.2 Motivations

The intricacy of Web service composition lies in the number of distinct facets it must simultaneously account for. As the **first facet**, services must be combined so that their operation inputs and outputs are properly linked, i.e. the output produced by a given service is usable as input by the next services in the composition, eventually leading up to the desired overall output. As the **second facet**, the composition must meet any specified user constraint. A constraint is defined as a user restriction that must be met in order for a composition solution to be considered valid, and this mainly concerns execution flow features (e.g. the composition must have multiple execution options – branches – according to a given condition) [82, 76, 38]. As the **third facet**, the composition must achieve the best possible overall Quality of Service (QoS) with regards to attributes such as the time required to execute the

composite services, the financial cost of utilising the service modules, and the reliability of those modules. As the **fourth facet**, the environment of the composition must be recognised as dynamic, with QoS values that fluctuate and services that become unavailable/available over time. These facets are discussed in more detail below:

1. **Functionality:** Functional solutions can be obtained by using two main approaches: *immediate* and *gradual*. Whenever a solution is built using immediate approaches, this solution is fully functional, i.e. the inputs of all atomic services are fully satisfied, and so are the overall task outputs given the provided inputs. This is typically done by employing AI planning algorithms that verify the correctness and completeness of the connections between services at every building step [82]. When a solution is built using gradual approaches, on the other hand, there are no guarantees that it will be fully functional. As opposed to performing checks at every building step, gradual approaches rely on the notion of improving solutions over multiple iterations, each time penalising incorrect and/or incomplete connections between services. In practice, these approaches are implemented using Evolutionary Computation techniques with penalties enforced through the fitness function [72]. The simplest way of establishing that the inputs of an atomic service have been satisfied is by verifying that there is an exact type match between each target input and its corresponding incoming connection (i.e. the output of a previous service). However, a more sophisticated matching approach relies on measuring the semantic similarity between the output and input types in question with the help of an ontology [9]; the smaller the distance between the two values, the better the match.
2. **Conditional constraints:** User constraints and preferences are a common requirement when performing Web service composition. For example, consider the case of a user who wishes to book transportation and accommodation for a trip, including transportation from the airport to the hotel, a room at the hotel for a given number of days, and transportation back to the airport [10]. In this scenario, the user is likely to have certain conditional constraints; for example, if the hotel booking is for a two-star hotel, then the transportation to and from the airport should be arranged using a taxi service, otherwise a shuttle service should be used. Different techniques have been explored to achieve compositions that consider such constraints, including the use of AI planning with rules encoding user constraints [9] and the representation of the composition as a constraint satisfaction problem to be processed with a solver engine [38]. This territory has not been widely explored by applying Evolutionary Computation (EC) techniques; yet, due to their flexibility and efficiency, it would be interesting to focus on the investigation of ways in which to extend them to apply these constraints. It must be noted that constraints relating to Quality of Service (QoS) values are not included in this facet.
3. **Quality of Service:** Quality of Service (QoS) refers to the non-functional (quality) attributes associated with a Web service, such as its expected execution time when answering requests, its availability, and its scalability [42]. The overall quality of compositions can be measured by aggregating the individual QoS values of its constituting atomic services, meaning that it is possible to optimise the composition's QoS by selecting the right combination of constituting services. As mentioned earlier, the selection of a set of services in order to maximise the overall QoS can be mapped into a classic optimisation scenario, and this has been done extensively in the literature. A variety of EC techniques have been used in this context [81], as well as other optimisation techniques such as integer linear programming [90]. Interestingly, AI planning

techniques were also used to this end, even though in this case they are incapable of also considering the conditional constraints facet [20]. In the realm of QoS optimisation, multi-objective techniques are often employed to the problem of Web service composition, since the optimisation of potentially conflicting QoS attributes such as time and cost is more intuitively performed using independent objective functions [50].

4. **Dynamic composition:** In a more realistic scenario, Web service compositions exist within a dynamic environment where the quality and availability of the atomic services in the repository varies as time passes. In such a dynamic environment, providing a static composition solution to a task is no longer enough, since this solution may decrease in quality throughout time and/or become non-executable if some of its composing services go offline. Thus, the focus of dynamic Web service composition is on monitoring and updating composition solutions as they become outdated [44]. The majority of techniques aimed at updating outdated or faulty compositions rely on building solutions that present constructs allowing for dynamic adaptation [2], but not many EC-based approaches have been tried in this area.

The above discussion references several techniques that have been proposed to address the composition problem. These techniques produce promising results, however they do not account for all of these composition facets at once. For example, AI planning techniques for composition focus on guaranteeing functional correctness (first facet), and either fulfilling constraints (second facet) or optimising QoS (third facet); similarly, EC techniques such as Genetic Algorithms (GA) and Genetic Programming (GP) focus on QoS in addition to functional correctness, but do not include the fulfilment of constraints.

1.3 Research Goals

The overall goal of this thesis is to propose a hybrid Web service composition approach that considers elements from all the four facets described above when generating solutions. More specifically, this approach combines elements of AI planning, to ensure functional correctness and constraint fulfilment, and of Evolutionary Computation, to evolve a population of near-optimised solutions from a QoS standpoint. The research aims to determine a flexible way in which planning and EC can be combined to allow the creation of solutions to solve composition problems that require multiple execution paths. The work conducted in this thesis intends to address the following research questions:

- (i) *What would be the components of a technique to create composition representations that allow for different execution paths to be optimised simultaneously with regards to quality?*

Traditionally, AI planning and EC have been employed separately to solve the problem of Web service composition. On the one hand, AI planning techniques are outstanding at ensuring the creation of composition solutions that have appropriate connections between the outputs and inputs of the composing services, and also at ensuring the creation of branches according to the constraints provided. On the other hand, EC techniques are ideal for encountering solutions with a good overall quality amongst a very large array of possibilities. Given the strengths of each set of techniques, it is ideal to combine these two composition strategies into a single approach that offers both sets of capabilities, and thus considers more service composition facets simultaneously. In this hybrid approach, users would be able to specify the conditions of when branching should occur, the order in which these conditions should be observed, and the outputs each execution branch should produce; the technique then returns a suitable solution.

- (ii) *What is the best way of independently optimising the Quality of Service (QoS) attributes of a composition with multiple execution branches?*

Existing optimisation approaches have focused on optimising the QoS of composite services that have only one execution path, by relying either on a single or on multiple objective functions. In the case of single-objective approaches, the different QoS attributes are combined through a weighted sum that produces a unified quality score used for ranking candidate solutions; multi-objective approaches, on the other hand, evaluate each QoS attribute using a separate score and divide candidate solutions into groups by comparing each of these dimensions simultaneously. When optimising services with multiple execution paths, the quality of different branches must be considered, either by aggregating the QoS values of each branch or by considering them independently.

- (iii) *What is the best semantic Web service selection method for a hybrid planning-EC technique?*

When building a candidate composition solution, atomic services must be selected according to the compatibility of their inputs. The simplest form of selection is when the inputs of services are matched according to their exact type, though recently more sophisticated semantic approaches that also allow for inexact matches have also been investigated. In a typical selection scenario, a concrete service is chosen to fulfil the functionality specified by an already-defined abstract service, which restricts the complexity of the problem. However, the concept of an abstract service cannot be used when selecting services during an AI planning-based composition, which decreases the efficiency of this approach. Thus, a better way of performing semantic selection must be designed for use with AI planning-based composition approaches.

- (iv) *What are the adaptations necessary to make an EC-based composition technique suitable to a dynamic environment?*

Evolutionary Computation (EC) approaches to Web service composition have been extensively investigated in static scenarios, where the quality and availability of services are assumed to remain constant, however the use of these techniques has only been superficially explored in dynamic scenarios. Despite this lack of research, EC-based approaches show promise in the area of dynamic Web service composition for two reasons: firstly, they allow a composite service system to *self-heal* by maintaining a population of solutions and using them as alternative compositions in the case of failure; secondly, they support *dynamic adaptation* by allowing solutions to be further evolved in order to account for changes in the QoS values of atomic services. Given these advantages, it is beneficial to investigate the use of EC approaches in a dynamic composition context.

The research goal described and outlined above can be achieved by completing the following set of objectives, which are intended to be used as research guides throughout this project:

1. *Determine the best direct representation for the planning-EC composition technique.*

The creation of a hybrid composition technique that combines elements from planning and evolutionary computation requires a decision on the representation each composition solution should adopt. Undoubtedly the simplest model would be that of a linear vector, with each element being a service that should be included into the composition, however this linear structure does not satisfactorily encode the relationships and links between different services. Another problem with a vector is that the EC techniques

which use such a representation do not allow for structures of varying lengths, a requirement when performing fully automated composition. A tree or graph representation would be better suited to the task at hand, since such structures are naturally capable of representing the links between the composing services and also multiple execution branches. Choosing between these two structures presents some trade-offs that must be carefully considered, such as the existence of EC techniques that support that structure and the computational cost they incur.

2. *Develop an indirect representation for the planning-EC composition technique.*

The structures discussed in Objective 1 may be referred to as *direct representations* of solutions, since the genotype and the phenotype of each solution are the same, meaning that solutions are represented directly as they should be interpreted. Despite the intuitiveness of direct representations, it has been argued in the problem domain of scheduling that indirect representations, where the final result must be decoded from a population candidate, have been shown to have better search performance than their direct counterparts [31, 16]. Due to this evidence, an indirect candidate representation should be proposed and compared to the direct representation. This indirect representation could be based on the Web service composition Particle Swarm Optimisation (PSO) approach proposed in [17].

3. *Develop a hybrid representation for the planning-EC composition technique.*

In case there is a trade-off between the execution time and the quality of the solutions produced by the direct and indirect representations developed in Objectives 1 and 2, a hybrid representation should be developed to combine the strengths of these two approaches. This representation could comprise the structure used in the direct representation with the addition of weights from the indirect representation. Comparisons should be performed to determine whether the hybrid solution presents any performance or quality gains.

4. *Propose a many-objective approach to optimise the quality of candidates with multiple execution branches.*

Two factors must be taken into account when optimising Web service compositions with multiple execution branches: firstly, each Quality of Service (QoS) attribute must be optimised independently, since they may be conflicting to each other; secondly, each branch of the composition must also be independently considered, since the QoS values for each branch may vary significantly despite their common root. Given these two factors, a many-objective technique should be employed in order to optimise the quality of solutions while at the same time considering these two factors independently.

5. *Develop an EC technique for performing semantic Web service selection in the context of a planning-based composition technique.*

As discussed in Objective 1, the composition technique investigated in this thesis combines elements of AI planning and of EC approaches to achieve compositions that are correct, have conditional branching, and can be optimised according to their QoS attributes. In planning-based approaches to service composition, however, the semantic selection process of candidate atomic services is inefficient, since there are many possible service matches to consider at each planning step. In fact, it may be unfeasible to consider all possible service matches in an exhaustive manner when handling larger service repositories, which invites the use of non-exhaustive approaches such as EC techniques to accomplish this task. Thus, this objective entails the use of an optimisation to discover semantically matching services to be included in a composition candi-

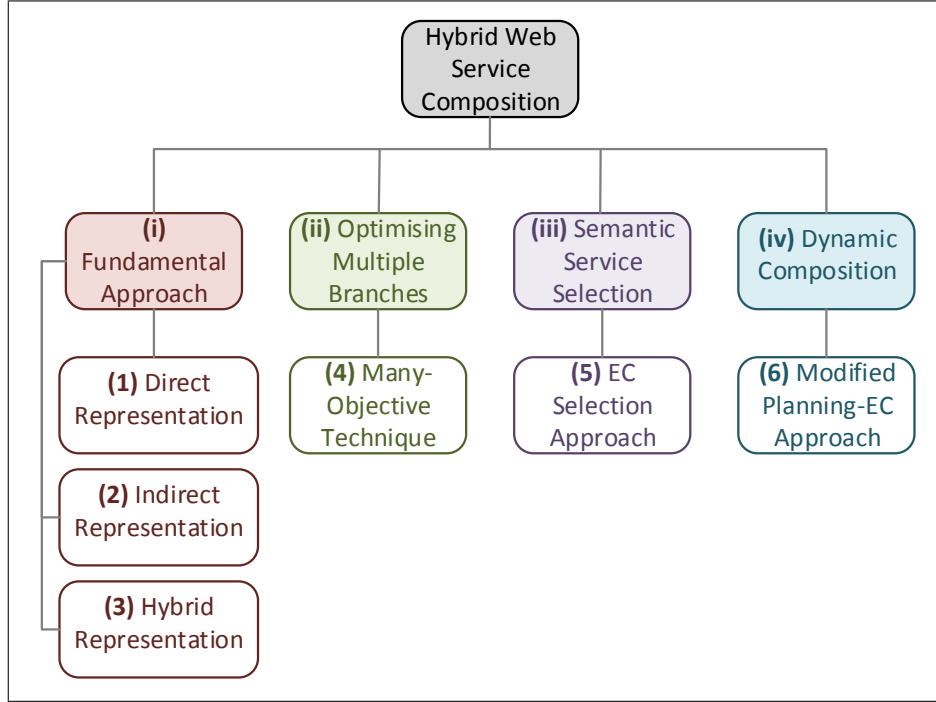


Figure 1.1: Research questions and their corresponding objectives.

date. Note that this objective aims to investigate the use of an optimisation technique nested within the service selection step of the planning-EC approach.

6. *Modify the planning-EC composition technique to work in a dynamic environment.*

In a static scenario the planning-EC technique is executed once for a given composition task, returning a composite result under the assumption that the quality levels and the availability of the atomic services included in that result will remain constant. In a more realistic dynamic scenario, however, the quality of the services in the repository may fluctuate and occasionally services may become unavailable. To account for these setbacks, solutions must be corrected and updated in response to changes in the environment. In order to do this, the planning-EC technique is to be modified to retain a population of candidates as alternatives in case of failure, and further generations are to be evolved as the QoS values of services in the repository change, thus leveraging the natural features of Evolutionary Computation.

As shown in Figure 1.1, Objectives 1, 2, and 3 are the core of this project, and answer research question (i) by providing the basis for an optimisation technique that admits Web service composition solutions with multiple execution branches. Objective 4, on the other hand, answers research question (ii), which aims to determine the ways in which the branches of a solution with conditional constraints can be independently optimised. Objective 5 answers research question (iii), leading to the investigation of a semantic selection approach for planning-based composition techniques. Finally, Objective 6 answers question (iv) by proposing one way in which EC-based techniques can be used in a dynamic Web service composition scenario.

1.4 Published Papers

1.5 Organisation of Proposal

The remainder of the proposal is organised as follows: Chapter ...; Chapter ...; Chapter

Chapter 2

Literature Review

This chapter is divided into...

2.1 An Overview of Web Service Composition

At its basic level, a Web service composition is the connection of several atomic services in different configurations in order to reach a result, given that there are multiple services offering the same functionality. The key aspect of compositions is that, in order to achieve the desired result, atomic services must all be executed in a particular order, forming a workflow of tasks. A workflow-based Web service composition approach can usually be decomposed into a series of steps, reflecting the process required to produce a solution [61]. These steps are shown in Figure 2.1 and discussed below:

1. *Goal specification:* The initial step in a Web service composition is to gather the user's goal for the solution to be produced. This is typically done through the generation of an abstract workflow that records the desired data flow and functionality details. This workflow is generated by the user and is referred to as *abstract*, since it contains a series of tasks that can be accomplished by employing a number of different existing Web service implementations. In addition to this workflow, the QoS requirements are determined based on the user's preferences.
2. *Service discovery:* Once an abstract workflow and a set of preferences have been provided, the next step is to discover candidate concrete services that are functionally and non-functionally suitable to fill the workflow slots in a service repository. The focus at this stage is to find candidates that provide the functionality required to fulfil the tasks, regardless of their quality levels.
3. *Service selection:* After pools of candidate services have been identified for each workflow slot, a technique is employed to determine which discovered services best fulfil each slot. The result of this process is the creation of a concrete Web service composition.
4. *Service execution:* The creation of the composition is followed by the execution of an instance of this composed Web service.
5. *Service maintenance and monitoring:* During execution, the created instance is constantly monitored for failures and/or changes to the composing atomic services, and corrective actions are dynamically carried out as necessary.

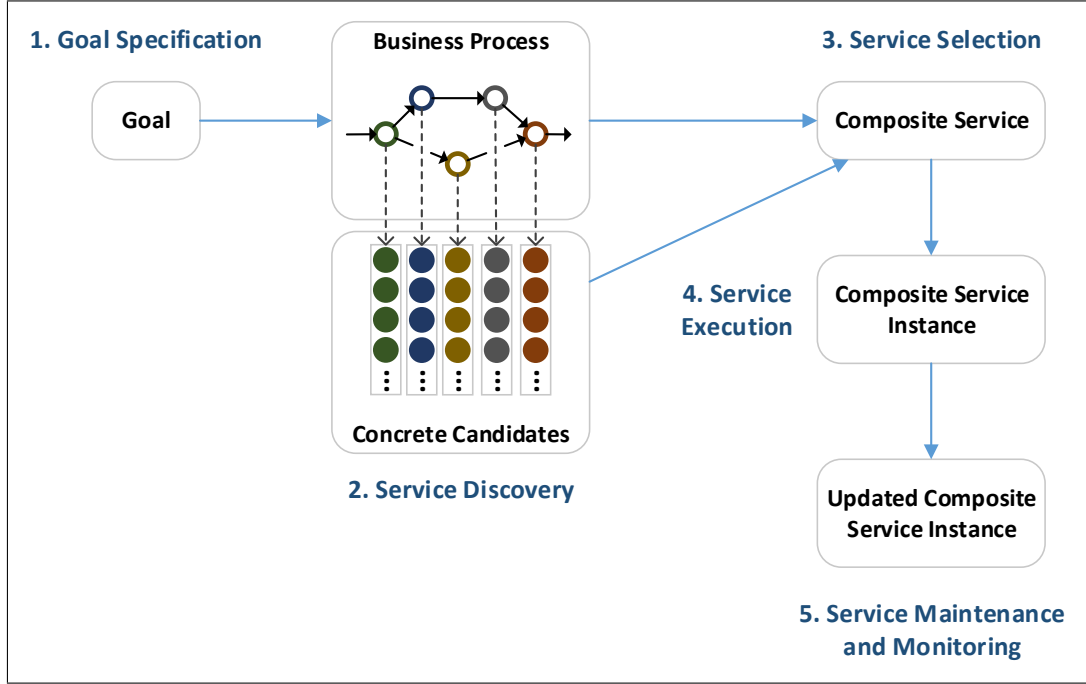


Figure 2.1: Typical steps in a workflow-based automated Web service composition solution [61].

It is important to draw a distinction between semi-automated and fully automated composition approaches [71]. The steps discussed above are typical of a **semi-automated approach**, where an abstract workflow is provided in the goal specification stage and the composition algorithm is only required to complete the abstract slots of this workflow. In a **fully automated approach**, on the other hand, an abstract workflow is not provided during the goal specification stage, and instead is calculated at the same time that services are selected based on user preferences such as the desired overall composition inputs and outputs. Consequently, service discovery may at times be also executed in tandem with the selection stage. Fully automated approaches have been shown to be more flexible than approaches with fixed abstract workflows (i.e. semi-automated approaches) with regards to solution optimisation [17], thus they are the focus of this project. It must also be noted that this work is focused on exploring new techniques to performing service selection, but not service discovery, maintenance, and monitoring.

Besides identifying the atomic Web services most suitable to fulfilling key composition tasks, the selection process often takes additional user constraints into account. A survey of the literature in the area shows that two types of constraints are commonly taken into account. The first group comprises the creation of service compositions that are optimised according to the Quality of Service (QoS) constraints on its constituent atomic services, where QoS attributes may be thought of as features that indicate the quality of a given Web service, such as the time it requires to return a response and its financial cost of execution [17]. The majority of works in this area use maximising and minimising functions for the different QoS attributes, meaning that they attempt to obtain solutions with the best possible qualities without any thresholds [84, 64, 90, 26, 94, 81, 61, 39, 46]. However, there also exist approaches focused on what may be referred to as a Service Level Agreement (SLA)-aware optimisation, which is solutions must meet certain predetermined QoS thresholds in order to be considered valid (e.g. the financial cost of each service used in a composition must not exceed 50 dollars) [90, 89, 13, 6].

The second group comprises the creation of service compositions that have multiple execution branches, indicating user preferences at runtime [82, 10, 38, 55, 7]. For example, the output of a service *A* determines which service to execute next; if the output is greater than a certain threshold, then service *B* should be executed; otherwise, service *C* should be executed instead. These preferences are expressed logically as conditions, and affect the workflow used to establish the connections between services rather than the individual services themselves. These two types of user constraints are discussed in more detail in subsequent Sections in this chapter.

2.1.1 Web Service Composition in Practice

The research material published on Web service composition is highly theoretical and frequently employs layers of abstractions and simplifications intended to make the problem at hand manageable. However, it is also important to investigate how these ideas fit into practice, which is the objective of this Subsection. As mentioned earlier, while significant amounts of work are being performed in the field of automated Web service composition, these approaches ultimately remain on the theoretical level due to difficulties that have not yet been fully addressed. On the one hand, a study [52] has shown that the composition of services is fraught with issues. A central problem is that there are discrepancies between the concepts used by different services: the schemas used differ from each other, even if the services handle exactly the same domain. Another problem is the existence of services that produce too much data, low-quality data, or that incur too much latency. These characteristics may slow down the execution of the composition and even require human intervention, defeating automation efforts.

On the other hand, an automated framework for Web service compositions which can be applied to real services has already been proposed [54]. The functionality of this framework was demonstrated by creating a composition that uses the Amazon Virtual Cart service, the Amazon Book Search service and a bank's Point of Sale Web service. The authors of the framework point out that it is very challenging to find service documentation that is comprehensive enough, and that functionality details had to be identified from natural language explanations intended for developers who are working manually. Additionally, the specification of control flow requirements must be performed manually using a logic programming language. Despite these difficulties, the evaluation of this framework shows that a composition can be found within seconds when using it as opposed to an estimated 20 hours of manual development, showing that research on Web service composition provides some immediate benefits.

These two works provide opposing views on the viability of automated Web service composition, however a more recent approach [60] presents the interesting middle ground of user-centered design. This system combines manual and automated techniques, providing a browsing tool that allows users to explore the repository and gain some understanding of the offered QoS value ranges of services before having to write any QoS requirements, as users may often be unaware of what a reasonable QoS value would be for a given dataset. Users are also supported through the selection process by utilising a UI that diminishes their cognitive overload, and helps them express their requirements using a standardised language. In addition to these tools, this approach also proposes a clustering algorithm that groups service candidates according to their range of QoS values, with the objective of providing selection options when users have fuzzy (i.e. soft) requirements. The strength of this approach is that it does not undervalue human intervention in the composition process, instead providing tools that empower system users. As Web service composition always requires some degree of user involvement, at the very least when setting composition require-

ments, this type of approach may prove to become an increasingly popular composition solution.

2.2 Single-Objective QoS-Aware Composition Approaches

This section discusses composition approaches that optimise solutions according to their overall QoS. These approaches can be divided into biologically inspired methods, which use Evolutionary Computation to reach their goal, and general optimisation approaches that do not turn to evolutionary techniques. These two groups are quite distinct, but they have the commonality of using an objective function as the guide with which to measure the quality of the candidate solutions.

2.2.1 Biologically-Inspired Approaches

Biologically-inspired Web service composition approaches rely on evolutionary computation algorithms, which implement their search strategies by drawing inspiration from nature, namely the behaviour of social animals such as bees, fish, and ants. An important distinction between these different bio-inspired approaches is in their representations of the composition problem. These varying representations are discussed throughout this Subsection, and visually summarised in figure 2.2. **Genetic Algorithms (GA)** are a popular choice for tackling combinatorial optimisation problems, and thus have been widely applied to the problem of Web service composition [81]. The encoding scheme for a composition is commonly done as a vector of integers, where each integer corresponds to a candidate Web service, even though some authors have attempted to use matrix representations that also include semantic information about services. A population of candidates is evolved for several generations using generic operators, typically crossover and mutation: in crossover, equivalent sections of the vectors in two distinct candidates are swapped; in mutation, a section of one candidate's crossover is modified at random in order to introduce some genetic diversity. An observed problem with the GA technique is that it tends to prematurely converge to solutions, thus preventing the exploration of further possibilities. **Particle Swarm Optimisation (PSO)** bears similarities with GA, also relying on a vector representation for candidates. However, instead of employing genetic operators to carry out the search process, PSO uses the concept of position updates to move candidate particles across the search space. As PSO may also present the problem of not fully optimising solutions (i.e. converging prematurely), hybrid approaches have been attempted to improve its efficiency and optimisation power [81]. A key limitation of both GA and PSO is in the underlying vector representation used by candidates, since it makes it very challenging to encode workflow information and thus perform any type of fully automated Web service composition.

Ant Colony Optimisation (ACO) has also been proposed as a solution to QoS-aware Web service composition [94]. ACO is particularly suitable to a directed acyclic graph (DAG) representation, in other words, the workflow composition representation commonly used in the field. In ACO, the Web service workflow is built to be traversed by a group of ants (agents). At each fork in the graph, the ants choose which path to follow based on probabilities that take into account the strength of the pheromones left by other ants, and also a heuristic function for that particular graph. The pheromones left by the ants are updated after all ants have toured through the graph once, with paths of higher fitness resulting in a larger pheromone increment for the edges of those paths. Meanwhile, the pheromone level of all edges gradually decreases (i.e. evaporates) after each tour of the ants. The graph representation in this technique follows the abstract workflow idea, with a pool of concrete Web services associated to each abstract Web service. Each pool of candidates is a layer that

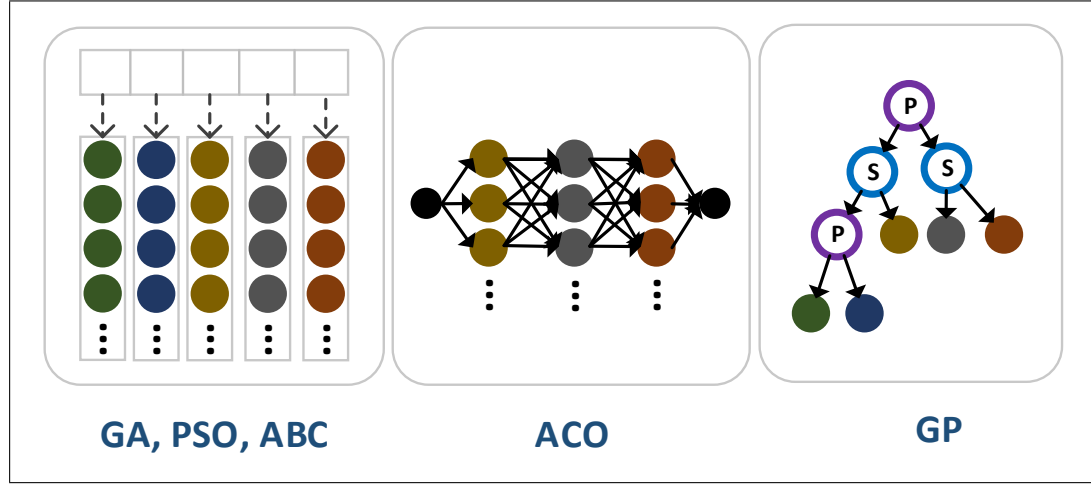


Figure 2.2: The different problem representations employed by biologically-inspired Web service composition approaches.

is fully connected to the layers of any following abstract services, so that an optimal path can be chosen from the edges laid out. For each concrete Web service, the heuristic factor is calculated based on its QoS values, and the fitness function also measures QoS attributes. As with GA and PSO, this representation also amounts to the idea of semi-automated Web service composition, even though in this case the encoding of workflow information leading to a fully-automated approach would seem to be trivial.

The works in [94] and [84] apply the **Artificial Bee Colony (ABC)** algorithm to Web service composition. The ABC algorithm simulates the behaviour of bees search for food sources. The position of food corresponds to candidate solutions in the search space, encoded as a service vector, and there are three types of bees dedicated to searching: *Employed bees* exploit the neighbourhood of a single food source already found; *Onlooker bees* exploit the neighbourhood of different food sources depending on the dance behaviour displayed by employed bees; *Scout bees* are the bees sent to random food sources after the neighbourhood they were previously exploiting does not present any food sources that are better than the original. The roles of the bees change according to the colony's needs, which is a feature unique to this algorithm. One of the issues with this approach, as pointed out by the authors, is the search space it explores. The problem with the typical search space for Web service composition is that it is organised based on the proximity of the components included into the composition. For example, given two adjacent solutions a and b in the search space, a will only have one service that is different from all the services included in b . Despite being neighbours, however, the fitness values of a and b may be radically different, and as the optimisation occurs according to a fitness function, this means that the search space is not entirely continuous. These works modify the traditional ABC algorithm to take this problem into account, filtering the neighbourhood of each solution during the search and excluding radically different neighbours from consideration.

Genetic Programming (GP) is, from this project's perspective, possibly the most interesting technique for the problem of Web service composition. That is because GP, unlike the previously discussed techniques, is capable of supporting fully automated Web service composition. In GP approaches [3, 72], workflow constructs are typically represented as the GP tree's non-terminal tree nodes while atomic Web service are represented as the terminal nodes. In this context, workflow constructs represent the output-input connections between two services. For example, if two services are sequentially connected, so that output of

service A is used as the input of service B , this would be represented by a sequence workflow construct having A as the left child and B as the right one. The initial population may be created randomly, in which case the initial compositions represented in that generation are very unlikely to be executable due to their mismatched inputs and outputs, or it may be created using an algorithm that restricts the tree structure configurations allowed in the tree to feasible solutions only. A fitness function is employed for the QoS optimisation of candidates, and the genetic operators employed for this evolutionary process are crossover, where two subtrees from two individuals are randomly selected and swapped, and mutation, where a subtree for an individual is replaced with a randomly generated substitute. One of the difficulties of tackling the problem of Web service composition using GP is that it does not intrinsically support the use of constraints [16], meaning that even if all candidates in a population meet the feasibility constraints, there is no guarantee that subsequent generations will maintain conformance to them. The approaches discussed above handle this problem in one of two ways: by *indirect constraint handling*, where feasibility constraints are incorporated into the fitness function so that the optimal function value reflects the satisfaction of all constraints, or by *direct constraint handling*, where the basic GP algorithm is adapted at the initialisation and genetic operation stages to ensure that the feasibility constraints are met. Indeed, the tree representation of an underlying workflow composition may pose difficulties whenever constraint verification is necessary.

Graph Variations of GP

Genetic programming using candidates constructed as graphs (instead of trees) would be ideal for the problem of Web service composition, since dependencies between services could be intuitively encoded. Even though variations of GP with graph candidates do exist, they have not been employed in this domain, therefore the focus of this section is on the techniques and not on the composition problem. Galvan-Lopez [24] proposes a general-purpose EC technique that is modification of the usual GP tree representation, allowing the creation of graphs. The key extension proposed here is the addition of pointer functions to certain non-terminal nodes, meaning that they can have connections to nodes in independent subtrees. Program inputs are provided as terminal nodes (similarly to the output-as-root representation discussed earlier), but instead of having a single output location at the root node, they may have other output locations as well, inserted as necessary amongst the non-terminal nodes of the tree. Since the overall tree structure is maintained, it becomes possible to perform the crossover and mutation operations similarly to their implementation in the case of a simple tree. The main difference is that any pointers present in the original tree may not be modified when these operations are applied. In order to ensure that the number of outputs in a tree remains correct throughout these genetic operations, each node in a tree is classified beforehand to establish which ones may be selected for the crossover operation. This representation makes it easier to evolve graphs, however it does not support strongly-typed GP, which is of interest to our research. In addition, each tree is still required to have a single root, meaning that there are connections between the different output nodes. In our problem domain, this is a hindrance because the output nodes must be completely independent from each other.

Miller and Thomson [59] present Cartesian Genetic Programming (CGP), a popular technique for evolving graph structures. The simplicity of SGP lies in the fact that it can represent the genotype of candidates as a string of fixed length, meaning that crossover and mutation operations are trivial to implement provided that they obey some simple constraints. The core idea of CGP is to create a two-dimensional array of programmable nodes of a predefined size. Each node has a predefined number of outputs, and the overall array has a

predefined number of inputs and outputs. Then, as this structure is evolved, the functions inside of each node can be reprogrammed, and so can the inputs those nodes require. From that point onward, the structure can be optimised according to the algorithm's fitness function. One important observation is that CGP may have unexpressed genes, meaning that not all the nodes in the two-dimensional array are necessarily components of the final answer. One limitation of this approach is that CGP does not easily handle strongly-typed GP, thus restricting the range of problems it can be applied to. Another problem is that it requires predefined numbers of nodes and node inputs/outputs, making it difficult to represent compositions with varying numbers of services and service inputs/outputs.

Mabu, Hirasawa and Hu [53] introduce a technique for the evolution of graph-based candidates, called Genetic Network Programming (GNP). GNP has a fixed number of nodes within its structure, categorised either as processing nodes (responsible for processing data) or as judgment nodes (perform conditional branching decisions), and works by evolving the connections between these fixed nodes. Connections are represented in a linear gene structure, and the number of outgoing connections from a node is dependent on the type of that node: processing nodes have a single outgoing connection, while judging nodes have more than one (depending on the number of branches desired). Because of the linear representation of connections between nodes, the genetic operators employed for the evolutionary process are quite simple. The mutation operator randomly chooses the destination of a node's outgoing connection; the crossover operator swaps two nodes with the same label from two different solutions, taking their outgoing edges with them. Li, Yang and Hirasawa [48] extend the basic GNP idea by using the Artificial Bee Colony (ABC) approach to evolve candidates. While these approaches present the advantage of simple genetic operations, the number of nodes and outgoing edges per node must be fixed throughout the evolutionary process, meaning that it suffers from the same limitations discussed above.

Poli [67] presents an EC algorithm where solutions are represented directly as graphs. This graph is mapped to a grid, where each row represents a layer of the graph, and columns can be used freely to accommodate the nodes in a layer. By using this representation, it is possible to perform relatively simple mutation and crossover operations to a graph without compromising its syntactic correctness. For the crossover operation, a crossover point (node) is selected in each candidate, and all the ancestors of that node are identified. These two subgraphs of ancestors are then swapped, a process that may overwrite nodes from the other graph but that maintains any original edges. Whenever this swap causes parts of the subgraphs to be placed outside of the grid, these nodes are "wrapped around" their respective rows, meaning that they are moved to the other extreme of the row. For the mutation operator, an existing subgraph is selected and replaced with a newly created one, using the same general mechanism as the crossover. Once again, while this approach makes the implementation of genetic operations simple, it does not allow for correctness constraints (i.e. input-output connections) to be maintained. Additionally, this representation assumes that the duplication of nodes is acceptable, which may pose some problems in the domain of Web service composition.

Kantschik and Banzhaf [37] propose linear-graph GP, a structure that allows programs with multiple execution paths to be optimised using evolutionary computing. This approach is an extension of the simple linear program representation where each element of a vector contains one program instruction. In this representation, a program graph contains several nodes, and each of these nodes contains a portion of a linear program that may be executed according to preceding branching conditions. Within each of these nodes, there exists also a branching node that is responsible for determining which outgoing edge to follow (i.e. which execution path to choose after executing the current node). The branching functions contained in each of these nodes may read the values of variables manipulated

in the linear program portion that is located in the same graph node. The mutation operator in this approach may alter an individual entry within a linear program vector, the branching function within a branching node, or the number of outgoing edges of a given graph node; the crossover operator may exchange individual entries within linear vectors of two candidates, or exchange a group of contiguous graph nodes between two graphs. Despite allowing branching constructs, this approach is not useful to our research because non-sequential relationships between different Web services cannot be encoded into linear representations.

The works in [27, 11, 62] present a graph-based genetic algorithm that is used to evolve representations of molecules. Atoms are represented as nodes, and their bonds as edges. Two types of genetic operations are supported: mutation, which can be the appending or removing of a node and its connecting bonds, and crossover, where edges are removed from each candidate until each graph is divided into two disconnected subgraphs that are then reconnected to create new child candidates. The reason why these genetic operators can be used without compromising the structure of the molecule is that the only restriction when creating a new connection is the valence of a given atom (i.e. the number of bonds it can make), but bonds do not need to be directed edges and cyclic structures are allowed. In the Web service composition domain, however, the need for additional restrictions means that these genetic operators are no longer suitable.

2.2.2 Other Optimised Composition Approaches

Other methods exist for producing optimised Web service composition results in addition to biologically inspired approaches, and a subset of them is presented in this section. **Tabu search** [28] is a combinatorial optimisation strategy for identifying an optimised solution amongst a group of candidates, typically in problems where exhaustive search is prohibitively expensive. In Tabu search, an objective function (either linear or nonlinear) is used to measure the goodness of solutions, encouraging solutions with the least penalty (i.e. optimal solutions). Then, a range of moves that lead from one candidate solution to another is defined. For a particular candidate solution, there is a set of moves that can be applied to it, and this is known as the neighbourhood function. One of the biggest advantages of Tabu search is that, unlike the hill climbing technique, it can avoid being stuck at local optima when searching. The technique keeps track of a set of tabu moves, which are moves that violate a given collection of tabu conditions. The objective of having a tabu set is to prevent the search from reaching solutions whose best next move has already been visited (i.e. prevent cyclic search moves). Due to its relative simplicity and its flexibility of implementation, tabu search has a wide range of practical applications for combinatorial optimisation problems. For example, a technique that combines Tabu search and a hybrid Genetic Algorithm (GA) implementation has been proposed recently [64]. For the Tabu search component of the technique, a move is defined as a change in one of the concrete services used as the solution, and the Tabu set is a fixed-size set of the latest n solutions visited. For the hybrid GA, the same basic idea of the traditional GA is applied (set candidate sizes, two-point crossover, mutation that randomly chooses another concrete service), however a local improver is also employed. This improvement process randomly explores a percentage of a solution's neighbourhood. A drawback of using Tabu search for Web service composition is that the structure used for the representation of candidates is typically linear [64, 5], which restricts the problem model to address semi-automated compositions only.

Integer Linear Programming (ILP) has also been applied to composing Web services [90]. ILP is flexible in the way it represents problems, therefore a fully automated Web service composition approach that also takes non-functional attributes into account when

constructing the best solution can be solved using it. An objective function is defined for the achieving the optimal QoS values, and several other functions are used to restrict the functionality of the solutions (i.e. restrict the search space). For linear programming, we first determine the "corners" of the restricted search space (i.e. where two constraint lines meet), and then apply the objective function to each of these solutions. One of these "corner" solutions is the optimal one, provided that all boundary functions are linear, so the best objective function score indicates the final solution. While ILP applied to Web service compositions is guaranteed to find an optimal solution, it has been shown to be very inefficient in comparison to EC-based approach as the complexity of the problem increases [3]. Another problem with the ILP approach is that it is likely to pose difficulties when used for creating compositions with multiple execution paths, as branching constraints would have to be represented using a linear function.

Structural Equation Modelling (SEM), a statistical analysis method for forecasting values according to current measurements, has been used in a service selection strategy that considers the future trends of the QoS values of the candidate services [46] as opposed to relying on static QoS snapshots. By being able to predict the QoS trends of services, it is possible to create a composition that is optimal at execution time. One key advantage of SEM is that it is capable of accommodating changes in the user preferences (weights) associated with each QoS attribute, and can use the errors in QoS measurement histories to create a forecast model. The proposed approach was tested using different composition algorithms (e.g. ILP) enhanced with SEM, with a dataset that simulates the changes of QoS parameters over time. Results show that the prediction model for QoS values is initially quite inaccurate, but the accuracy increases for both algorithms as days go by and a history of values is built. While this method is very effective at predicting QoS values to aid in the optimisation of solutions, it cannot be used alone to compose Web services. Thus, SEM is not considered further in this work.

Finally, **Algebraic Expressions (AE)** have been employed to the problem of Web service composition [23, 39]. A formal representation of a Web service composition is used in this approach, relying on algebraic constructs to describe the behaviour of atomic Web services and to constraint the characteristics of a correct composition solution. One of the main advantages of AE is that this technique is expressive enough to emulate the behaviour of Web service composition languages such as BPEL4WS, thus it is possible to design and verify composition solutions entirely through AE. A more flexible composition option, explored in [23], involves constructing a mapping between algebraic expression and BPEL, to allow for an automated translation between these two representations. The work in [39] goes further, proposing a composition algorithm that also performs QoS optimisation based on algebraic expressions. Despite promising, the disadvantage of this technique is that it requires the composition task to be described in formal terms that are challenges to system users without the necessary background.

2.3 Multi-objective, Top-K, and Many-objective Composition Approaches

The Web service composition techniques discussed up to this point have the commonality of being single-objective approaches, meaning that they optimise composition solutions based on a one maximising or minimising function. In order to produce a single numerical score while at the same time accounting for multiple QoS attributes, the majority of the objective functions employed by those approaches employs a Simple Additive Weighting (SAW) [33] technique, meaning that all individual QoS attribute scores of a composition are summed

after having been scaled by weights that reflect the importance of each quality feature. The objective function used in [17] is a classic example of this SAW technique:

$$fitness_i = w_1 A_i + w_2 R_i + w_3 (1 - T_i) + w_4 (1 - C_i) \quad (2.1)$$

$$\text{where } \sum_{i=1}^4 w_i = 1$$

Despite being frequently used in Web service optimisation problems, the SAW technique presents a fundamental limitation: it does not handle the independent and often conflicting nature of the different QoS attributes very well. For example, consider the trade-off between a composition's financial cost and its execution time. Services that have been implemented to respond after very short execution times are likely to be financially more expensive and vice-versa, a trade-off that is not well represented by the SAW model. To overcome this limitation, researchers have developed techniques that allow each QoS attribute to be optimised with an independent function, creating a set of candidate solutions that show the various quality trade-off amongst promising solution candidates.

Multi-objective (MO) techniques have been extensively employed in QoS-aware Web service composition [50, 94, 92, 89, 88, 13]. The basic idea is to optimise solutions according to a series of independent objective functions that measure the different QoS attributes to be considered. Candidates are then compared based on whether they *dominate* each other, that is, whether the quality scores of one candidate are clearly superior than those of another. For example, imagine a scenario where two composition candidates *A* and *B* are evaluated according to their execution cost *c* and time *t*. If *A* has (*c* = 3, *t* = 1) and *B* has (*c* = 4, *t* = 1) we say that *A* is a *dominant solution* in relation to *B*, since *A* has the same execution time and lower execution cost; however, if *A* has (*c* = 3, *t* = 1) then we say that it is a *non-dominant* solution, since its execution time is longer despite having a better execution cost. When comparing a population of candidates, multi-objective techniques produce a set of globally dominant solutions (but non-dominant amongst themselves) called a *Pareto front*. A Pareto front is useful because it presents a set of solutions with quality trade-offs between them, allowing the composition requestor to make the final choice. Initial approaches to multi-objective Web service composition and selection have focused on employing versions of GA [50], however other EC algorithms have also been considered. The work in [94], for example, proposes the use of a multi-objective Ant Colony Optimization (ACO) algorithm for QoS-aware Web service composition, an algorithm that is particularly suitable to an directed acyclic graph (DAG) composition representation. ACO generally works by building a graph of components (in this case, Web services) that is then traversed by a group of ants (agents). At each fork in the graph, the ants choose which path to follow based on probabilities that take into account the strength of pheromones left by other ants. This approach was tested against a multi-objective GA technique using the same fitness function, with results showing that the GA approach converges faster but requires a population 6 times larger than the ACO approach.

HMDPSO, a hybrid multi-objective discrete particle swarm optimisation (PSO) algorithm [89], is another interesting example of MO applied to Web service composition. The type of composition proposed in this work is SLA-aware, meaning that for the same composition solution there are different user levels with distinct SLA (quality) needs. Each particle is represented as an array that contains concrete candidates, and is divided into multiple parts, each part representing a solution for a different user level. While the paper does not explicitly explain why solutions for multiple user levels are combined in a single particle, it is thought that this is done to reduce the computational expenses associated with MO algorithms. The multi-objective PSO algorithm proposed in this paper updates the positions

of particles through the use of crossover operators, and also employs a mutation strategy to prevent stagnation in local optima. The fitness function is responsible for verifying the dominance of a solution over others, and a domination rank is calculated for each solution, with a value of 1 indicating that a solution is non-dominated and a higher values recording the number of dominating others for a given solution. Experiments were conducted to compare the performance of the HMDPSO to that NSGA-II (a GA-based MO algorithm) for the same composition tasks, and results show that HMDPSO with local search is superior to NSGA-II for all objectives and all SLA levels. This is thought to be the case in part because of the more granular fitness function employed by HMDPSO, which differentiates the domination rank of each candidate.

A problem related to multi-objective optimisation is that of selecting the **Top-K** best solutions to a composition problem. According to publications in this area [93], full reliance on multi-objective techniques may not be useful in practice because there are too many possible solutions in the resulting Pareto front, and also because it is difficult for requesters to translate their QoS preferences into weights. Due to this problem, techniques to restrict the number of individuals in the solution set must be applied. Besides using a top-K approach, a ranking of QoS attributes can also be compiled by users to describe the QoS values that are the most important to them [93]. This ranking is used to compute a preference-aware skyline, where a service is only part of the solution set if it is not dominated by another service on a specific number of preferences (this number does not necessarily need to correspond to the full number of QoS attributes considered). The work in [19] is an example of applying a Top-K approach to fully automated composition in large-scale service sets. According to the authors, the main advantage of providing K composition solutions instead of only one is that multiple options are available in case of network failures or similar problems. The issue of large-scale service sets is dealt with by employing a multi-threaded solution that is capable of computing candidates in parallel. Before identifying the Top-K solutions, two preprocessing steps are executed: in the first step, candidate Web services are indexed according to the concepts produced by their outputs (following semantic relationships); in the second step, services that are not useful for the composition task are discarded from consideration using a filtering algorithm. Then, the solutions are identified by employing a MapReduce framework in conjunction with graph planning techniques to compare all potential candidates.

The performance of multi-objective algorithms tends to degrade when optimising solutions according to more than three quality criteria [18], and thus alternative techniques must be investigated to handle such scenarios. These are known as **many-objective techniques**. The work discussed in [18] is based on NSGA-II, an algorithm that allows for the independent optimisation of different quality attributes, and it compares this algorithm to two other techniques: *Maximum Ranking (MR)*, where each solution is ranked for each objective according to its value, from highest to lowest, and the solution's highest ranked objective is used as its overall rank; and *Average Ranking (AR)*, where the ranking values for each objective are computed and added, and each candidate is then ranked according to this aggregated value. The experiments conducted using these techniques show that NSGA-II has the worst performance overall for larger composition tasks when considering measurements such as generational distance (i.e. the distance of solutions from the known pareto set), hypervolume (i.e. the are covered by the solutions in the Pareto set), and spread (i.e. the distribution of the solutions within the Pareto front) [4, 35].



Figure 2.3: Basic example of Web service composition using the Graphplan algorithm.

2.4 AI Planning-based Composition Approaches

AI planning approaches to Web service composition ensure feasibility by building a composition solution step by step. This solution is represented as a graph, and may either be built to enforce a set of user constraints, or it may be used to find an optimal solution in terms of QoS. A number of works in this area [22, 79, 86, 82] have been based on a fast planning algorithm named Graphplan [8]. In this algorithm, a solution is constructed gradually, at each time adding a new atomic service to the composition. A service may only be added to the solution if all of its conditions are met, that is, all of its inputs are fulfilled. Finally, the execution of Graphplan is stopped once an atomic service has been added that leads to meeting the overall composition objectives (i.e. the composition now produces all of the required outputs). Figure 2.3 shows a basic example run of Graphplan applied to Web service composition. In step 1, a start node is added to the graph structure. This node produces the overall input values provided by the composition requestor, *ZipCode* and *Date*. In step 2, the service *LocationByZip* is connected, since its input of *ZipCode* can be fulfilled by the existing structure. However, the algorithm continues to execute, since the overall output has only been partially fulfilled (i.e. only *City* can be produced). Finally, in step 3, the service *Weather* is connected, since its inputs are fulfilled by both the start nodes and the *LocationByZip* service. As the overall output has been fulfilled, the graph's end node is also connected to the structure.

In [12], authors combine a planning algorithm and a graph search algorithm to achieve both QoS optimisation and feasibility in Web service compositions. The generic Graphplan algorithm first builds a representation of the search space as a planning graph, then finds a solution within this graph by traversing it backwards. This standard planning approach is modified to use Dijkstra's algorithm [75] when performing the backwards traversal, thus finding an optimised solution. The planning graph is extended to include labels associated with each proposition (i.e. each intermediate action between two vertices), where each label contains a layer number and associated execution costs. Dijkstra's algorithm is used to calculate the upcoming costs of each node in the graph. Then, a backtracking algorithm uses this information to determine the optimal solution.

The work in [20] proposes a planning-graph approach to creating a Web service composition technique that is capable of identifying the top K solutions with regards to overall QoS values. According to authors, AI planning is highly suitable to the domain of service composition, however many planning approaches are not efficiently executed. A notable exception to this trend is the planning-graph method, where the search space is greatly reduced through an initial search, thus allowing the remainder of the algorithm to focus on the

relevant areas of the search space. In this work, the planning-graph approach is employed in a three-part algorithm. In the first part, a forward search is executed from the input node aiming to find the output node, gradually filtering the services that could be used in the composition. In the second part, the optimal local QoS of each service in the remaining graph is calculated using functions that take into consideration the QoS of the services that could possibly feed its input, also executed from the input node towards the output. Finally, a backward search algorithm is executed to generate the top K solutions according to local values calculated in the previous step (a threshold is provided when running this algorithm to prune out the substandard composition options).

An automated Web service composition approach that uses a filtering algorithm to reduce the number of services considered for the composition, organising the remaining services as a graph according to the ways in which their inputs and outputs match, is proposed in [32]. Once the graph has been determined, a modified dynamic programming approach is applied to it in order to calculate the composition with the optimal QoS. Dynamic programming is a method that breaks problems into smaller subproblems that are then solved, ultimately leading to the solution of the parent problem. In this case, the optimal QoS of each atomic service in the graph is calculated, taking into account its input dependencies. At the end of this process, the overall optimal QoS values are known and the subgraph containing the solution can be extracted by searching the graph backwards. Experiments with the WSC2009 dataset show that the algorithm has good execution times for various dataset sizes, demonstrating its scalability. This work was extended in [34], with the presentation of a composition tool called QSynth, and the performance of formal comparisons with other state-of-the-art approaches, also producing superior results.

A number of other works in the area employ formal AI planning techniques and frameworks to create compositions [7]. The work in [76] presents an approach to include user preferences into the process of Web service composition. This is accomplished by relying on a framework written in Golog, a language created for agent programming. Golog is used to specify the particular attributes of generic workflows that represent commonly requested composition procedures (an example of a generic workflow would be one that is dedicated to booking inter-city transportation). The syntax of a logic-based language used to specify user preferences is described, allowing for branching according to conditions, and for expressing preferences over alternative services. Despite supporting branching, only one set of final outputs is allowed, meaning that the branches must be merged before reaching the end node of the composition workflow.

An approach for modelling the data flow between Web services through the use of *domain objects* is presented in [40]. For example, the travel domain contains objects such as flight ticket, hotel reservation, etc. The key idea is to use these objects to connect the composition needs to the services that can address them. In order to do so, authors annotate how each Web service operation relates to a given domain object. By creating this abstraction layer of objects, it is possible to reduce the composition's dependency on implementation details for correct execution. Now services can be thought of as having an object port proxy that leads to specific service ports. Compositions can then be achieved by identifying the necessary domain objects for the required task. The paper goes on to show how this technique can be integrated into existing service composition techniques, in this case AI-planning based, through the creation of a formal framework. This framework was implemented and run with a virtual travel agency scenario. According to the authors, the implementation of such framework was not trivial, however it successfully demonstrates that service implementations can be modified without impacting the overall composition.

A Web service composition approach that allows users to specify constraints on the data flow of the solutions (i.e. which routes a message is allowed to take and which manipula-

tions it can undergo) is presented in [55]. For example, consider a Web service composition that is supposed to book a holiday for a customer using a flights, accommodation, and a map service. If it is possible to book a suitable flight but it is not possible to book a hotel, the customer should not accept the offer. This is the type of requirement addressed by this approach using a data flow modelling language. This is a visual language that supports the definition of inputs/outputs, forking messages, merging messages, operations on messages, etc. By connecting these elements we obtain *data nets* whose satisfiability can be clearly verified. The composition of Web services is performed using a planning framework that is capable of interpreting and respecting the constraints of a data net. At the time this paper was written, this approach had not yet been implemented or tested.

2.4.1 Hybrid Approaches

Hybrid approaches combine elements of AI planning and optimisation techniques for solving the composition problem with functionally correct, optimised solutions [15, 70, 87, 14]. These hybrid approaches are quite similar to each other, relying on a directed acyclic graph as the base representation for a candidate solution, and then applying the optimisation techniques to this structure. However, despite incorporating the use of planning techniques, they do not include any discussion on the issue of producing solutions that satisfy user constraints or preferences. Another commonality between these works is that they require the use of SAWSDL-annotated datasets for testing, but these are not widely available to the research community. Therefore, authors developed their own datasets, and utilised each dataset's optimal task solution as the benchmark with which to evaluate the success of their implementation. More specifically, authors calculated the percentage of runs that culminated in the identification of the global optimum as the recommended solution.

In [68], an approach that combines AI planning and an immune-inspired algorithm is used to perform fully automated QoS-aware Web service composition, also considering semantic properties. One significant contribution of this work is the proposal of an Enhanced Planning Graph (EPG), which extends the traditional planning graph structure by incorporating semantic information such as ontology concepts. Given this data structure, the composition algorithm selects the best solution configuration from a set of candidates. A fitness function considering QoS values and semantic quality is used to judge the best solution, and a clonal selection approach is employed to perform the optimisation. Candidates cells (solutions) are cloned, matured (mutated by replacing services with others from the same cluster in the EPG) and the cell most suited to combating the invading organism (i.e. the best solution) is discovered.

The work in [69] proposes the employment of the Firefly meta-heuristic technique for performing QoS-aware Web service composition, in conjunction with an AI planning strategy that uses an EPG as the basis for solutions. The firefly meta-heuristic is based on the behaviour of mating fireflies, which emit a flashing light to attract potential mates. Each artificial firefly investigates the search space, with each position representing a composition solution. The brightness of the firefly is represented by the fitness of the current solution (location) associated with it. Fireflies are attracted to others according to their brightness, which varies with distance. Finally, fireflies move towards the individuals they are attracted to, meaning that small modifications occur in the current solution. The fitness function takes into account the QoS attributes of the composition.

2.5 Semantic Composition Approaches

QoS-Aware Semantic Service Selection: An Optimization Problem [26]

This work uses a QoS ontology as the basis for the proposal of a framework that selects services according to user preferences, representing this scenario as an optimisation problem and applying a variety of techniques to solve it. Since the ontology is represented as an XML document, XSL transformations are used in order to create the initial representation to the optimisation problem, allowing a common ontology to be used for a variety of representations. Once a user expresses his/her preferences, the framework converts this information into a set of utility functions that act as optimisation constraints, and the main optimisation function maximises/minimises the quality as required. This framework was not implemented or tested in this paper.

A Semantic Selection Approach for Composite Web Services using OWL-DL and Rules [9]

The motivation of this paper is to present a selection strategy for Web services that considers more than just functional (e.g. input and output) and non-functional (e.g. QoS) attributes for Web services. This paper addresses semi-automatic service composition approaches, i.e. those approaches where an abstract workflow of services has already been provided. In order to consider more than just functional and non-functional attributes, this work associates objects with each service that contain additional information useful for the selection process. These objects have an independent ontology that describes how they interrelate (e.g. a service for making medical appointments has associated objects such as doctor, patient, hospital/clinic, etc). When performing a composition, these objects must be compatible. To accomplish this, a framework with service providers, ontology providers, information agents and a composer is proposed. This framework takes selection constraints set by the user into account. A prototype was implemented to evaluate this framework, and results showed that the proposed approach scales better with regards to execution time than the naive approach (does not check constraints before considering possible compositions) as the number of candidate services increases. That is because it is able to filter out candidates that do not conform to the previously defined constraints. A possibility for future work is the combination of this semantic selection approach with an optimisation approach such as PSO or ACO for performing the concrete composition.

Genetic Algorithm for Context-Aware Service Composition Based on Context Space Model [95]

This paper proposes a composition approach in which the context of services is taken into consideration. By context the authors mean inter-relationships and conflicting relationships between services. Each group of service candidates has an associated group of rules that state these relationships. The fitness function is defined as a multi-objective function that takes into account the length of the service sequence and the distance between the current solution and the target solution (?). The genetic operators take relationships between services into account, and rely on functions that resolve any conflicts. Experiments indicate that better fitness can be obtained by running on larger datasets, and for solutions that have length of between 12 and 16 atomic services. The paper is quite brief, so examples of the relationships between services are not provided.

Bringing Semantics to Web Services with OWL-S [56]

This paper presents and explains the OWL-S (Web ontology language for Services) standard. OWL-S was created to address a limitation with WSDL, which is its lack of a formal description of the steps involved in using a Web service. This means that the only way to obtain such information is by reading natural language documents, and that automating tasks that depend on this information is difficult. To address this need, OWL-S provides a formal specification of the workings of a service. This standard also allows for the association of classes with each Web service, and supports the use of ontologies that record the

relationships between members of the different classes, thus establishing a common vocabulary to be used in inter-service interactions. These features are conducive to automating the handling of Web services, and facilitating the discovery of those that are relevant for a specific task.

OWL-S offers three distinct ontologies. The first one is the *service profile*, which describes what can be accomplished by the service. The goal of this ontology is to facilitate advertisement of a service's capabilities, and consequently its discovery. Information at this level includes required inputs and outputs, Quality of Service attributes, and the class that service belongs to. The second one is the *process model*, which records how the service should be used from a conceptual point of view. The focus of this ontology is in describing the patterns of interaction that apply to this service, and these can be atomic processes (single input-output interchange), or composite processes (a set of atomic processes orchestrated by using flow control structures). The third one is the *grounding*, which describes how users should interact with the service at a concrete (i.e. implementation) level. This ontology maps the the OWL-S atomic processes to their corresponding WSDL operations, providing a translation between the semantic and the implementation levels. One limitation of OWL-S is that it is not expressive enough to encode preconditions and effects, thus requiring the additional use of languages that are capable of conveying first-order logic. Another problem is that the ontology manipulation languages offered by the OWL-S standard are not intuitive to human users. The paper concludes by presenting several technologies that were developed based on the OWL-S standard.

Web Services Composition Handling User Constraints: Towards a Semantic Approach [10]

This paper presents a theoretical plan for a framework capable of defining user constraints utilising a semantic approach. It plans to address what are called hard constraints, including service properties (e.g. service A must have a specific property with value V) and choreography details (e.g. a particular service B must follow service A). The key idea is to implement a multi-layer framework that allows users to specify a list of abstract services in their semi-automated composition, as well as a list of constraints. Based on those constraints, suitable concrete candidates are found for each abstract Web service provided. The composition algorithm was not yet finished at the time the authors wrote this paper, so there are no concrete results to support the validity of this design.

Composing Semantic Web Services Under Constraints [38]

This work focuses on performing Web service compositions by modelling them as a constraint satisfaction problem, providing particular user constraints. A crucial point of this work is that it utilises the semantic information associated with each atomic Web service to verify that the user constraints, also semantically defined, have been met. A framework is used for defining this composition approach, based on the semi-automated composition paradigm. The user begins by selecting the appropriate abstract workflow with which the composition will be performed. Once this has been determined, an engine is employed to identify suitable concrete candidates for each workflow slot. Finally, information is gathered about each of the candidates and a plan-generating module is invoked to selecting the services that should make up the composition, taking into consideration the constraints defined earlier. This generator identifies the optimal constrained solution and translates it into an executable format (BPEL). Details are provided on the syntax used for modelling constraints and translating them into the constraint language used by the plan-generating module. This approach was tested using different candidate set sizes, different of abstract workflow sizes, and different numbers of user constraints. Results showed that the execution time for larger numbers of candidates only increases slightly (0.06 seconds are added to the execution time for a 100-fold growth in the number of candidates); for a larger workflow, the time increased linearly in relation to the growing complexity of the structure; finally, the

execution time also increases linear as the amount of defined constraints grows.

A Framework for the Semantic Composition of Web Services Handling User Constraints [25]

This paper proposes a framework for performing semantic Web service composition that allows user constraints to be specified. Initially, services are grouped into distributed "service communities" according to their OWL-S semantic descriptions, where each community has services that cater for similar domains and consequently present similar functionalities. Then, when users need to request a composition they formulate their need, including the necessary constraints, using terms from the semantic service community descriptions. Effectively, they create an abstract workflow for semi-automated composition by using the community descriptions and specifying their own preferences for the services to be selected. These constraints may be restrictions in input value ranges, in the output value ranges, restrictions that apply to local parameters (?), or a combination thereof. In this approach, a language called KIF is chosen to express the constraints according to the corresponding OWL-S service descriptions. These constraints are simplified into mono-service (so that they only require one service to be checked at a time) and a group of local composer agents (one for each community – i.e. slot in the abstract workflow) collaborate to compute the concrete composition, verifying locally that the constraints can be satisfied. This approach can also handle world-altering services and non-deterministic behaviour because it makes use of statecharts to model the behaviour of each Web service. However, the framework has not been tested and it is not clear how the composer agents could be implemented.

SAWSDL: Semantic Annotations for WSDL and XML Schema [43]

This paper discuss SAWSDL, a standard for semantically annotating WSDL schemas. The advantage of having such a standard is that semantic information facilitates the automation of processes that handle Web services, such as discovery and composition. SAWSDL builds on WSDL by embedding pointers in the WSDL description that refer to semantic concepts, and as a W3C standard it is meant to be used as the common ground between all of the existing semantic Web service frameworks (e.g. OWL-S). These pointers are called annotations, and point to concepts in a higher-level ontology and semantic description layer. More specifically, SAWSDL works by making two types of extensions to WSDL:

- *Model references*: an attribute that is added to any WSDL element to create a reference to one or more semantic concepts. The corresponding concepts are listed as the values of this attribute.
- *Schema mappings*: attributes used for lowering semantic data to a non-semantic equivalent, and also for lifting non-semantic data into a semantic equivalent. For example, this may be used for passing messages from a Web service to a client software. In this case, the Web service will lower its semantic data into a message, and the client software will lift the semantics from that message later.

Evidently, a concrete service ontology must be defined for these references and mappings to work. SAWSDL supports the creation of two types of semantics describing the different aspects of services:

- *Information semantics*: an ontology for instance data, typically used when mapping data to concepts.
- *Functional semantics*: an ontology for the capabilities provided by services, providing information regarding the invocation of a service. Typically used during service discovery.

A DAML-Based Repository for QoS-Aware Semantic Web Service Selection [77]

This paper proposes an approach to select services based on QoS attributes that are semantically represented. In order to do so, an ontology of service categories and quality attributes is created, where services are organised according to classes in a tree, and each class has its respective quality attributes. In this representation, the root of the ontology is a general service, and its attributes are those classically discussed in literature (price, cost, time, availability, etc). Lower ontology nodes, on the other hand, represent more specific classes of services that consequently contain more specific measures of quality. For example, a service class dedicated to vehicle rental might have the best rental rate as one of its measures. A language called *dvQL* for performing semantic queries for service selection is proposed, based on DAML (a semantic language for the web). *dvQL* allows for the creation of queries including service category and attribute value information. While *dvQL* presents some limitations with regards to metadata handling, it is expressive enough to enable the construction of meaningful expressions. The implementation of this semantic service repository used a relational database with an in-memory ontology, and translated the queries to SQL.

A QoS-Aware Selection Model for Semantic Web Services [83]

This paper presents an approach to Web service selection that employs an ontology of QoS values represented using WSMO (Web Services Modeling Ontology). The objective behind the use of this ontology is to allow the definition of more specific quality metrics, in addition to the those commonly used (time, cost, etc). The ontology supports a hierarchy of services with associated QoS values, and a taxonomy of terms that are used within this hierarchy. Any metric can be defined and associated to a Web service by providing a description including the metric name, its value type, measurement unit, and many other details. At selection time the user expresses his/her requirements about a service, including specific quality properties and whether they are necessary or optional, and that information is matched against the profile of each candidate service (which includes functional and non-functional information). Then, the process for filtering candidates is run in two stages, first matching services according to their functional aspects and then according to their quality attributes. The algorithm used for matching services according to qualities creates an array of quality values for each candidate, with all entries normalised according to the minimum and maximum values observed within the candidate set. For each quality value, a ratio of how closely it matches the user's requirements is calculated according to the requirements (the user either wants the smallest possible value, the largest possible value, or the closest possible match to a given value). Finally, the overall evaluation result is obtained by summing all of these proximity ratios. Experiments to test this approach were conducted using a hypothetical dataset in the domain of telephony, showing that the selection algorithm is in line with human preferences and intuition.

A Four-level Matching Model for Semantic Web Service Selection Based on QoS Ontology [30]

This paper proposes a semantic Web service selection method that performs matching based on four distinct levels:

1. *Application Domain Matching*: identifies the domain that best matches the user request through the use of category ontologies, and retrieves a list of potential service description candidates. This is performed by calculating a similarity degree between the user request and the semantic information associated with each domain.
2. *Description Matching of Service*: based on a given domain ontology, vectors are created for each potential service candidate description and another vector is created to represent the user requirements. A Vector Space Model is created, employing cosine similarity and TF-IDF to select the best service description.
3. *Function Matching of Service*: information from service providers is compared to the

service description using a similarity measure, and a set of all services whose functionality fulfils the requirements is returned.

4. *QoS Matching*: a matching matrix is created, similarly to what is explained in [83], and the optimal candidate is returned.

Web Service Composition with User Preferences [49]

This paper presents a semantic Web service composition approach in which user preferences may be defined as part of the composition task. An important distinction is made between user constraints, where unless all rules are met a composition solution cannot be produced, and user preferences, where solutions are produced even if not all rules are satisfied (in this case, the best solutions are those that meet as many rules as possible). A planning approach is adopted in this work, requiring the modelling of user preferences using the PDDL3 language, and the availability of service descriptions in the OWL-S standard. These representations are then translated into a formal representation that is compatible for use with a Hierarchical Task Network (HTN – a technique for automated planning). SCUP, which is the planning algorithm for satisfying as many user preferences as possible, is presented, and it generally consists of checking each step of a solution solution including all possible constraints, then progressively discarding the constraints that cannot be satisfied from consideration. Results of tests on a prototype of SCUP are presented, and they show that SCUP is capable of violating less user preferences than SGPlan (a state-of-the-art technique for preference-based planning) for the same tasks and datasets.

An Automated Composition of Information Web Services based on Functional Semantics [74]

This paper points out that many Web service composition approaches present the limitation of only taking inputs and outputs as a measure of compatibility between services. To overcome this problem, it proposes a method that also considers the semantics of services and the data types accepted by their operations, thus improving the ability of compositions to present semantic correctness. In this method, services are classified according to a functionality tag consisting of an {action, object} pair (e.g. a service for calculating the distance between two cities has the tag {Calculate, Distance}). A service relation graph is created to illustrate the dependencies between concepts, and it is divided in three parts: a graph showing relationships between actions, another graph showing relationships between objects (contains input/output relationships), and a mapping between items in these graphs. The relationships between these items are determined using domain ontology trees, with the assumption that these trees have already been provided. Given these dependencies, an algorithm is used to find a composition path. The path is found through the action graph, and object connections are made based on the object graph (and not only object names). This approach was compared to a previously proposed approach, and it was shown to require substantially less time to execute for larger datasets.

World-altering Semantic Web Services Discovery and Composition Techniques - A Survey [73]

This paper points out that there are many approaches dedicated to discovering and composing semantically-annotated information-providing services (those services that do not change the state of the world), but the same is not true for world-altering services (those that do change the state of the world), primarily because accurately describing their preconditions and effects presents an additional challenge. It then presents a brief survey of techniques that do take this type of service into account. Many approaches annotate services with preconditions and effects so that discovery can occur taking these factors into account. Similarly, the authors identify composition techniques that consider the matching of preconditions and effects in addition to the usual matching of inputs with outputs.

Towards a Capability Model for Web Service Composition [47]

This paper points out that many papers have been published on techniques for achieving

Web service composition, but more work needs to be carried out in proposing ways for casual users to specify their composition preferences, as current approaches require domain knowledge and formal descriptions. It then puts forth a three-layer model that helps in extracting natural language-like user requirements, deriving constraints from these preferences, and generating a composition accordingly. The three layers are described as follows:

- *Objective Layer*: Allows users to specify each of the composition objectives by using a verb-noun pair (these words must be associated to an agreed-upon ontology).
- *Profile Layer*: This layer is created by service providers and contains a list of capabilities and attributes that specify what each Web service can achieve and its relevant characteristics. Attributes have names and values, while capabilities are also expressed using verb-noun pairs.
- *Composition Layer*: Records the different capability relationships between Web services. Four relations are supported: competition, cooperation, support and generalisation/specification.

A process is run involving all of these layers to discover the capabilities to satisfy each of the composition objectives, with the user selecting the appropriate capabilities that are initially discovered. Finally, Web services associated with those capabilities are discovered and used for the final composition. Future work should automate the capability matching process so it no longer requires human supervision.

2.6 Dynamic, Distributed, and Cloud Composition

Dynamic Service Composition [41]

This paper presents an approach for performing dynamic service composition. As opposed to static composition, where the services that compose the solution are selected before execution time and instantiated at runtime, in dynamic compositions the atomic services are selected at the time of execution. The advantage of dynamic composition is that it can use the current state of the service environment as opposed to relying on information retrieved at an earlier time. Its drawback is that it requires a composition algorithm that is fast enough to create a solution that satisfies a pending request. The problem of service composition is often modelled as a graph in which a path is to be found, and this approach is also widely employed by researchers focusing on dynamic composition. Forward and backward chaining approaches based on input/output matches are used for path exploration, often relying on heuristics to encourage the exploration of the most promising paths and to reduce the number of services considered. In the approach proposed by this paper, both forward chaining and backward chaining are employed simultaneously, with the intention of having their paths meet in the middle. By doing so, the number of branches to be considered is greatly reduced. Experiments were performed using reduced service sets (24 services and 19 services, as shown in the paper's appendix), and results showed that the bidirectional search implemented requires the exploration of a consistently smaller number of services when compared to the exclusively forward and exclusively backward approaches (in these exclusive approaches, the number of services explored varied greatly according to the dataset).

Facing Uncertainty in Web Service Composition [1]

This paper proposes a Web service composition preservation approach that results in solutions that are capable of evolving and adjusting themselves according to unexpected changes in the environment, removing the closed world assumption employed by many researchers in the area. For example, if one of the third-party services fails in a manner that has not

been foreseen, current compositions are not able to adapt even though it would be beneficial if they did. In order to maintain compositions functional in the face of unexpected events, tactics (strategies) must be defined for preserving the initial composition requirements, even though it is not known ahead of time which failure events will trigger which tactics. The paper proposes the use of an evolution layer that is capable of reconfiguring solutions according to tactics to prevent a set of requirements from changing. This layer constantly monitors the composition to determine if there are any unknown events, which requirements they affect, and which tactics should be carried out. A prototype for this system was implemented and tested, with results showing that its performance is acceptable and that it scales well. However, the effectiveness of this approach depends on how well the predefined tactics match the complexity of the environment the composition is being executed in.

Probabilistic Top-K Dominating Services Composition with Uncertain QoS [85]

This paper proposes a multi-objective Web service composition approach that takes into account the fact that QoS attributes are dynamic in Web services (this is called an uncertain QoS model). According to the range of changes in the QoS attributes of a particular service, it is possible to calculate the probabilities of one service dominating another in the Pareto set. Top-k domination refers to the ability to find the k solutions that dominate the highest amount of other solutions in the search space (a technique for doing this is presented in the paper). Future work includes the use of different QoS models to represent the dynamic nature of Web services, and to increase the efficiency with which composition services are retrieved.

A Novel Adaptive Web Service Selection Algorithm Based on Ant Colony Optimization for Dynamic Web Service Composition [80]

This paper presents a multi-objective Web service selection approach that optimises solutions according to two functions, a static one that calculates the overall quality of service (QoS) of a solution, and a dynamic one which calculates the *trust degree* of a solution at a given time. The trust degree is defined as the number of successful executions of a service over its total number of executions, information which can be obtained by analysing execution logs. The optimisation algorithm used in this approach is an adaptive form of ant colony optimisation (ACO), where pheromones are adjusted according to the trust degree (calculated anew at every iteration) and the QoS is used by each ant as the heuristic for choosing the next node to visit. Each node in the graph explored by the ants represents an abstract service, with multiple concrete service candidates which can provide that functionality. The algorithm works by generating a set of solutions and then identifying its Pareto subset by comparing all solutions. The Pareto subset is used in the next iteration, for updating the path pheromones. A case study is presented, comparing the performance of adaptive ACO to that of the standard ACO algorithm, with results showing that adaptive ACO has a higher accuracy percentage than the standard ACO.

Revenue Optimization of Service Compositions using Conditional Request Retries [6]

This paper proposes a retry mechanism in a Web service composition that optimises the cost of running its instances according to changes in its environment. Often, providers offering Web service compositions that accomplish a specific test have Service Level Agreements (SLAs) with users specifying performance guidelines that should be met. However, due to the dynamic environment in which Web services exist, breaches to the SLA may occur and result in losses for the provider. To counteract this, the authors propose the use of conditional retry mechanisms in compositions. The retry mechanism is executed when a given service does not provide a response within a specified deadline, according to a decision policy associated with each composition. A formal definition and formal testing are presented for the proposed mechanism.

TQoS: Transactional and QoS-Aware Selection Algorithm for Automatic Web Service Composition [21]

This paper proposes a semi-automated Web service composition approach that not only takes the functionality and quality of the services into account, but also their transactional properties. In this work, transaction properties are defined as the behaviour of Web services when interacting with one another. Knowing the behaviour of Web services is important for estimating how reliable their execution is and which ones might require recovery strategies. A service is considered *retrieable* if it can terminate successfully after multiple invocations, *compensatable* if there is another service that can semantically undo its effects, and *pivot* if its effects cannot be undone once it is executed but if it fails there are no effects. The system takes an abstract workflow and a set of user preferences as its inputs, where the user preferences contain QoS weights and risk levels corresponding to the transitional requirements for the composition. Then, a planner engine assigns one concrete Web service to each abstract slot of the provided workflow. Whenever a service is assigned to a part of the workflow, its transactional properties influence any subsequent services, thus the risk must be recalculated along with the overall QoS. Experiments were run for various risk scenarios, and computation time was found to remain under 2 seconds even for the largest dataset (comprising 3602 atomic Web services), at the same time meeting user preferences.

QoS-Aware Services Composition using GRASP with Path Relinking [65]

This paper presents an approach to QoS-aware Web service composition that is focused on *rebinding*, that is, deciding which concrete services to bind to each abstract task at runtime in order to take the current state of the environment into consideration. This approach considers global QoS constraints (e.g. the overall composition price must be lower than 5), local QoS constraints (e.g. the individual price for a service must be no higher than 1), and service dependency constraints (e.g. as many services as possible should be used from the same provider). Two techniques are employed during the composition process: GRASP, which is used to construct initial solutions, and Path Relinking, which is used to perform improvement on these solutions. GRASP (Greedy Randomized Adaptive Search Procedure) iteratively builds a *valid* solution vector, adding one candidate to fulfil each solution slot at a time and ensuring that this candidate respects the pre-established user constraints. The order of candidate addition matters, so it is performed randomly each time. Once a valid solution has been built, GRASP identifies a list of replacement candidates for each task slot, including the most promising candidates while also respecting the constraints observed by the valid solution. Path relinking explores the neighbouring solutions of the initial valid solution, seeking to further improve it. To do so, it slowly modifies solutions by changing services from one task slot at a time. The objective function used in this paper encourages the improvement of QoS values at the same time it enforces the relevant user constraints. The proposed approach was compared against a hybrid tabu search/simulated annealing method and against GA, using generated datasets containing up to 252 services and having the stopping criteria be the execution time. Results showed that the proposed approach produces solutions with a significantly better fitness the majority of the time, but the TS/SA hybrid surpasses it for highly constrained problems.

CSCE: A Crawler Engine for Cloud Services Discovery on the World Wide Web [63]

This paper presents a crawler for discovering cloud services available on the web, an area that has not been widely researched. The difficulties in cloud service discovery come from the fact that standards for publishing and for describing cloud services have not yet been agreed upon, and no cloud service registries are widely utilised. Due to its comprehensive architecture comprising of discovery layers, the crawler engine encountered thousands of cloud services on the web, thus allowing the characteristics of current cloud services to be analysed. Analysis on the retrieved results shows that services are commonly placed in one

of the following three categories: IaaS, PaaS, and SaaS. The majority of service providers offer exclusively IaaS, and only 7% focus on PaaS. One important finding is that the great majority of cloud services is not implemented using Web service languages such as WSDL, and that formal descriptions are generally lacking. This illustrates the need for the development of standards for cloud services.

Cloud Computing Service Composition: A Systematic Literature Review [36]

This paper provides an overview of the approaches published between 2009 and 2013 on cloud computing service composition. This type of composition is very similar to Web service composition, however it also takes into account the characteristics that are unique to cloud services: their dynamic contracting/expanding nature, the increased security they require, the conflicts that may exist between different services, etc. The current solutions found in the literature can be divided into five groups:

- *Classic and graph-based algorithms*: includes linear programming approaches, backtracking algorithms (incrementally build candidates to solutions, abandoning partial candidates), and methods with multiple processing layers.
- *Combinatorial algorithms*: includes the use of perceptrons and neural networks, genetic algorithms and particle swarm optimisation.
- *Machine-based approaches*: includes the employment of finite state machines. The finite state machines encode the order in which candidates should be linked.
- *Structures*: involves the use of a B+ tree to simplify the retrieval of cloud services to be used in the composition. This indexing accounts for multiple QoS attributes.
- *Frameworks*: proposes frameworks that detail the steps for the construction of composition solutions. Are typically divided in distinct phases.

Most authors in the field have been focused on the improvement of composition algorithms and the proposal of frameworks/structures for composition, however there is a shortage of proposed datasets for testing, as well as difficulty performing fair comparisons of different proposed technique. This scarcity of works points out to an existing gap in the field that needs to be fulfilled in future works.

Location: A Feature for Service Selection in the Era of Big Data [96]

This paper argues that with the rising popularity of Big Data, the importance of considering the location of Web services when selecting them is also increasing. That is because the transmission cost for large amounts of data has a big impact in the overall efficiency of the composition, as opposed to the typical composition scenario with much lower transmission needs. In this work, services are considered to be located in an area, which is defined as a particular subnetwork, cloud or computer. Thus, the task of finding compositions with the smallest possible distance between areas becomes an optimisation problem. A measure for the distance between services is presented and used in the fitness function employed in the selection, which calculates the distance between all consecutive services in a sequential construct. The problem is tackled using an integer programming (IP) technique devised by the authors for finding the shortest path between areas in the sequential composition. A case study is presented to illustrate this approach, and a performance study on the algorithm is carried out.

A Location & Time Related Web Service Distributed Selection Approach for Composition [51]

This paper proposes a framework for Web service composition approach that considers the time factor (i.e. the hour of the day) and the location of Web services when proposing a solution. The key is to provide decentralised service orchestration, allowing reduced network

traffic and consequently resulting in higher efficiency. In this work, the QoS attributes of services are divided into those that are affected by location and time and those that are not, and services are divided according to the service broker they belong to. The quality of services is tracked throughout time, allowing predictions to be made on their future quality attributes. A seemingly planning-inspired algorithm is proposed to perform the composition, producing a solution that may not be optimal but which satisfies the user requirements. In order to allow for a decentralised execution, the solution is turned into a series of sequences that can be executed independently. At execution time, the services with the best quality predictions are selected to execute the required atomic functionality.

Full Solution Indexing Using Database for QoS-Aware Web Service Composition [45]

This paper presents a QoS-aware Web service composition technique that makes use of a relational database, thus preventing the entire RAM of a machine from being consumed with the creation of a single composition to the detriment of all other processes. Compositions are represented in this approach as graphs, and the various intermediate approaches involved in calculating these graphs are stored in the database as services, paths and semantic concepts. When a user request is issued through a query, algorithms build this information into a single graph that supports parallel and sequence constructs. Experiments were performed using the WSC dataset, and results show that the approach performs well in comparison to another method which does not employ a database. In particular, the proposed approach is capable of finding solutions that require fewer services, at the same time requiring less rounds of checking for the algorithm. However, the paper provides no figures on memory usage, an issue that was one of the main motivations for this research.

Optimizing Web Service Composition for Data-intensive Applications [91]

This paper presents a method to create optimal scaled Web service compositions for data-intensive processing. The core idea is to create replicas of services to augment their throughput, but to balance the creation of these replicas with the incurring cost of doing so. A mathematical model called BROCS (Benefit Ratio of Composite Services) is proposed to calculate the optimal degree of parallelism (i.e. the best cost-benefit ratio for parallelism). The authors tested their approach using microblogging data and a microblog searching Web service, using a range of different replication values for services and identifying the ideal number of replicas to use. When validating against the results obtained from using the range of replication values previously described, it was concluded that the proposed BROCS model successfully identified the optimal amount of service replicas to utilise in that scenario.

An Approach for Mining Service Composition Patterns from Execution Logs [78]

This paper presents an approach for taking advantage of recurring patterns in existing Web service compositions when creating new ones. The rationale for doing so is that composition patterns that are widely used have two fundamental qualities: firstly, they have been shown to be considered accurate/relevant enough by multiple composition producers; secondly, these patterns have been already tested in a range of situations (i.e. the applications they are a part of). A pattern may be thought of as a collection of Web service operations and the control flow structures that connect them. Patterns are valuable because they successfully capture the knowledge of those who originally built the composition, and allow this knowledge to be reused if a similar composition problem presents itself. Many different approaches have been proposed for the extraction of patterns, and they can be divided into two groups: top-down approaches, where the business specifications of different companies are studied in order to create patterns, and bottom-up approaches, where the logs of composite applications are examined with the objective of discerning useful business processes. In this work, a bottom-up method for pattern extraction is presented, consisting of four steps. In the first step, application logs are collected and processed into a common representation. In the second step, services that are frequently associated are identified, since they may indi-

cate a recurring pattern, and these groups are further processed until the final collection of patterns has been created. In the third step, the control flow structures between the services in a pattern are determined. Finally, functionally similar patterns that use different component services are identified, and abstract higher-level patterns are created based on them. This approach is validated with a case study which shows that all stages described can be effectively performed.

Bibliography

- [1] ALFÉREZ, G. H., AND PELECHANO, V. Facing uncertainty in web service compositions. In *Web Services (ICWS), 2013 IEEE 20th International Conference on* (2013), IEEE, pp. 219–226.
- [2] ALFÉREZ, G. H., PELECHANO, V., MAZO, R., SALINESI, C., AND DIAZ, D. Dynamic adaptation of service compositions with variability models. *Journal of Systems and Software* 91 (2014), 24–47.
- [3] AVERSANO, L., DI PENTA, M., AND TANEJA, K. A genetic programming approach to support the design of service compositions. *International Journal of Computer Systems Science & Engineering* 21, 4 (2006), 247–254.
- [4] BADER, J. M. *Hypervolume-based search for multiobjective optimization: theory and methods*. Johannes Bader, 2010.
- [5] BAHADORI, S., KAFI, S., FAR, K., AND KHAYYAMBASHI, M. Optimal web service composition using hybrid ga-tabu search. *Journal of Theoretical and Applied Information Technology* 9, 1 (2009), 10–15.
- [6] BERG, H. V. D., ET AL. Revenue optimization of service compositions using conditional request retries. In *Web Services (ICWS), 2013 IEEE 20th International Conference on* (2013), IEEE, pp. 1–9.
- [7] BERTOLI, P., KAZHAMIKIN, R., PAOLUCCI, M., PISTORE, M., RAIK, H., AND WAGNER, M. Control flow requirements for automated service composition. In *Web Services, 2009. ICWS 2009. IEEE International Conference on* (2009), IEEE, pp. 17–24.
- [8] BLUM, A. L., AND FURST, M. L. Fast planning through planning graph analysis. *Artificial intelligence* 90, 1 (1997), 281–300.
- [9] BOUSTIL, A., MAAMRI, R., AND SAHNOUN, Z. A semantic selection approach for composite web services using OWL-DL and rules. *Service Oriented Computing and Applications* 8, 3 (2014), 221–238.
- [10] BOUSTIL, A., SABOURET, N., AND MAAMRI, R. Web services composition handling user constraints: towards a semantic approach. In *Proceedings of the 12th International Conference on Information Integration and Web-based Applications & Services* (2010), ACM, pp. 913–916.
- [11] BROWN, N., MCKAY, B., GILARDONI, F., AND GASTEIGER, J. A graph-based genetic algorithm and its application to the multiobjective evolution of median molecules. *Journal of chemical information and computer sciences* 44, 3 (2004), 1079–1087.

- [12] CHEN, M., AND YAN, Y. Qos-aware service composition over graphplan through graph reachability. In *Services Computing (SCC), 2014 IEEE International Conference on* (2014), IEEE, pp. 544–551.
- [13] CHEN, Y., HUANG, J., AND LIN, C. Partial selection: An efficient approach for qos-aware web service composition. In *Web Services (ICWS), 2014 IEEE International Conference on* (2014), IEEE, pp. 1–8.
- [14] CHIFU, V. R., POP, C. B., SALOMIE, I., SUIA, D. S., AND NICULICI, A. N. Optimizing the semantic web service composition process using cuckoo search. In *Intelligent distributed computing V*. Springer, 2012, pp. 93–102.
- [15] COTTA, C., AND FERNÁNDEZ, A. J. Memetic algorithms in planning, scheduling, and timetabling. In *Evolutionary Scheduling*. Springer, 2007, pp. 1–30.
- [16] CRAENEN, B., EIBEN, A., AND MARCHIORI, E. How to handle constraints with evolutionary algorithms. *Practical Handbook Of Genetic Algorithms: Applications* (2001), 341–361.
- [17] DA SILVA, A. S., MA, H., AND ZHANG, M. A graph-based particle swarm optimisation approach to qos-aware web service composition and selection. In *Evolutionary Computation (CEC), 2014 IEEE Congress on* (2014), IEEE, pp. 3127–3134.
- [18] DE CAMPOS, A., POZO, A. T., VERGILIO, S. R., AND SAVEGNAGO, T. Many-objective evolutionary algorithms in the composition of web services. In *Neural Networks (SBRN), 2010 Eleventh Brazilian Symposium on* (2010), IEEE, pp. 152–157.
- [19] DENG, S., HUANG, L., TAN, W., AND WU, Z. Top- k automatic service composition: A parallel method for large-scale service sets. *Automation Science and Engineering, IEEE Transactions on* 11, 3 (July 2014), 891–905.
- [20] DENG, S., WU, B., YIN, J., AND WU, Z. Efficient planning for top-k web service composition. *Knowledge and information systems* 36, 3 (2013), 579–605.
- [21] EL HADAD, J., MANOUVRIER, M., AND RUKOZ, M. Tqos: Transactional and qos-aware selection algorithm for automatic web service composition. *Services Computing, IEEE Transactions on* 3, 1 (2010), 73–85.
- [22] FENG, Y., NGAN, L. D., AND KANAGASABAI, R. Dynamic service composition with service-dependent qos attributes. In *Web Services (ICWS), 2013 IEEE 20th International Conference on* (2013), IEEE, pp. 10–17.
- [23] FERRARA, A. Web services: a process algebra approach. In *Proceedings of the 2nd international conference on Service oriented computing* (2004), ACM, pp. 242–251.
- [24] GALVAN-LOPEZ, E. Efficient graph-based genetic programming representation with multiple outputs. *International Journal of Automation and Computing* 5, 1 (2008), 81–89.
- [25] GAMHA, Y., BENNACER, N., NAQUET, G., AYEB, B., AND ROMDHANE, L. B. A framework for the semantic composition of web services handling user constraints. In *Web Services, 2008. ICWS'08. IEEE International Conference on* (2008), IEEE, pp. 228–237.
- [26] GARCÍA, J. M., RUIZ, D., RUIZ-CORTÉS, A., AND PAREJO, J. A. Qos-aware semantic service selection: An optimization problem. In *Services-Part I, 2008. IEEE Congress on* (2008), IEEE, pp. 384–388.

- [27] GLOBUS, A., LAWTON, J., AND WIPKE, T. Automatic molecular design using evolutionary techniques. *Nanotechnology* 10, 3 (1999), 290.
- [28] GLOVER, F. Tabu search-part i. *ORSA Journal on computing* 1, 3 (1989), 190–206.
- [29] GOTTSCHALK, K., GRAHAM, S., KREGER, H., AND SNELL, J. Introduction to web services architecture. *IBM systems Journal* 41, 2 (2002), 170–177.
- [30] GUO, G., YU, F., CHEN, Z., AND XIE, D. A four-level matching model for semantic web service selection based on qos ontology. In *Information Science and Engineering (ISISE), 2010 International Symposium on* (2010), IEEE, pp. 630–634.
- [31] HART, E., ROSS, P., AND CORNE, D. Evolutionary scheduling: A review. *Genetic Programming and Evolvable Machines* 6, 2 (2005), 191–220.
- [32] HUANG, Z., JIANG, W., HU, S., AND LIU, Z. Effective pruning algorithm for qos-aware service composition. In *Commerce and Enterprise Computing, 2009. CEC'09. IEEE Conference on* (2009), IEEE, pp. 519–522.
- [33] HWANG, C.-L., AND YOON, K. Lecture notes in economics and mathematical systems. *Multiple Objective Decision Making, Methods and Applications: A State-of-the-Art Survey* 164 (1981).
- [34] JIANG, W., ZHANG, C., HUANG, Z., CHEN, M., HU, S., AND LIU, Z. Qsynth: A tool for qos-aware automatic service composition. In *Web Services (ICWS), 2010 IEEE International Conference on* (2010), IEEE, pp. 42–49.
- [35] JOSHI, A., ROWE, J. E., AND ZARGES, C. Improving the performance of the germinal center artificial immune system using\ epsilon-dominance: A multi-objective knapsack problem case study. In *Evolutionary Computation in Combinatorial Optimization*. Springer, 2015, pp. 114–125.
- [36] JULA, A., SUNDARARAJAN, E., AND OTHMAN, Z. Cloud computing service composition: A systematic literature review. *Expert Systems with Applications* 41, 8 (2014), 3809–3824.
- [37] KANTSCHIK, W., AND BANZHAF, W. Linear-graph gp-a new gp structure. In *Genetic Programming*. Springer, 2002, pp. 83–92.
- [38] KARAKOC, E., AND SENKUL, P. Composing semantic web services under constraints. *Expert Systems with Applications* 36, 8 (2009), 11021–11029.
- [39] KATTEPUR, A., GEORGANTAS, N., AND ISSARNY, V. Qos composition and analysis in reconfigurable web services choreographies. In *Web Services (ICWS), 2013 IEEE 20th International Conference on* (2013), IEEE, pp. 235–242.
- [40] KAZHAMIKIN, R., MARCONI, A., PISTORE, M., AND RAIK, H. Data-flow requirements for dynamic service composition. In *Web Services (ICWS), 2013 IEEE 20th International Conference on* (2013), IEEE, pp. 243–250.
- [41] KHAKHKHAR, S., KUMAR, V., AND CHAUDHARY, S. Dynamic service composition. *International Journal of Computer Science and Artificial Intelligence* (2012).
- [42] KO, J. M., KIM, C. O., AND KWON, I.-H. Quality-of-service oriented web service composition algorithm and planning architecture. *Journal of Systems and Software* 81, 11 (2008), 2079–2090.

- [43] KOPECKY, J., VITVAR, T., BOURNEZ, C., AND FARRELL, J. Sawsdl: Semantic annotations for wsdl and xml schema. *Internet Computing, IEEE* 11, 6 (2007), 60–67.
- [44] LI, G., LIAO, L., SONG, D., AND ZHENG, Z. A fault-tolerant framework for qos-aware web service composition via case-based reasoning. *International Journal of Web and Grid Services* 10, 1 (2014), 80–99.
- [45] LI, J., YAN, Y., AND LEMIRE, D. Full solution indexing using database for qos-aware web service composition. In *Services Computing (SCC), 2014 IEEE International Conference on* (2014), IEEE, pp. 99–106.
- [46] LI, M., ZHU, D., DENG, T., SUN, H., GUO, H., AND LIU, X. GOS: a global optimal selection strategies for qos-aware web services composition. *Service Oriented Computing and Applications* 7, 3 (2013), 181–197.
- [47] LI, W., BADR, Y., AND BIENNIER, F. Towards a capability model for web service composition. In *Web Services (ICWS), 2013 IEEE 20th International Conference on* (2013), IEEE, pp. 609–610.
- [48] LI, X., YANG, G., AND HIRASAWA, K. Evolving directed graphs with artificial bee colony algorithm. In *Intelligent Systems Design and Applications (ISDA), 2014 14th International Conference on* (2014), IEEE, pp. 89–94.
- [49] LIN, N., KUTER, U., AND SIRIN, E. *Web service composition with user preferences*. Springer, 2008.
- [50] LIU, S., LIU, Y., JING, N., TANG, G., AND TANG, Y. A dynamic web service selection strategy with qos global optimization based on multi-objective genetic algorithm. In *Grid and Cooperative Computing-GCC 2005*. Springer, 2005, pp. 84–89.
- [51] LIU, Z., AND LU, T. A location & time related web service distributed selection approach for composition. In *Grid and Cooperative Computing (GCC), 2010 9th International Conference on* (2010), IEEE, pp. 296–301.
- [52] LU, J., YU, Y., ROY, D., AND SAHA, D. Web service composition: A reality check. In *Web Information Systems Engineering-WISE 2007*. Springer, 2007, pp. 523–532.
- [53] MABU, S., HIRASAWA, K., AND HU, J. A graph-based evolutionary algorithm: genetic network programming (gnp) and its extension using reinforcement learning. *Evolutionary Computation* 15, 3 (2007), 369–398.
- [54] MARCONI, A., PISTORE, M., AND POCCIANI, P. Automated web service composition at work: the amazon/mps case study. In *Web Services, 2007. ICWS 2007. IEEE International Conference on* (2007), IEEE, pp. 767–774.
- [55] MARCONI, A., PISTORE, M., AND TRAVERSO, P. Specifying data-flow requirements for the automated composition of web services. In *Software Engineering and Formal Methods, 2006. SEFM 2006. Fourth IEEE International Conference on* (2006), IEEE, pp. 147–156.
- [56] MARTIN, D., BURSTEIN, M., MCDERMOTT, D., MCILRAITH, S., PAOLUCCI, M., SYCARA, K., MCGUINNESS, D. L., SIRIN, E., AND SRINIVASAN, N. Bringing semantics to web services with owl-s. *World Wide Web* 10, 3 (2007), 243–277.
- [57] MAUÉ, P., AND ROMAN, D. The envision environmental portal and services infrastructure. In *Environmental Software Systems. Frameworks of eEnvironment*. Springer, 2011, pp. 280–294.

- [58] MILANOVIC, N., AND MALEK, M. Current solutions for web service composition. *IEEE Internet Computing* 8, 6 (2004), 51–59.
- [59] MILLER, J. F., AND THOMSON, P. Cartesian genetic programming. In *Genetic Programming*. Springer, 2000, pp. 121–132.
- [60] MOBEDPOUR, D., AND DING, C. User-centered design of a qos-based web service selection system. *Service Oriented Computing and Applications* 7, 2 (2013), 117–127.
- [61] MOGHADDAM, M., AND DAVIS, J. G. Service selection in web service composition: A comparative review of existing approaches. In *Web Services Foundations*. Springer, 2014, pp. 321–346.
- [62] NICOLAOU, C. A., APOSTOLAKIS, J., AND PATTICHIS, C. S. De novo drug design using multiobjective evolutionary graphs. *Journal of Chemical Information and Modeling* 49, 2 (2009), 295–307.
- [63] NOOR, T. H., SHENG, Q. Z., ALFAZI, A., NGU, A. H., AND LAW, J. Csce: A crawler engine for cloud services discovery on the world wide web. In *Web Services (ICWS), 2013 IEEE 20th International Conference on* (2013), IEEE, pp. 443–450.
- [64] PAREJO, J. A., FERNANDEZ, P., AND CORTÉS, A. R. Qos-aware services composition using tabu search and hybrid genetic algorithms. *Actas de los Talleres de las Jornadas de Ingeniería del Software y Bases de Datos* 2, 1 (2008), 55–66.
- [65] PAREJO, J. A., SEGURA, S., FERNANDEZ, P., AND RUIZ-CORTÉS, A. Qos-aware web services composition using grasp with path relinking. *Expert Systems with Applications* 41, 9 (2014), 4211–4223.
- [66] PERREY, R., AND LYCETT, M. Service-oriented architecture. In *Applications and the Internet Workshops, 2003. Proceedings. 2003 Symposium on* (2003), IEEE, pp. 116–119.
- [67] POLI, R. *Parallel distributed genetic programming*. Citeseer, 1996.
- [68] POP, C. B., CHIFU, V. R., SALOMIE, I., AND DINSOREANU, M. Immune-inspired method for selecting the optimal solution in web service composition. In *Resource Discovery*. Springer, 2010, pp. 1–17.
- [69] POP, C. B., ROZINA CHIFU, V., SALOMIE, I., BAICO, R. B., DINSOREANU, M., AND COPIL, G. A hybrid firefly-inspired approach for optimal semantic web service composition. *Scalable Computing: Practice and Experience* 12, 3 (2011).
- [70] POP, C. B., VLAD, M., CHIFU, V. R., SALOMIE, I., AND DINSOREANU, M. A tabu search optimization approach for semantic web service composition. In *Parallel and Distributed Computing (ISPDC), 2011 10th International Symposium on* (2011), IEEE, pp. 274–277.
- [71] RAO, J., AND SU, X. A survey of automated web service composition methods. In *Semantic Web Services and Web Process Composition*. Springer, 2005, pp. 43–54.
- [72] RODRIGUEZ-MIER, P., MUCIENTES, M., LAMA, M., AND COUTO, M. I. Composition of web services through genetic programming. *Evolutionary Intelligence* 3, 3-4 (2010), 171–186.

- [73] SABOOHI, H., AND KAREEM, S. A. World-altering semantic web services discovery and composition techniques-a survey. In *The 7th International Conference on Semantic Web and Web Services (SWWS)* (2011), pp. 91–95.
- [74] SHIN, D.-H., AND LEE, K.-H. An automated composition of information web services based on functional semantics. In *Services, 2007 IEEE Congress on* (2007), IEEE, pp. 300–307.
- [75] SKIENA, S. Dijkstra’s algorithm. *Implementing Discrete Mathematics: Combinatorics and Graph Theory with Mathematica*, Reading, MA: Addison-Wesley (1990), 225–227.
- [76] SOHRABI, S., PROKOSHYNA, N., AND MCILRAITH, S. A. Web service composition via the customization of golog programs with user preferences. In *Conceptual Modeling: Foundations and Applications*. Springer, 2009, pp. 319–334.
- [77] SOYDAN BILGIN, A., AND SINGH, M. P. A daml-based repository for qos-aware semantic web service selection. In *Web Services, 2004. Proceedings. IEEE International Conference on* (2004), IEEE, pp. 368–375.
- [78] UPADHYAYA, B., TANG, R., AND ZOU, Y. An approach for mining service composition patterns from execution logs. *Journal of Software: Evolution and Process* 25, 8 (2013), 841–870.
- [79] WANG, A., MA, H., AND ZHANG, M. Genetic programming with greedy search for web service composition. In *Database and Expert Systems Applications* (2013), Springer, pp. 9–17.
- [80] WANG, D., HUANG, H., AND XIE, C. A novel adaptive web service selection algorithm based on ant colony optimization for dynamic web service composition. In *Algorithms and Architectures for Parallel Processing*. Springer, 2014, pp. 391–399.
- [81] WANG, L., SHEN, J., AND YONG, J. A survey on bio-inspired algorithms for web service composition. In *Computer Supported Cooperative Work in Design (CSCWD), 2012 IEEE 16th International Conference on* (2012), IEEE, pp. 569–574.
- [82] WANG, P., DING, Z., JIANG, C., AND ZHOU, M. Automated web service composition supporting conditional branch structures. *Enterprise Information Systems* 8, 1 (2014), 121–146.
- [83] WANG, X., VITVAR, T., KERRIGAN, M., AND TOMA, I. A qos-aware selection model for semantic web services. In *Service-Oriented Computing–ICSOC 2006*. Springer, 2006, pp. 390–401.
- [84] WANG, X., WANG, Z., AND XU, X. An improved artificial bee colony approach to qos-aware service selection. In *Web Services (ICWS), 2013 IEEE 20th International Conference on* (2013), IEEE, pp. 395–402.
- [85] WEN, S., TANG, C., LI, Q., CHIU, D. K. W., LIU, A., AND HAN, X. Probabilistic top-k dominating services composition with uncertain qos. *Service Oriented Computing and Applications* 8, 1 (2014), 91–103.
- [86] XIA, Y.-M., AND YANG, Y.-B. Web service composition integrating qos optimization and redundancy removal. In *Web Services (ICWS), 2013 IEEE 20th International Conference on* (2013), IEEE, pp. 203–210.

- [87] XIANG, C., ZHAO, W., TIAN, C., NIE, J., AND ZHANG, J. Qos-aware, optimal and automated service composition with users' constraints. In *e-Business Engineering (ICEBE), 2011 IEEE 8th International Conference on* (2011), IEEE, pp. 223–228.
- [88] XIANG, F., HU, Y., YU, Y., AND WU, H. Qos and energy consumption aware service composition and optimal-selection based on pareto group leader algorithm in cloud manufacturing system. *Central European Journal of Operations Research* 22, 4 (2014), 663–685.
- [89] YIN, H., ZHANG, C., ZHANG, B., GUO, Y., AND LIU, T. A hybrid multiobjective discrete particle swarm optimization algorithm for a sla-aware service composition problem. *Mathematical Problems in Engineering* 2014 (2014).
- [90] YOO, J. J.-W., KUMARA, S., LEE, D., AND OH, S.-C. A web service composition framework using integer programming with non-functional objectives and constraints. *algorithms* 1 (2008), 7.
- [91] YU, D., LI, C., AND YIN, Y. Optimizing web service composition for data-intensive applications. *International Journal of Database Theory & Application* 7, 2 (2014).
- [92] YU, Q., AND BOUGUETTAYA, A. Efficient service skyline computation for composite service selection. *Knowledge and Data Engineering, IEEE Transactions on* 25, 4 (2013), 776–789.
- [93] ZHANG, S., DOU, W., AND CHEN, J. Selecting top-k composite web services using preference-aware dominance relationship. In *Web Services (ICWS), 2013 IEEE 20th International Conference on* (2013), IEEE, pp. 75–82.
- [94] ZHANG, W., CHANG, C. K., FENG, T., AND JIANG, H.-Y. Qos-based dynamic web service composition with ant colony optimization. In *Computer Software and Applications Conference (COMPSAC), 2010 IEEE 34th Annual* (2010), IEEE, pp. 493–502.
- [95] ZHANG, Z., ZHENG, S., LI, W., TAN, Y., WU, Z., AND TAN, W. Genetic algorithm for context-aware service composition based on context space model. In *Web Services (ICWS), 2013 IEEE 20th International Conference on* (2013), IEEE, pp. 605–606.
- [96] ZHILING, L., YING, L., AND JIANWEI, Y. Location: a feature for service selection in the era of big data. In *Web Services (ICWS), 2013 IEEE 20th International Conference on* (2013), IEEE, pp. 515–522.