ServLog: A Unifying Logical Framework for Service Modeling and Contracting

Dumitru Roman\textsuperscript{a} and Michael Kifer\textsuperscript{b}
\textsuperscript{a} SINTEF, Forskningsveien 1, 0314 Oslo, Norway
E-mail: dumitru.roman@sintef.no
\textsuperscript{b} Stony Brook University, Stony Brook, NY 11794-2424, U.S.A.
E-mail: kifer@cs.stonybrook.edu

Abstract.
Implementing semantics-aware services, which includes semantic Web services, requires novel techniques for modeling and analysis. The problems include automated support for service discovery, selection, negotiation, and composition. In addition, support for automated service contracting and contract execution is crucial for any large scale service environment where multiple clients and service providers interact. Many problems in this area involve reasoning, and a number of logic-based methods to handle these problems have emerged in the field of Semantic Web Services. In this paper, we lay down theoretical foundations for service modeling, contracting, and reasoning, which we call ServLog, by developing novel techniques for modeling and reasoning about service contracts with the help of Concurrent Transaction Logic. With this framework, we significantly extend the modeling power of the previous work by allowing expressive data constraints and iterative processes in the specification of services. This approach not only captures typical procedural constructs found in established business process languages, but also greatly extends their functionality, enables declarative specification and reasoning about services, and opens a way for automatic generation of executable business processes from service contracts.

Keywords: service modeling and contracting, service process, constraints, Semantic Web service, automated reasoning, declarative modeling of processes.

1. Introduction

The emerging field of service science \cite{33} calls for the development of novel techniques to support various service-related tasks such as service modeling and discovery, service process specification, automated contracting for services, service enactment and monitoring. These issues have been partially addressed by a number of pioneering projects in the area of Semantic Web Services, such as WSMF \cite{15}, OWL-S \cite{23},\textsuperscript{1} WSMO \cite{27},\textsuperscript{2} SWSL,\textsuperscript{3} DIP \cite{36},\textsuperscript{4} and SUPER \cite{21,22},\textsuperscript{5} and more recently \cite{30}. Nevertheless, many core issues remained largely unsolved. The present paper builds on the previous efforts while primarily addressing the behavioral aspects of services, including service contracting and service contract execution. It complements approaches such as OWL-S and WSMO, which

\footnotesize
\begin{itemize}
  \item \textsuperscript{1}http://www.daml.org/services/owl-s/
  \item \textsuperscript{2}http://www.wsmo.org/
  \item \textsuperscript{3}http://www.w3.org/Submission/SWSF-SWSL/
  \item \textsuperscript{4}http://dip.semanticweb.org/
  \item \textsuperscript{5}http://ip-super.org/
\end{itemize}

\textsuperscript{1570-0844/0-1900/$27.50 \textcopyright$ 0 – IOS Press and the authors. All rights reserved}
primary focused on semantic annotations for Web services, brings new insights, and points to new directions for research in Semantic Web Services.

In a service-oriented environment, interaction is expected among large numbers of clients and service providers, so making contracts through human interaction is not feasible. To enable automatic establishment of contracts, a formal contract description language is needed, and a reasoning mechanism must verify that the contract terms are fulfilled. A contract specification has to describe the functionality of a service, values to be exchanged, procedures, and guarantees. The present paper develops just such a unifying logical framework, called $\text{ServLog}$. $\text{ServLog}$ is based on Concurrent Transaction Logic (CTR) [10] and continues the line of research that employs CTR as a unifying formalism for modeling, discovering, choreographing, contracting, and enactment of Web services [13,32,14,28,29]. Transaction Logic has also been successfully used in a number of other domains ranging from security verification policies to reasoning about actions and other service-related issues [6,26,18,7,17]. The present work builds on the results reported in [28,29], but greatly extends that previous work through generalization and addition of new modeling and reasoning techniques. All this is achieved while at the same time significantly simplifying the technical machinery. With $\text{ServLog}$ we lay down the theoretical foundations for service contracting. Specifically, we extend the expressive power of the constraint language used for specifying contracts, allow iterative processes, and allow to pass arguments to processes. We also extend our reasoning techniques to deal with the new expressive variety of modeling primitives, making it possible to address an array of issues in service contracts, ranging from complex process descriptions to temporal and data constraints. The inference procedure for CTR developed here also contributes to the body of results about CTR itself — it covers CTR conjunctive formulas that enable execution of constrained transactions, which previous CTR proof theory was not able to handle. We also develop a logical language for specifying and enacting processes of great complexity.

The remainder of this paper is organized as follows. Section 2 informally describes the basic techniques from service contract specification used in $\text{ServLog}$ and introduces the problem of service contracting and service contract execution. Section 3 gives a short introduction to CTR to keep the paper self-contained. Section 4 formally defines modeling constructs. Section 5 describes the reasoning procedure of $\text{ServLog}$—the key component of service contracting and contract execution in our framework. In Section 6 we discuss related work and compare it to ours. Section 7 concludes the paper.

2. Service Modeling, Contracting, and Contract Execution

Figure 1 depicts the main aspects of service behavior addressed by $\text{ServLog}$. The service process is described through its control and data flows—a specification that tells how to interact with the service and how data flows among tasks. A service process may also expose inner workings of the service (interactions that are not between the service and the client but between internal tasks of the service or 3rd parties) for common situations in which the service allows the client to express constraints on how the service process is executed.

The Business Process Model and Notation (BPMN)\(^6\) is an OMG standard for business process modeling and provides a graphical notation for specifying business processes. BPMN distinguishes between public and private processes, choreographies, and collaborations; it can provide different views of internal and external interactions. BPMN supplies a comprehensive set of graphical notations for business users to model various aspects of services and their interactions. Another approach to modeling service processes is the WSMO model of choreography, \(^7\) which is limited to server-side interactions. In contrast, the model of the W3C Choreography Group includes both service-side interactions and client-side interactions.\(^8\) The W3C Choreography model is symmetric and can be represented in $\text{ServLog}$ via service process, policy, and contract components on both the client and the service side, as in Figure 1. This model can be further extended to multiagent interactions, but this is beyond the scope of this paper and is orthogonal to the problems considered here.

\(^6\)http://www.omg.org/spec/BPMN/2.0/
\(^7\)http://www.wsmo.org/TR/d14/
\(^8\)http://www.w3.org/TR/2006/WD-ws-cdl-10-primer-20060619/
There is a number of modeling languages for capturing interactions between services and clients (or among internal tasks within the same service), some focusing on specific features and targeting different audiences (e.g. business analysts, Web service developers, etc.). At a higher level of abstraction, however, all interactions can be represented through control and data dependencies between tasks. ServLog captures this level of abstraction in a logic and enables powerful forms of automated reasoning about it.

The service policy component in Figure 1 is a set of constraints on a service process and on its input. The contract requirements included on the client side of the figure represent the contractual requirements of the user that go beyond the basic functions of a service. (Examples of basic services are selling books and helping with travel arrangements, while an example of a service requirement is the request that the amount should be charged only after shipping.) In ServLog, service processes are described via control and data flow graphs, while service policy and client contract requirements are described via constraints. We will now discuss these modeling aspects in more detail using a typical order placement scenario depicted in Figure 2.

**Service process.** Figure 2 shows a service process as a hierarchical control and data flow graph, called a process graph, of a fairly complex virtual manufacturing service. The control flow aspect of process graphs is typically used to specify local execution dependencies among the interactions of the service; it is a good way to visualize the overall flow of control. Data flow complements the control flow by specifying the data dependencies among the interactions.

With Figure 2 we are not attempting to suggest yet another notation for service processes; the purpose is to introduce and explain our running example in a compact and focused way. Representing the same information, say, in BPMN would have been much bulkier and would still force us to introduce additional notation to compensate for features (such as data flow) that BPMN lacks.

**Control flow.** The nodes in a service process graph represent interaction tasks, which can be thought of as functions that take inputs and/or produce outputs. Some tasks are meant to be executed by the service and some by the client. The distinction between service and client tasks is part of the service process description. In general there can be several actors involved, some acting as clients in one context and services in another.

In Figure 2, tasks are represented as labeled rectangles. The label of a rectangle is the name of the task represented by that rectangle, and a graph inside the rectangle is the definition of that task. Such a task is called composite because it has nontrivial internal structure. Tasks that do not have associated graphs are primitive. A service process graph can thus be viewed as containing a hierarchy of tasks. The top-level graph shown at the top is the root of the hierarchy. In our example, the tasks of the top-level graph include handle_items, handle_shippers, and...
The top-level graph and each composite task has an initial and a final interaction task, the successor task(s) for each interaction task, and a sign that tells whether all these successors must be executed concurrently (represented by AND-split nodes), or whether only one of the alternative branches needs to be executed non-deterministically (represented by OR-nodes). For instance, in the top-level graph, all successors of the initial interaction place_order must be executed whereas in the definition of accept_producer either accept or reject is to be executed. The node contact_producer (inside manage_item) is also an OR-split, but with a twist. The upper branch going out of this node represents the situation where the client is informed about a producer. The lower branch, has no interactions and it joins with the upper branch. Since this is an OR-node, it means that the upper branch (inform_client) is optional: the service may or may not inform the client about a producer.

Some composite tasks in Figure 2 are marked with the suffix "*" — these can execute multiple times. We call these tasks iterative and differentiate them from non-iterative tasks. Iteration is indicated through recursive occurrences of the same tasks — by placing tasks inside their own definitions. For example, handle_items is an iterative task where
a sequence of sub-tasks, select_item followed by manage_item can be executed multiple times (for example for each item in the purchase order). Iteration is indicated by a recursive occurrence of a handle_items box to the right of the manage_item box. Note that recursive occurrences of tasks may be followed by other tasks, which gives us a general mechanism for capturing different kinds of iterations, including loops and nesting.

It should now be clear how the control flow aspect of the service process graph in Figure 2 represents the virtual manufacturer scenario described earlier: first the order is placed (place_order), then the items in the purchase order are processed (handle_items), delivery is arranged (handle_shippers), and payment is settled (handle_payment). These three tasks are executed in parallel. Once all of them complete, handling of the order is finished (end_order). The other parts of the figure show how each of the above tasks is executed. The important thing to observe here is that some tasks are complex and some primitive; some are to be executed in parallel (the AND-nodes) and some in sequence; some tasks have non-deterministic choice (the OR-nodes) and some are iterative.

**Data flow.** Interaction with a service typically involves passing data and the flow of that data is normally captured using data dependencies between tasks. Such dependencies complement the control flow and complete the description of the service process graph.

Since tasks can be conceptualized as functions that take input and produce output, arguments are attached to tasks to capture both input and output. In Figure 2, each task-label has one or more arguments. For example, handle_payment has the arguments (Order#,Price), meaning that, to execute handle_payment, Order# and Price must be provided.

Data-passing between tasks is captured via the shared argument names and through a shared data space (e.g., a database) of the workflow process. Data-passing through shared arguments is possible between a task and its direct successors, or within the definition of the same composite task. The scope of arguments is relevant in this case: argument names of a task are logical variables. When they are shared with the task’s direct successors, they refer to the same data items (i.e., data is passed from tasks to their direct successors using shared arguments). This aspect should be familiar from basic logic and Logic Programming. For example, data identified by Order# in handle_items(Order#) is the same as the data identified by Order# in place_order(Order#,Price), i.e., place_order passes Order# to handle_items. In case of a composite task, the names of its arguments are global to that task’s definition, meaning that if subtasks in its definition use the same argument names as the composite task then they refer to the same data items. In this way, the composite task passes data to its subtasks. For example, the composite task handle_items passes Order# to its manage_item subtask.

Note also that sharing via shared variables is bi-directional, as pure logic has no notion of explicit input and output. However, for practical reasons, some logic programming systems accept mode specification and do mode-inference, which allows the user to identify the producers of data.

Data-passing through shared data space is used when passing data is not possible through shared arguments due to the difference in scope of the arguments. This often occurs when data needs to be shared between tasks that have no control dependencies in the control flow part of the service process graph. For example, if select_item(Order#,Item) needs to pass Order# and Item to book_shipper(Ord#,Shipper,Itm), data-passing through shared argument names is not an option, since these subtasks appear inside composite tasks that are not related via control dependencies. To capture such data-passing, a shared database can be used as follows: select_item can store Order# and Item, and book_shipper can read them later. This is depicted by the dashed arc going from select_item to book_shipper. The label on the arc represents the data items that are being passed. In our case, Order# is being passed as Ord# and Item as Itm.

Data items can be consumable or non-consumable. In case of data-passing through shared argument names, data is non-consumable: data-producing tasks share data via all of their out-ports with the receiving tasks. The latter can further share this data with their descendants, and so on. In our example place_order produces Order#-items, and that item is then shared with with handle_items, handle_shippers, and handle_payment. Note the emphasis on sharing: all tasks involved work on the same copy of the data—the phenomenon that is familiar from basic predicate logic and logic programming. In case of data-passing through a shared database, data can be consumable or non-consumable. It is consumable when each query to the databases is followed by deletion of the queried data item. It is non-consumable data if such deletion is not implied. Note that in our example all data items are consumable.
A final remark about the service process graphs concerns transition conditions on arcs. Examples of such conditions are \textit{producer(Item,Producer)} in the definition of \textit{manage item} and \textit{itemsAvailable(Order#,Producer,Item)} or \textit{itemsUnavailable(Order#,Producer,Item)} in the definition of \textit{inform client}. Another example of transition condition is a test of the form \textit{Quantity} \textgreater 3. A transition condition signifies that in order for the next interaction with the service to take place, the condition must be satisfied. Transition conditions are Boolean tests attached to the arcs in the service process graphs. These tests can also query the underlying database. Only the arcs whose transition conditions evaluate to true can be followed at run time. For uniformity, \textit{ServLog} treats transition conditions formally as a separate type of task.

\textbf{Service Policies and Client Contract Requirements.} Apart from the local dependencies represented directly in control flow graphs, \textit{global constraints} often arise as part of policy specification. Another case where global constraints arise is when a client has specific requirements to the interaction with the service. These requirements usually have little to do with the functionality of a service (e.g., handling orders); instead they represent guarantees that the client wants before entering into a contract with the service. We call such constraints \textit{client contract requirements}. In Figure 3 we give an example of global constraints that represent service policies and client contract requirements for our running example. To make the example more interesting, we use iterative tasks.

\begin{itemize}
    \item \textbf{Service policy}
    \begin{enumerate}
        \item A shipper is booked only if the user accepts at least 7 items.
        \item If pay-per-item is chosen by the user, then the payment must happen immediately before each item delivery.
        \item Payment guarantee must be given before the client is informed about the availability of items.
    \end{enumerate}
    \item \textbf{Client contract requirements}
    \begin{enumerate}
        \item All items in the same order must be shipped at the same time.
        \item If full payment is chosen by the client, then it must happen only after all purchased items are delivered.
        \item Before the client purchases items, the service must book a shipper.
    \end{enumerate}
\end{itemize}

Fig. 3. Global behavioral constraints on iterative processes.

Constraints can be imposed on separate tasks (e.g., a task must or must not execute, \textit{may} execute a certain number of times) or it can involve several tasks (a task must execute in a certain relationship to another task, e.g., before, after, between). Furthermore, constraints can be combined using Boolean connectives (e.g., a task must execute but after its execution some other task must not execute or must execute some number of times).

Other constraints may involve data only. Examples of such constraints include service pre- and post-conditions. For instance, the requirement that “a confirmation number must be available after the execution of the \textit{book_shipper} service” is a post-condition for the \textit{book_shipper} service, where the confirmation number is a data item in the constraint. Since data in such constraints arise as a result of interactions, this kind of constraint can be seen as a special case of constraints on interactions. Other types of constraints involve Quality of Services (QoS) and Service Level Agreements (SLAs). For instance, “availability provided by the \textit{book_shipper} service is always greater than the requested availability” is a QoS requirement. \textit{ServLog} can also model QoS constraints but this is outside the scope of this paper.

\textbf{Service Contracting and Service Contract Execution.} With a modeling mechanism in place, we define service contracting and service contract execution in \textit{ServLog} as follows:

\begin{itemize}
    \item \textit{Service contracting}: Given a service process (i.e., control and data flow) and a set of service policies and client contract requirements (i.e., constraints), decide whether an execution of the service process that satisfies both the service policies and the client contract requirements exists. Note that it does not matter in the end if this is actually executed but the important aspect is that there is at least one and execution of the contract can proceed.
Service contract execution: Execute tasks in the process in a way where client and service take turns as prescribed by the control and data flows and the constraints. When a step is proposed, the logic’s proof theory verifies if acceptance of that step still leaves the possibility of a successful execution of the entire service process that satisfies all the constraints. If so, the step is accepted and executed; it is rejected otherwise. A list of possible allowed steps can also be suggested by the system at each turn.

To solve the above two problems, Section 4 formally defines service processes, service policies, and client contract requirements using Concurrent Transaction Logic (CTR). Section 5 then extends the original proof theory of CTR to be suitable for the task at hand.

3. Overview of CTR

Concurrent Transaction Logic (CTR) [10] is an extension of the classical predicate logic, which allows programming and reasoning about state-changing processes. Here we summarize the relevant parts of CTR’s syntax and give an informal account of its semantics. For details we refer the reader to [10].

Basic syntax. The atomic formulas of CTR are identical to those of the classical logic, i.e., they are expressions of the form \( p(t_1, \ldots , t_n) \), where \( p \) is a predicate symbol and the \( t_i \)'s are terms constructed of constants, variables, and function symbols. Complex formulas are built with the help of connectives and quantifiers. Apart from the classical \( \lor, \land, \neg, \forall, \exists \), CTR has two additional infix connectives, \( \odot \) (serial conjunction) and \( \otimes \) (concurrent conjunction), and a modal operator \( \circ \) (isolated execution). For instance, \( \circ(p(X) \otimes q(X)) \mid (\forall Y(r(Y) \lor s(X, Y))) \) is a well-formed formula in CTR.

Informal semantics. Underlying the logic and its semantics is a set of database states and a collection of paths. For this paper, the reader can think of states as just relational databases, but the logic is more general and can deal with a wide variety of states. Formally, in this paper, a state is a pair consisting of a state identifier and a relational database.

A path is a finite sequence of state identifiers (constants used to refer to the actual states). For instance, if \( s_1, s_2, \ldots , s_n \) are state identifiers, then \( \langle s_1 \rangle, \langle s_1 s_2 \rangle \), and \( \langle s_1 s_2 \ldots s_n \rangle \) are paths of length 1, 2, and \( n \), respectively.

As in classical logic, CTR formulas take truth values. However, unlike classical logic, the truth of CTR formulas is determined over paths, not at states. If a formula, \( \phi \), is true over a path \( \langle s_1 s_2 \ldots s_n \rangle \), it means that \( \phi \) can execute starting at state \( s_1 \). During the execution, the current state will change to \( s_2, s_3, \ldots \), etc., and the execution terminates at state \( s_n \). In such a case we will also call the path \( \langle s_1 \ldots s_n \rangle \) an execution of \( \phi \).

Although we are interested in execution of CTR formulas over paths, if a formula involves the concurrency operator, the subformulas may be executed in an interleaved fashion, like database transactions. For instance, if \( \phi = (p \otimes q \otimes r) \mid (u \otimes v) \), the concurrency operator means that legal executions of \( \phi \) consist of an execution of some part of \( p \otimes q \otimes r \), e.g., \( p \), then of some execution of \( u \otimes v \), e.g., \( u \) then again of some part of \( p \otimes q \otimes r \), such as \( q \) or even \( r \), then the remaining part of \( u \otimes v \), i.e., \( v \), etc. The concurrency operator does not preclude the two parts of \( \phi \) from executing one after another (in any order), but this type of non-interleaved execution is less interesting. In the first, interleaved execution, the two parts of \( \phi \) execute not on paths but on multi-paths, i.e., on sequences of paths. Execution of one part of \( \phi \) may be broken by executions of another part, so the intervening gaps in the execution of \( p \otimes q \otimes r \) are filled by executions of \( u \otimes v \) (and vice versa).

A multi-path (or an m-path) is a sequence \( (\pi_1, \ldots , \pi_k) \) of paths. If \( \mu = (\pi_1, \ldots , \pi_k) \) and \( \mu' = (\pi'_1, \ldots , \pi'_n) \) are two m-paths, their concatenation, \( \mu \cdot \mu' \), is the m-path \( (\pi_1, \ldots , \pi_k, \pi'_1, \ldots , \pi'_n) \) and their interleaving, \( \mu \upharpoonright \mu' \), is an m-path of the form \( (\kappa_1, \ldots , \kappa_{k+n}) \) such that it is a topological sort of the two sequences \( \mu \) and \( \mu' \). For example, one interleaving of \( (s_1, s_2, s_3), (s_0, s_7) \) and of \( (s_4, s_5), (s_8, s_9) \) is \( (s_1, s_2, s_3), (s_4, s_5), (s_0, s_7), (s_8, s_9) \). Another interleaving is \( (s_1, s_2, s_3), (s_4, s_5), (s_8, s_9), (s_6, s_7) \). Yet another is \( (s_1, s_2, s_3), (s_6, s_7), (s_4, s_5), (s_8, s_9) \), a degenerate interleaving.

Finally, a path \( \pi = (s_1, \ldots , s_m) \) is a merge of an m-path \( (\pi_1, \ldots , \pi_n) \) if there are integers \( 1 = i_0 \leq i_1 \leq i_2 \leq \ldots \leq i_{n-1} \leq i_n = m \) such that \( \pi_1 = \langle s_{i_0}, \ldots , s_{i_1} \rangle, \pi_2 = \langle s_{i_1}, \ldots , s_{i_2} \rangle, \ldots , \pi_{n-1} = \langle s_{i_{n-2}}, \ldots , s_{i_{n-1}} \rangle, \pi_n = \langle s_{i_{n-1}}, \ldots , s_{i_n} \rangle \). Note that for the merge to be possible, the end-state of each path \( \pi_i \) in the m-path must be the start-state
of the subsequent path \( \pi_{l+1} \) for each \( 1 \leq l < n \). For instance, \((s_1, s_2, s_3, s_4, s_5, s_6)\) is a merge of the m-path
\((\langle s_1, s_2 \rangle, \langle s_2, s_3, s_4 \rangle, \langle s_4 \rangle, \langle s_4, s_5, s_6 \rangle)\).

A multipath structure is a mapping that, for each multipath \( \pi \), tells which ground atomic formulas are true on \( \pi \). Informally, this can be understood as telling which ground atomic transactions can execute along \( \pi \). Note that CTR formulas hold truth values not over states, but over multipaths.

First, we connect truth over path of length 1 to database states.

- If \( s \) is a state identifier and \( p \) is a fact that is true in the database associated with \( s \) then \( p \) is true over the path \((s)\) (and the m-path \((\langle s \rangle)\)).

CTR connectives are used to construct composite formulas out of the atomic ones, and the statements below define which composite formulas are true on which multipaths.

- \( \phi \otimes \psi \): execute \( \phi \) then execute \( \psi \). Model-theoretically: \( \phi \otimes \psi \) is true over an m-path \( \pi \) in a multipath structure if \( \phi \) is true over a prefix m-path of \( \pi \), \( \pi_1 \) (in that same structure), and \( \psi \) is true over the suffix m-path, \( \pi_2 \). That is, if \( \pi = \pi_1 \cdot \pi_2 \).

- \( \phi \mid \psi \): \( \phi \) and \( \psi \) execute concurrently, in an interleaved fashion. That is, \( \phi \mid \psi \) is true over an m-path \( \pi \) in a multipath structure if \( \phi \) is true over an m-path \( \pi_1 \) (in that same structure), \( \psi \) is true over an m-path \( \pi_2 \), and \( \pi \) is one of the interleavings \( \pi_1 \parallel \pi_2 \).

- \( \phi \vee \psi \): \( \phi \) and \( \psi \) execute along the same path. That is, \( \phi \vee \psi \) is true on an m-path \( \pi \) if both \( \phi \) and \( \psi \) are true on \( \pi \). In practice, this is best understood in terms of constraints on execution. For instance, \( \phi \) can be thought of as a non-deterministic transaction and \( \psi \) as a constraint on the execution of \( \phi \). It is this feature of the logic that lets us specify constraints as part of service contracts.

- \( \phi \vee \psi \): execute \( \phi \) or execute \( \psi \) non-deterministically. That is, \( \phi \vee \psi \) is true on an m-path \( \pi \) if either \( \phi \) or \( \psi \) is true on \( \pi \).

- \( \neg \phi \): execute in any way provided that this will not be a valid execution of \( \phi \). That is, \( \neg \phi \) is true on any m-path on which \( \phi \) is not true. Negation is an important ingredient in temporal constraint specifications.

- \( \varnothing \phi \): execute \( \phi \) in isolation, i.e., without interleaving with other concurrently running tasks. That is, \( \varnothing \phi \) is true on any singleton m-path (an m-path that contains just one path) where \( \phi \) is true. Note: \( \varnothing \phi \) is never true on an m-path that consists of more than one path, so the execution of \( \varnothing \phi \) cannot be broken by other executions. This operator enables us to specify non-interleaved parts of service contacts.

When considering the entire service, we are interested in its executions over paths, not m-paths: executions over m-paths are used only to represent concurrently running subtasks of the service. To complete the picture, we define truth of CTR formulas over paths:

- \( \phi \) is true over a path, \( \pi \), if it is true over an m-path, \( \pi' \), and \( \pi \) is a merge of \( \pi' \).

CTR contains a special propositional constant, \( \text{state} \), which is true only on paths of length 1, that is, on database states. In service processes, \( \text{state} \) is often used as the exit condition for iterative tasks. Another propositional constant that we will use to represent constraints is \( \text{path} \), defined as \( \text{state} \lor \neg \text{state} \); this constant is true on every path.

**Concurrent-Horn subset of CTR.** The implication \( p \leftarrow q \) is defined as \( p \lor \neg q \). The form and the purpose of the implication in CTR is similar to that of Datalog: \( p \leftarrow q \) can be thought of as the name of a procedure and \( q \) as the definition of that procedure. However, unlike Datalog, both \( p \) and \( q \) take truth values over execution paths, not at individual states.

More precisely, \( p \leftarrow q \) means: if \( q \) can execute along a path \((s_1, ..., s_n)\), then so can \( p \). If \( p \) is viewed as a task name, then the meaning can be re-phrased as: one way to execute task \( p \) is to execute its definition, \( q \).

To specify service processes we use concurrent-Horn goals and concurrent-Horn rules.

**Definition 3.1 (Concurrent-Horn goal)** A concurrent-Horn goal is either an atomic formula or has the form \( \phi \otimes \psi \), \( \phi \mid \psi \), \( \phi \vee \psi \), or \( \varnothing \phi \), where \( \phi \) and \( \psi \) are concurrent-Horn goals.

When confusion does not arise, we will often talk about CTR goals, omitting the “concurrent-Horn” adjective. □
Concurrent-Horn goals occur in our setting in two places: as bodies of the rules that are used to define composite tasks and as formulas that are formal embodiments of control flow graphs. In the latter case, we will be interested in finding out whether a control flow graph can be enacted. Such a question corresponds to proving a statement of the form $\exists X \phi$, where $\phi$ is a CTR goal and $X$ is the set of variables that occur in $\phi$.

**Definition 3.2 (Concurrent-Horn rule)** A concurrent-Horn rule is a CTR formula of the form

$$\forall X (\text{head} \leftarrow \text{body})$$

where $\text{head}$ is an atomic formula, $\text{body}$ is a concurrent-Horn goal, and $X$ is the set of variables that occurs in $\text{head}$ and $\text{body}$. $\square$

Since all variables in a rule are quantified the same way (universally outside of the rule), we will usually omit explicit quantifiers—a common practice that simplifies the notation.

The concurrent-Horn fragment of CTR has an SLD-style proof procedure that proves concurrent-Horn formulas and executes them at the same time [10]. The present paper significantly extends this proof theory to formulas that contain the $\land$ connective thus enabling execution of constrained transactions, which are non-Horn. We also deal with a much larger class of constraints than [13,28], including iterative processes.

**Elementary updates.** In CTR, elementary updates are ground (i.e., variable-free) atomic formulas that change the underlying database state. Semantically they are represented by binary relations over state identifiers. For instance, if $⟨s_1 \ s_2⟩$ belongs to the relation corresponding to an elementary update $u$, it means that $u$ can cause a transition from state $s_1$ to state $s_2$. We will conveniently represent this kind of situation using the following notation:

$$s_1 \xrightarrow{u} s_2$$

Usually the binary relations that represent elementary updates are defined outside of CTR. In that case, they are called transition oracles [8,10,9]. Transition oracles can be defined using formal English or a number of other formal languages. They can also be represented in CTR as partially defined actions [26]. In either case, the elementary updates can be defined to perform any kind of transformation. For instance, they can add or delete single tuples or sets of tuples, add and delete entire relations, and so on.

**Constraints.** Because formulas are defined on paths, CTR can express a wide variety of constraints on the way formulas may execute. One can place existential constraints on execution (these are based on serial conjunction), or universal constraints, which are based on serial implication. To express the former, we use the propositional constant path introduced above. For example, $\text{path} \otimes \psi \otimes \text{path}$ is a constraint that is true on a path if $\psi$ is true somewhere on that path. To express universal constraints, the binary connectives “$\leftarrow$” and “$\Rightarrow$” are used, which are defined via $\otimes$ and $\neg$ as follows: $\psi \leftarrow \phi \overset{\text{def}}{=} \neg(\psi \otimes \neg \phi)$ and $\psi \Rightarrow \phi \overset{\text{def}}{=} \neg(\psi \otimes \neg \phi)$. A moment’s reflection should convince the reader that $\psi \leftarrow \phi$ means that whenever $\phi$ occurs then $\psi$ must have occurred just before it and that $\psi \Rightarrow \phi$ means that whenever $\psi$ occurs then $\phi$ must occur right after it. Thus, $\text{path} \Rightarrow \psi \leftarrow \text{path}$ constrains executions to be such that every subpath encountered in the course of the execution satisfies $\psi$ (including subpaths of the form $⟨s⟩$, where $s$ is an arbitrary intermediate state).

### 4. Formalizing Service Contracts in ServLog

This section formally defines the core modeling elements of ServLog. First we define service processes directly in CTR. Section 4.2 then introduces service policies and contract requirements as constraints that can be expressed in CTR. Section 4.3 then defines an important notion, which we call the service contract assumption.
4.1. Modeling Service Processes

Definition 4.1 (Task) A task is represented by a predicate. The name of the predicate is the name of that task and the arity specifies the number of arguments that the predicate takes. For notational simplicity, we assume that each predicate name has exactly one arity, so each task is uniquely defined by its name.

In service contract specifications, the actual invocations of tasks are represented by task atoms.

Definition 4.2 (Task atom) A task atom is a statement of the form

\[ p(t_1, \ldots, t_n) \]  

where \( p \) is a task predicate of arity \( n \)(\( n \geq 0 \)) and \( t_1, \ldots, t_n \) are terms (defined as in first-order logic) that represent the arguments that \( p \) takes. The terms representing the arguments of the task are placeholders for data items that the task manipulates (its inputs and outputs).

For brevity, we will often write task atoms as \( p(T) \), \( p(U) \), etc., where \( T \), \( U \), ... stand for the tuples of arguments that the task takes.

When confusion does not arise, the term “task” will refer both to tasks and task atoms.

We distinguish between two main types of task predicates: composite and primitive. Composite task predicates are the ones defined by rules (i.e., they are allowed in the rule heads) and primitive tasks are not allowed in the rule heads. The primitive task predicates are further subdivided into update-tasks, query-tasks, and builtin test tasks. The update task predicates are those used as the elementary updates of CTR; the query tasks are the ones whose predicates are used to represent the facts stored in database states; and the builtin tests use predicates whose truth is independent of the database state. These three categories of predicates are disjoint. The transition conditions on the arcs of service process graphs, which were introduced in Section 2, are also treated in ServLog as tasks: specifically as query tasks or builtin tests.

In this paper, we will be using the builtins “=” (identity), “!=” (distinct values), “>”, “<”, and others, as needed. The identity predicate \( a = b \) is true if and only if \( a \) and \( b \) are the same ground (i.e., variable-free) term and \( a \neq b \) holds if \( a \) and \( b \) are distinct ground terms. Clearly, the truth of these predicates is independent of the database state (or of a path) where the predicate is evaluated.

From now on, when talking about CTR goals and rules, we assume that the atomic formulas there are task atoms only. In addition, the predicates occurring in the rule heads must correspond to composite tasks only.

Definition 4.3 (Task occurrence) A task \( p \) occurs in a CTR goal \( \Omega \) if \( \Omega \) contains a task atom \( p(T) \) for some \( T \).

Definition 4.4 (Immediate subtask) Let \( p \) and \( q \) be a pair of tasks and \( R \) be a set of rules. We say that \( p \) is an immediate subtask of \( q \) with respect to \( R \) if and only if \( R \) contains a rule of the form \( q(T) \leftarrow \Omega \) and \( p \) occurs in \( \Omega \).

Definition 4.5 (Subtasks) Let \( p \) and \( q \) be a pair of tasks and \( R \) be a set of rules. Then \( p \) is a subtask of \( q \) with respect to \( R \) if and only if \( p \) is either an immediate subtask of \( q \) or there is an immediate subtask \( r \) of \( q \) such that \( p \) is a subtask of \( r \).

Definition 4.6 (Non-iterative rule) A rule in \( R \) is non-iterative if and only if it has the form

\[ q(T) \leftarrow \Omega \]

where \( q \) does not occur in \( \Omega \) and none of the tasks that occur in \( \Omega \) have \( q \) as a subtask.

\[ \]  

[10] Recall that these predicates can be represented either outside of CTR, as transition oracles [8,10,9], or inside CTR, as partially defined actions [26].
Here is an example of a non-iterative rule ($p$ and $r$ are assumed to be primitive tasks here): $q(X) \leftarrow p(X, Y) \otimes r(Y)$. As seen in this example, variables in ServLog are represented as symbols prefixed with “?”. 

**Definition 4.7 (Iterative rule)** A rule $q(T) \leftarrow \Omega$ in R is iterative if and only if $q$ either occurs in $\Omega$ directly or it is a subtask of a task that occurs in $\Omega$. 

Here are examples of iterative rules:

- $q(X) \leftarrow (p(X, Y) \otimes q(Y) \otimes r(Y, Z)) \lor s(Y)$
- $q(X) \leftarrow p(X, Y) \otimes q(Y) \otimes qq(Y, Z)$
- $qq(X) \leftarrow (pp(X, Y) \otimes q(Y)) \lor t(X)$

Note that here $q$ and $qq$ are both iterative tasks that are mutually dependent on each other (are subtasks of each other). In practice, however, the most common form of iterative tasks is a loop of the form

$$q(T) \leftarrow \Phi \otimes q(U)$$
$$q(T) \leftarrow \Psi$$

where $\Phi$ and $\Psi$ do not depend on $q$.

**Definition 4.8 (Service process)** A service process is a pair ($\Omega$, R), where $\Omega$ is a CTR goal and R is a set of iterative and non-iterative rules whose heads are task atoms of the tasks that occur in $\Omega$ or are subtasks of these tasks. 

Recall that primitive tasks come in three guises: updates, queries, and builtins. Similarly, we classify composite tasks based on the rules that define them. Namely, if a task is defined by at least one iterative rule (i.e. there exists an iterative rule with the task as its head), we call the task iterative; if it is defined only by non-iterative rules then the task is non-iterative.

Equipped with this mechanism for defining service processes in ServLog, one can capture a wide range of control and data flow constructs that often appear in business process languages and notations. For example, the service process introduced in Figure 2 is represented in ServLog as shown in Figure 4.

The top-level graph is specified as a CTR goal at the very beginning. The tasks appearing in that goal are defined by the rules that follow. This service process illustrates data flow through variables as well as via a shared database. For example, passing data from handle_items to manage_shipper is done via the underlying database by having handle_items insert selected_item(?Order#, ?Item) and then querying this data item by the task manage_shipper. To fully capture the dataflow, we introduce three additional database query predicates: producer(?Item, ?Producer), which returns a producer for the given item, itemAvailable(?Order#, ?Producer, ?Item), which is true if the given item is produced by the given producer and the item is in stock, and itemUnavailable(?Order#, ?Producer, ?Item), which is the negation of itemAvailable. While the set of available producers is relatively static, the relation itemAvailable can be modified by the task contact_producer(?Producer, ?Item). For instance, after contacting the producer an item might become reserved.

Data flow types supported by ServLog are simple yet powerful: tasks can share data through shared variables or through the underlying shared database — the former is standard in classical logic and in logic programming languages, the latter is a feature of CTR. For instance, in the looping subprocess manage_item in Figure 4, the same data is passed through the shared variables ?Order# and ?Item to the query producer, other tasks (e.g., contact_producer), and also to the recursive (looping) invocations of manage_item itself. This data comes from the task handle_items and then is passed along to the top-level invocation of manage_item. Inside manage_item, new data is obtained by the query producer and then is passed to the subtasks contact_producer and inform_client through the shared variable ?Producer.
Definition 4.9 (Task Pattern) A task pattern has the form $p(t_1, ..., t_n)$ where each $t_i$ is either a regular ground term (of the kind that may occur in a task atom) or a placeholder. A placeholder is either a named logical variable (which

Goal:
place_order(Order#, Price) @
(handle_items(Order#))
| handle_shippers(Order#)
| handle_payment(Order#, Price) @
end_order(Order#)

Rules:
handle_items(Order#) ←
| select_item(Order#, Item) @
| insert_item(Order#, Item) @  // elementary update to insert selected_item(...) 
| manage_item(Order#, Item) @
| handle_items(Order#) @

handle_shippers(Order#) ←
| contact_shipper(Order#, Shipper) @
| manage_shipper(Shipper, Order#) @
| deliver(Order#) @

handle_payment(Order#, Price) ← pay(Order#)
handle_payment(Order#, Price) ← payment_guarantee(Order#, Credit#
inform_client(Order#, Producer, Item) ←
| itemAvailable(Order#, Producer, Item) @
| inform_availability(Order#, Producer, Item) @
| acceptProducer(Order#, Producer) @

inform_client(Order#, Producer, Item) ←
| itemUnavailable(Order#, Producer, Item) @
| acceptProducer(Order#, Producer) @
| accept(Order#, Producer) @
| reject(Order#, Producer) @
| pay(Order#) ← full_payment(Order#) @
| pay_per_item(Order#) @
| pay_one_item(Order#) @ payment_per_item(Order#) @ state

Fig. 4. ServLog representation of a service process.

4.2. Modeling Constraints

We now formalize service policies and contract requirements as constraints in ServLog. Note that constraints are not defined directly as CTR formulas (unlike task definitions). The main reason for this is that constraints represent patterns that executions of service processes must follow and specialized language constructs for such patterns make specification of constraints easier to understand. Nevertheless, the constraints of ServLog can be expressed as CTR formulas (see Appendix C), so CTR is indeed used here as a unifying formalism for both service task definition and constraints. Recall that whereas CTR can represent constraints, they are not Concurrent Horn formulas and are therefore not handled by the existing CTR proof theory — the extension of the CTR proof theory to handle constraints is proposed in Section 5.2 and is an important contribution of this paper.

A constraint specifies the rules governing the occurrences of various tasks during the execution of a service. Each occurrence of a task is represented by a pattern, which specifies the task name and various conditions on the arguments with which that task can be invoked during the execution. These conditions can require that certain arguments must be bound to specific values and they can also require that certain arguments must be shared within a task occurrence or across the occurrences of different tasks.
will be designated with the prefix ‘_’, e.g., _ord#) or a don’t care placeholder, denoted by ‘_’. Each occurrence of a don’t care placeholder represents a new logical variable that does not occur in other patterns.

We will often need to perform two operations: matching and refinement. The former is the usual matching operator of first-order logic: it is a substitution, \( \theta \) such that \( \theta(pattern) = task\_atom \). Since task\_atom is ground, \( \theta \) will normally be a ground substitution. In this case we will say that the pattern and the ground task match. Note that different occurrences of ‘_’ may be mapped by \( \theta \) to different constants, since such occurrences represent completely different logical variables. Refinement is defined next.

**Definition 4.10 (Refinement)** The refinement operation takes a ground task atom and a pair of task patterns and yields another task pattern as follows:

\[
\text{refine(out\_pattern}; \text{in\_ground\_task}, \text{in\_pattern}) = \text{refined\_pattern}
\]

Here the arguments in\_ground\_task and in\_pattern must have the same task name and in\_ground\_task must match in\_pattern. The task-pattern out\_pattern may have a different task name. The result of the operation, refined\_pattern, is defined as follows: Let \( \theta \) be the substitution that matches in\_pattern against in\_ground\_task. If out\_pattern has variables other than those in in\_pattern, \( \theta \) can map them to anything (to some other variable or constant). Then

\[
\text{refined\_pattern} = \theta(out\_pattern)
\]

One can verify by direct inspection that the task atom \( p(2, 1, abc, cde, 13, cde, 13, 5) \) matches the pattern \( p(_1, \_abc, \_foo, \_bar, \_foo, \_bar, _) \), that refine \( p(_ff, 5, _) ; p(1, 2, 3), p(\_, ff, 3) \) = \( p(2, 5, _) \), and that refine \( q(_f, 5, _, _, h) ; p(1, 2, 2, 3), p(\_, f, \_, f, 3) \) = \( q(2, 5, 2, _, h) \). The last example also illustrates the situation where in\_pattern has named placeholders that do not occur in out\_pattern; the number of arguments in the input and output patterns can also differ.

**Definition 4.11 (Constraints)** In this definition, we will use \( \mathcal{T}, \mathcal{\Pi}, \) etc., to represent tuples that include placeholders as some of the arguments in task patterns. The uppercase symbols \( \mathcal{T}, \mathcal{U}, \) etc., will denote tuples of arguments in task atoms (i.e., they do not contain placeholders). The task names \( p, q, r \), and the task patterns mentioned in the constraints, below, do not need to be distinct.

The set \( \mathcal{CONSTR} \) of constraints supported by ServLog is formally defined as follows. For each constraint we first give its syntax followed by a brief informal explanation and then provide a formal semantic definition. Appendix C provides alternative representation of these constraints as CTR formulas.

1. **Existence constraints:**
   
   - **atleast\(_n(p(\mathcal{T}))\):** task \( p \) must execute at least \( n \) times \((n \geq 1)\).

     Formally, an execution \( \{s_1, \ldots, s_m\} \) satisfies this constraint if and only if there are ground task atoms \( p(\mathcal{T}_1), \ldots, p(\mathcal{T}_n) \) that executed at some states \( s_{i_1}, \ldots, s_{i_n} \) (i.e., \( s_{i_1} \xrightarrow{p(\mathcal{T}_1)} s_{i_1+1}, \ldots, s_{i_n} \xrightarrow{p(\mathcal{T}_n)} s_{i_n+1} \) such that \( p(\mathcal{T}) \) matches \( p(\mathcal{T}_1), p(\mathcal{T}_2), \ldots, p(\mathcal{T}_n) \).\(^{11}\)

   - **absence(p(\mathcal{T})):** task \( p \) must not execute.

     Formally, an execution \( \{s_1, \ldots, s_m\} \) satisfies this constraint if and only if there is no state \( s_i \) in that execution such that \( s_{i} \xrightarrow{p(\mathcal{T})} s_{i+1} \) and \( p(\mathcal{T}) \) matches \( p(\mathcal{T}) \).

\(^{11}\)The notation \( s \xrightarrow{p} s' \) was introduced in (2) in Section 3. Recall that since \( p(\mathcal{T}_1), \ldots, p(\mathcal{T}_n) \) cause state transitions, they are primitive update tasks.
- exactly\(_n\)(\(p(\mathcal{T})\)): task \(p\) must execute exactly \(n\) times (\(n \geq 1\)).

An execution \(\langle s_1, ..., s_m \rangle\) satisfies this constraint if and only if it satisfies \(\text{atleast}_n(p(\mathcal{T}))\) but not \(\text{atleast}_{n+1}(p(\mathcal{T}))\).

2. Serial constraints:

- \text{after}(p(\mathcal{T}) \rightarrow q(\mathcal{U}))\): whenever \(p\) executes, \(q\) must execute after it. Task \(q\) is not required to execute immediately after \(p\), and several other instances of \(p\) might execute before \(q\) actually does.

Formally, an execution \(\langle s_1, ..., s_m \rangle\) satisfies this constraint if and only if whenever there is a state \(s_i\) in this execution, such that \(s_i \xrightarrow{p(\mathcal{T})} s_{i+1}\) and \(p(\mathcal{T})\) matches \(p(\mathcal{T})\), there must be a state \(s_j\) in that same execution such that \(j \geq i + 1\), \(s_j \xrightarrow{q(\mathcal{U})} s_{j+1}\) and \(\text{refine}(q(\mathcal{U}); p(\mathcal{T}), p(\mathcal{T}))\) matches \(q(\mathcal{U})\).

For instance, if the above constraint has the form \(\text{after}(p(_\text{foo}_,\_ \rightarrow q(_\text{foo}_,\_))\) then the sequence \(p(a, 1), q(2, a)\) is a valid execution, but \(p(a, 1), q(2, b)\) is not.

- \text{before}(p(\mathcal{T}) \leftarrow q(\mathcal{U}))\): whenever \(q\) executes, it must be preceded by an execution of \(p\). Task \(p\) does not have to execute immediately prior to \(q\).

An execution \(\langle s_1, ..., s_m \rangle\) is said to satisfy this constraint if and only if whenever there is a state \(s_i\) in this execution such that \(s_i \xrightarrow{q(\mathcal{U})} s_{i+1}\) and \(q(\mathcal{U})\) matches \(q(\mathcal{U})\), there must be a state \(s_j\) in that same execution such that \(j \leq i - 1\), \(s_j \xrightarrow{p(\mathcal{T})} s_{j+1}\), and \(\text{refine}(p(\mathcal{T}); q(\mathcal{U}), q(\mathcal{U}))\) matches \(p(\mathcal{T})\).

For instance, if the constraint is \(\text{before}(p(_\text{foo}_,\_ \leftarrow q(_\text{foo}_,\_))\) then the sequence \(p(a, 1), q(2, a)\) is a valid execution, but \(p(a, 1), q(2, b)\) is not.

- \text{blocks}(p(\mathcal{T}) \rightarrow q(\mathcal{U}))\): if task \(p\) executes then task \(q\) cannot execute in the future.

Formally, an execution \(\langle s_1, ..., s_m \rangle\) satisfies this constraint if and only if whenever there is a state \(s_i\) in this execution such that \(s_i \xrightarrow{q(\mathcal{U})} s_{i+1}\) and \(q(\mathcal{U})\) matches \(q(\mathcal{U})\), there is no state \(s_j\) in that execution such that \(j > i\), \(s_j \xrightarrow{q(\mathcal{U})} s_{j+1}\), and \(\text{refine}(q(\mathcal{U}); p(\mathcal{T}), p(\mathcal{T}))\) matches \(q(\mathcal{U})\).

- \text{between}(p(\mathcal{T}) \rightarrow q(\mathcal{U}) \leftrightarrow r(\mathcal{V}))\): task \(q\) must execute between any two executions of \(p\) and \(r\). That is, after an execution of \(p\), any subsequent execution of \(r\) has to wait until \(q\) is executed.

An execution \(\langle s_1, ..., s_m \rangle\) satisfies this constraint if and only if whenever there are states \(s_i, s_k\) (\(k > i + 1\)) such that \(s_i \xrightarrow{q(\mathcal{U})} s_{i+1}\), \(s_k \xrightarrow{r(\mathcal{V})} s_{k+1}\), \(p(\mathcal{T})\) matches \(p(\mathcal{T})\), and \(r(\mathcal{V})\) matches \(r(\mathcal{V})\), then there must be a state \(s_j\) such that \(i < j < k\), \(s_j \xrightarrow{q(\mathcal{U})} s_{j+1}\), and both \(\text{refine}(q(\mathcal{U}); p(\mathcal{T}), p(\mathcal{T}))\) and \(\text{refine}(q(\mathcal{U}); r(\mathcal{V}), r(\mathcal{V}))\) match \(q(\mathcal{U})\).

- \text{not\_between}(p(\mathcal{T}) \rightarrow q(\mathcal{U}) \leftrightarrow r(\mathcal{V}))\): task \(q\) must not execute between any pair of executions of \(p\) and \(r\). Thus, if \(q\) executes after \(p\), no future execution of \(r\) is possible.

An execution \(\langle s_1, ..., s_m \rangle\) satisfies this constraint if and only if whenever there are states \(s_i, s_k\) (\(k > i + 1\)) such that \(s_i \xrightarrow{q(\mathcal{U})} s_{i+1}\), \(s_k \xrightarrow{r(\mathcal{V})} s_{k+1}\), \(p(\mathcal{T})\) matches \(p(\mathcal{T})\), and \(r(\mathcal{V})\) matches \(r(\mathcal{V})\), then there is no state \(s_j\) such that \(i < j < k\), \(s_j \xrightarrow{q(\mathcal{U})} s_{j+1}\), and both \(\text{refine}(q(\mathcal{U}); p(\mathcal{T}), p(\mathcal{T}))\) and \(\text{refine}(q(\mathcal{U}); r(\mathcal{V}), r(\mathcal{V}))\) match \(q(\mathcal{U})\).

3. Immediate serial constraints:
- right_after\((p(\overline{T}) \rightarrow q(\overline{T}))\): whenever \(p\) executes, \(q\) must execute immediately after it.

Formally, an execution \(\langle s_1, ..., s_m \rangle\) satisfies this constraint if and only if whenever there is a state \(s_i\) in this execution such that \(s_i \xrightarrow{p(\overline{T})} s_{i+1}\) and \(p(\overline{T})\) matches \(p(\overline{T})\), then \(s_{i+1} \xrightarrow{q(\overline{T})} s_{i+2}\) must hold and \(\text{refine}(q(\overline{T}); p(\overline{T}), p(\overline{T}))\) must match \(q(\overline{T})\).

- right_before\((p(\overline{T}) \leftarrow q(\overline{T}))\): whenever \(q\) executes, \(p\) must have been executed immediately before it.

Formally, an execution \(\langle s_1, ..., s_m \rangle\) satisfies this constraint if and only if whenever there is a state \(s_i\) in this execution such that \(s_i \xrightarrow{q(\overline{T})} s_{i+1}\) and \(q(\overline{T})\) matches \(q(\overline{T})\), then \(s_{i-1} \xrightarrow{p(\overline{T})} s_i\) must hold and \(\text{refine}(p(\overline{T}); q(\overline{T}), q(\overline{T}))\) must match \(p(\overline{T})\).

- not_right_after\((p(\overline{T}) \rightarrow q(\overline{T}))\): whenever \(p\) and \(q\) execute, \(q\) must not execute immediately after \(p\). That is, after \(p\) there must be an execution of a task other than \(q\) before \(q\) is allowed again.

An execution \(\langle s_1, ..., s_m \rangle\) is said to satisfy this constraint if and only if whenever there is a state \(s_i\) in this execution such that \(s_i \xrightarrow{p(\overline{T})} s_{i+1}\), where \(p(\overline{T})\) matches \(p(\overline{T})\), and \(s_{i+1} \xrightarrow{r(\overline{T})} s_{i+2}\) for some \(r(\overline{T})\), then \(\text{refine}(q(\overline{T}), p(\overline{T}), p(\overline{T}))\) must not match \(r(\overline{T})\).

4. **Composite constraints**: If \(C_1, C_2 \in \textit{Constr}\) then so are \(C_1 \land C_2\) (a conjunctive constraint) and \(C_1 \lor C_2\) (a disjunctive constraint).

Nothing else is in \(\textit{Constr}\). \(\square\)

Note the use of different arrows between the arguments in some of the constraints. The convention here is that the task at the tail of the arrow represents the condition of the constraint (\textit{if or whenever the task executes}) and the task at the head of the arrow indicates the effect of the constraint (\textit{the task must or must not execute in a given relationship to the task at the tail of the arrow}). We use strong arrows to indicate immediacy (\textit{execution must take place right before or after}) and dashed arrows indicate that immediacy is not required. Slashed arrows indicate negative relationships (e.g., the task in the head must \textit{not} execute). Note also that the negation of right_before can be defined as follows:

\[
\text{not_right_before}(p(\overline{T}) \leftarrow q(\overline{T})) \equiv \text{not_right_after}(p(\overline{T}) \rightarrow q(\overline{T})).
\]

The following is an example of a legal constraint in \(\textit{Constr}\):

\[
\text{atleast}_2(p(\_X, 1, \_X)) \land \text{exactly}_3(q(\_X, \_X)) \land \text{right_after}(p(\_X, \_X) \rightarrow r(\_X))
\] (5)

The constraint \(\text{atleast}_2(p(\_X, 1, \_X))\) requires that \(p\) executes at least twice and arguments 1 and 3 have the same value in each execution, while the second argument is the integer 1. In \(\text{right_after}(p(\_X, \_X) \rightarrow r(\_X))\), the placeholder \_X is shared between \(p\) and \(r\). This means that whenever \(p\) is executed, \(r\) must follow immediately and \(r\)'s second argument must be the same as the first argument in \(p\).

We can now show how the constraints from Figure 3 can be represented in \textit{ServLog}:
1. \((\text{atmost}_n(p(_,1)) \land \text{absence} (\text{book\_shipper}(_,_,_)))\) \\
\lor \ (\text{atleast}_1(p(_,1))) \\
\land \ \text{after}(\text{accept}(\text{Ord#},_), \text{book\_shipper}(\text{Ord#},_,_))

where \(\text{atmost}_n(p)\) is a shorthand for \\
absence(p) \lor \text{exactly}_1(p) \lor \text{exactly}_2(p) \lor \ldots \lor \text{exactly}_n(p)

2. \(\text{absence} (\text{pay\_per\_item}(\_))\) \\
\lor \text{right\_before}(\text{pay\_one\_item}(\text{Ord#}) \leftarrow \text{deliver}(\text{Ord#}))

3. before(\text{inform\_client}(\text{Ord#},_), \text{payment\_guarantee}(\text{Ord#},_))

4. \text{exactly}_1(\text{deliver}()\)

5. \(\text{absence} (\text{full\_payment}(\_)) \lor \text{blocks} (\text{full\_payment}(\text{Ord#}) \rightarrow \text{deliver}(\text{Ord#}))\)

6. before(\text{pay}(\text{Ord#}) \leftarrow \text{book\_shipper}(\text{Ord#},_,_))\)

Many other types of constraints can be naturally expressed in \(\text{CONSTR}\) as shown below:

- \(\text{atmost}_n(p(_,1))\) — task \(p\) can execute at most \(n\) times and each time the second argument must be 1. This constraint was defined in item 1 above.
- \(\text{absence} (p(_,4)) \lor \text{atleast}_1(q(2,,3))\) — if \(p\) is executed with its second argument 4, then \(q\) must also execute (before or after \(p\)) and its first and last arguments must be 2 and 3 respectively.
- \((\text{absence}(p()) \lor \text{atleast}_1(q())\)) \land \text{absence}(q()) \lor \text{atleast}_1(p())\) — if \(p\) is executed, then \(q\) must also be executed, and vice versa.
- \(\text{after}(p(\_X) \rightarrow q(\_X)) \land \text{before}(p(\_X) \leftarrow q(\_X))\) — every occurrence of task \(p\) must be followed by an occurrence of task \(q\) with the same argument and there must be an occurrence of \(p\) before every occurrence of \(q\) and their arguments must be the same.
- \(\text{absence} (p(\_)) \lor \text{between}(p(\_X) \rightarrow q(\_X) \leftarrow p(\_))\) — if task \(p\) is executed then \(q\) must execute after it, with the same argument, and before that there can be no other \(p\).
- \(\text{absence}(q(\_)) \lor (\text{before}(p(\_) \leftarrow q(\_) \land \text{between}(q(\_X) \rightarrow p(\_) \leftarrow q(\_)))\) — if task \(q\) is executed, it has to be preceded by an occurrence of \(p\). The next instance of \(q\) can execute only after another occurrence of \(p\).
- \(\text{between}(p(\_X) \rightarrow q(\_X) \leftarrow p(\_X)) \land \text{between}(q(\_X) \rightarrow p(\_X) \leftarrow q(\_X))\) — tasks \(p\) and \(q\) must alternate when they execute with the same argument.
- \(\text{right\_after}(p(\_X) \rightarrow q(\_)) \land \text{right\_before}(p(\_) \leftarrow q(\_))\) — executions of \(p\) and \(q\) must be next to each other with no intervening tasks in-between.
- \(\text{absence}(p(\_)) \lor \text{absence}(q(\_))\) — it is not possible for \(p\) and \(q\) to execute in the same service process run.
- \(\text{not\_between}(p(\_X) \rightarrow q(\_) \leftarrow p(\_X)) \land \text{not\_between}(q(\_X) \rightarrow p(\_) \leftarrow q(\_X))\) — \(q\) must not execute between any two executions of \(p\) with the same arguments, and \(p\) must not execute between any two executions of \(q\) with the same arguments.

4.3. The Service Contract Assumption

We now introduce modeling assumptions, which tighten the form of the constraints and tasks that are allowed in service processes. These assumptions do not limit the modeling power of the language in the sense that any service process can be represented by another process that satisfies these assumptions. However, these assumptions greatly simplify the reasoning system of Section 5.

Definition 4.12 (Service Contract Assumption) A service process \(G\) and a set of constraints \(C\) satisfy the service contract assumption if and only if the set of constraints \(C\) is based on primitive update tasks and the primitive tasks of \(G\) satisfy the independence assumption. \(\Box\)

The last two notions in this definition are spelled out in Definitions 4.13 and 4.14 below. Also recall that a primitive task is one that is not defined by a rule and a primitive update task is just an elementary CTR update.
Definition 4.13 (Constraints based on primitive update tasks) A set of constraints is based on primitive update tasks if and only if all tasks appearing in the set are primitive update tasks.

This restriction does not limit the modeling power of the language, since one can always instrument composite tasks in such a way that the resulting set of constraints will be based on primitive update tasks. More specifically, one can insert “bounding” primitive tasks at various locations in the definition of composite tasks, and then transform constraints on composite tasks into constraints on those bounding primitive tasks. These bounding tasks are defined as no-ops and their only purpose is to capture the various stages in the life cycle of a task. Examples include the beginning and end of a task, the beginning and end of an iteration, and so on.

The rationale behind restricting constraints to primitive tasks is that specifying constraints directly over composite tasks can be highly ambiguous. For instance, what should the sentence “task \( b \) must start after task \( a \)” mean exactly? Should \( b \) start after \( a \) begins or after \( a \) ends? Similar ambiguity exists with other constraints, such as before and between constraints. Requiring that constraints are based on primitive update tasks avoids ambiguity and complications without limiting the modeling power.

The following example illustrates the process of inserting bounding tasks to delineate the beginning and the end of a composite task:

- Every non-iterative composite task of the form \( p \leftarrow \Omega \) can be changed to:
  \[
  p \leftarrow p_{\text{start}} \otimes \Omega \otimes p_{\text{end}}
  \]

- Every iterative composite task, for example, of the form \( p \leftarrow (\Phi \otimes p) \lor \Psi \), can be changed to:
  \[
  p \leftarrow (p_{\text{start}} \otimes \Phi \otimes p \otimes p_{\text{end}}) \lor (p_{\text{start}} \otimes \Psi \otimes p_{\text{end}})
  \]

In fact, there are many other ways to insert bounding tasks, which would enable many more kinds of constraints. For instance,

\[
 p \leftarrow p_{\text{start}} \otimes \Phi_{\text{start}} \otimes \Phi \otimes \Phi_{\text{end}} \otimes p \otimes p_{\text{end}}
\]

A constraint such as \( \text{after}(p \rightarrow q) \), where \( p \) and \( q \) are composite tasks, can now be interpreted as \( \text{after}(p_{\text{start}} \rightarrow q_{\text{start}}) \), or \( \text{after}(p_{\text{end}} \rightarrow q_{\text{start}}) \), or \( \text{after}(p_{\text{start}} \rightarrow q_{\text{start}}) \land \text{after}(p_{\text{end}} \rightarrow q_{\text{end}}) \). By exposing the bounding primitive subtasks, ServLog enables many kinds of constraints that cannot be specified on composite tasks directly. For instance, \( \text{before}(p_{\text{start}} \rightarrow q_{\text{end}}) \) or \( \text{between}(p_{\text{start}} \rightarrow q_{\text{start}} \leftrightarrow p_{\text{end}}) \).

Definition 4.14 (Independence Assumption) Two primitive update tasks are said to be independent if and only if they are represented by disjoint binary relations over database states. A service process satisfies the independence assumption if and only if all its primitive tasks are independent of each other.

Independence implies that any transition between a pair of states is caused by precisely one primitive task, and no other task can cause that transition.

Any set of primitive tasks can be instrumented so that the tasks would become independent. For example, each primitive task, \( t \), can be instrumented to insert a unique token every time it executes. Specifically, \( t \) might insert \( \text{token}(t, 0) \) on the first execution and then \( \text{token}(t, 1) \), \( \text{token}(t, 2) \), etc., on subsequent executions. As a result, any transition between any pair of states would be possible by at least one primitive task.

Without the independence assumption it is hard to come up with an effective algorithm for checking satisfaction of constraints by service executions, and it is hard to develop a simple enough proof theory for finding service executions that satisfy such constraints. To see this, suppose that the independence assumption is not satisfied and

\[12\text{Recall from Section 3 that primitive tasks are represented in CTR by elementary updates, and an elementary update is a binary relation over database states.}\]
there are two distinct primitive update tasks such that \( s \xrightarrow{p} s' \) and \( s \xrightarrow{q} s' \) hold. Suppose that we are now trying to execute \( p \) at state \( s \). In the presence of constraints such as \texttt{before}(r \cdot q), \texttt{absence}(q), \) and the like, it would be hard to determine whether \( p \) can be executed, since one must first determine if execution of \( p \) amounts to execution of another, prohibited task, such as \( q \), in this example.

5. Reasoning about Contracts in \textit{ServLog}

We begin with an example that shows how service contracting and contract execution are intended to work. The example illustrates most of the aspects of service modeling introduced in Section 4: service processes (control and data flows), client contract requirements, and service policies. The last two are represented via constraints. Section 5.2 formalizes the decision procedure as an extension to the proof procedure of the original CTR.

5.1. Informal Example

For simplicity, the example uses propositional tasks only, but the proof procedure in Section 5.2 is designed to work for the more general case where tasks have arguments.

For concreteness and to illustrate the interactive aspect of our model, we assume the following division among the tasks involved:

- Service tasks: \( a, f, g \)
- Client tasks: \( d, e, h \)

\textit{Service process:}

\textbf{Process formula:}
\[
a \otimes (B | C) \otimes (g \lor h)
\]

\textbf{Rules:}
\[
B \leftarrow d \lor e
\]
\[
C \leftarrow (f \otimes C) \lor \text{state}
\]

\textit{Client contract requirements:}

\[
\texttt{absence}(e) \land \texttt{atleast}_2(f)
\]

\textit{Service policy:}

\[
\texttt{after}(d \cdot f) \land \texttt{absence}(g)
\]

\textit{Service contracting.} As explained at the end of Section 2, service contracting is an interactive decision procedure that checks whether an execution of the service process exists and satisfies both the service policies and the client contract requirements. In our example this amounts to finding an execution path of (6) such that the constraints (7) and (8) are satisfied. For example, \( \{a,d,f,f,h\} \) is an execution path for (6) path that satisfies the constraints, but \( \{a,f,f,d,h\} \) is an execution that violates \texttt{after}(d \cdot f). The proof procedure introduced in Section 5.2 is designed to find paths on which the constraints are satisfied, if such paths exist. Note it does not matter if the path that was found will actually be the one to be executed—all that is needed is to find out if the contract is satisfiable. If service contracting does not find a path that satisfies the constraints, it means that no execution will ever be successful and the contract is unsatisfiable.
Service contract execution. If a contract is satisfiable, its execution deals with the actual interactions performed by the client and the service. The idea is based on the same proof theory that is used for service contracting, but it is applied differently. When a task is chosen for execution by the client or the service, the proof theory of Section 5.2 checks if acceptance of that task still leaves the possibility of a successful execution of the remainder of the service process (that satisfies the constraints). If so, the task is accepted and executed; otherwise the task is rejected and a different task must be chosen for execution by the agent in question. Note that such a task must exist because when accepting the previous task we must have checked that some continuation is possible.

Contract execution for our example works as follows:

1. Suppose the service selects task $a$. (This is the only task that can possibly be chosen according to our process specification.) We already checked (while doing service contracting) that there is a legal execution that starts with $a$, so the task is accepted.
2. The remainder of the process is $((d \lor e) \land (f \lor C) \lor \text{state}) \lor (g \lor h)$ (where $B$ and $C$ are replaced with their definitions). The next step can be taken either by the client (e.g., by picking $d$ or $e$) or by service (e.g., by picking $f$). It is also possible for the system itself to execute an internal action by picking \text{state} (here the client and the service “take a short break”). Let us assume that the client takes initiative and picks $e$ for execution. The system checks if $e$ can be executed and finds that this would violate the constraint \text{absence}(e); so $e$ is rejected. Suppose that the client does not give up and selects $d$ for execution. Again, the system checks if $d$ is allowed to execute given the constraints. In this case no constraint forbids the execution of $d$ because $a, d, f, f, f, h$ is a valid execution of the remainder of the process, so $d$ is accepted.

It would be very inefficient if the system had to go back to the beginning of the path in order to check if acceptance of an action permits a valid continuation. To avoid this, we modify the constraints after acceptance of each action so that we will never have to look back in order to decide whether to accept or reject an action. For instance, after accepting $d$ the system revises the constraints by replacing after($d \rightarrow f$) with \text{atleast}_1(f). This is possible because after the execution of $d$ the constraint after($d \rightarrow f$) will be satisfied iff $f$ is executed at some point in the future, whence \text{atleast}_1(f). The system now checks whether an execution of the remaining process $((f \lor C) \lor \text{state}) \lor (g \lor h)$ exists under the updated set of constraints $\text{absence}(e) \land \text{atleast}_2(f) \land \text{atleast}_1(f) \land \text{absence}(g)$ (since we have two constraints \text{atleast}_2(f) and \text{atleast}_1(f), the last one can be removed).

3. For the remainder of the process, $((f \lor C) \lor \text{state}) \lor (g \lor h)$, only $f$ and \text{state} can be chosen. Let us assume that the internal action \text{state} is picked for execution. The system now checks whether an execution of the remaining part of the process, $(g \lor h)$, exists given the updated set of constraints $\text{absence}(e) \land \text{atleast}_2(f) \land \text{absence}(g)$. It is easy to see that $(g \lor h)$ cannot execute to satisfy \text{atleast}_2(f) because there is no $f$ in $(g \lor h)$. Therefore, \text{state} is rejected. The only other way to proceed is for the service to pick $f$ (recall that $f$ can be executed only by the service). Proceeding as before, the proof theory would check if $f$ can execute. There are no constraints to prevent that, so the system updates the constraints and checks if a legal continuation is possible. The update replaces \text{atleast}_2(f) with \text{atleast}_1(f), since executing $f$ means that the number of required occurrences of $f$ decreases by 1. The remaining part of the process, $C \lor (g \lor h)$ has a legal execution $[f]$ with respect to the updated set of constraints $\text{absence}(e) \land \text{atleast}_1(f) \land \text{absence}(g)$, so $f$ is accepted.

4. To proceed, we need to expand $C$ using the rule in (6), which gives us $((f \lor C) \lor \text{state}) \lor (g \lor h)$ — the same process as in the previous step. Although the set of constraints has now changed, because of \text{atleast}_1(f) the task $f$ must still be executed for the same reasons as in the previous step. This task is accepted because a legal continuation exists, and now the set of constraints gets changed to $\text{absence}(e) \land \text{absence}(g)$.

5. Now the remainder of the process is $(C \lor \text{state}) \lor (g \lor h)$ and \text{state} can be successfully picked for execution and the remainder of the process becomes $g \lor h$. Either $g$ can now be attempted by the client or $h$ by the service. However, the constraint $\text{absence}(g)$ prevents the former, so the service proceeds by picking $h$ and, since no constraint prevents it from going ahead, it is executed. At this point, the remainder of the process is empty and we are done.
Note that other executions are also possible and could have been taken. For instance, if the service continued to press initiative in step 2, it could have picked f and the execution could have become a,f,d,f,h or a,f,f,d,f,h.

As we saw above, both service contracting and contract execution rely on the same inference rules which do not explicitly differentiate between client and service tasks, however the difference plays out in the way the inference rules are applied.

5.2. Proof System

Let \( C \in C\text{ONSTR} \) be a constraint (which can be composite), where \( C\text{ONSTR} \) includes both the service policy and the client contract requirements. Let \( G \) be a service process and \( G \) and \( C \) satisfy the service contracts assumption. We consider the following reasoning problems in \( \text{ServLog} \):

1. **Contracting**: The problem of determining if contracting for a service is possible amounts to finding out if there is an execution of the CTR formula \( G \land C \). Formally, contracting aims to determine if there is a path on which \( G \land C \) is true in every multipath structure that makes all composite task definitions true.

2. **Contract Execution**: The problem of contract execution amounts to producing an interactive proof that \( G \land C \) can execute along some path. In that proof, the client and the service take turns that are prescribed by the process specification and by the ownership of the primitive tasks, as illustrated in the previous subsection. This proof must be constructive—a sequence of applications of the inference rules of CTR, which starts with an axiom and ends with the aforesaid formula \( G \land C \). Each such proof provides a way to execute the process without violating any of the constraints in \( C \).

The rest of this section develops a proof theory for formulas of the form \( G \land C \), where \( G \) is a service process and \( C \in C\text{ONSTR} \).

To simplify the matters, we will assume that the service process \( G \) has no disjunctions in the rule bodies and in its CTR goal part. This does not limit the generality, as such disjunctions can always be eliminated through a simple transformation similar to the one in classical logic. For instance, the disjunction in

\[
p \leftarrow q \land (r \lor s) \land t
\]

can be eliminated by transforming this rule into

\[
p \leftarrow q \land \text{newpred} \land t
\]
\[
\text{newpred} \leftarrow r
\]
\[
\text{newpred} \leftarrow s
\]

**Hot components.** We recall the notion of **hot components** of a CTR goal from [10]: hot(\( \psi \)) is a set of subformulas of \( \psi \) that are “ready to be executed” and corresponds to the notion of goal selection in SLD-style resolution proof theories. Hot components are defined inductively as follows:

1. hot(()) = \{ \}, where () is the empty goal
2. hot(\( \psi \)) = \{ \( \psi \) \}, if \( \psi \) is an atomic formula
3. hot(\( \psi_1 \otimes \ldots \otimes \psi_n \)) = hot(\( \psi_1 \))
4. hot(\( \psi_1 | \ldots | \psi_n \)) = hot(\( \psi_1 \)) \cup \ldots \cup hot(\( \psi_n \))
5. hot(\( \otimes \psi \)) = \{ \otimes \psi \}

**Additional constraints.** The set of constraints is changing as tasks in the process execute. The exact mechanism is explained in the inference rule “executing primitive tasks,” below. It involves three new constraints: force(\( p(T) \)), suspend(\( p(T) \)), and next_right_before(\( q(\pi) \rightleftarrows p(T) \leftleftarrows q(\pi) \)), plus a generalization of the constraint before, which allows exceptions. We did not introduce these before, since the new constraints are technical means by which the proof theory works and they are unlikely to be employed by users directly.

The meaning of the constraint force(\( p(T) \)), where \( p(T) \) is a task pattern for a primitive update, is that the very next task to be executed must match \( p(T) \). More precisely, an execution \( \langle s_1, s_2, \ldots, s_n \rangle \) satisfies force(\( p(T) \)) if \( s_1 \rightarrow s_2 \) holds and \( p(T) \) matches \( p(T) \).
The constraint \(\text{suspend}(p(\overline{t}))\) means that no task that matches \(p(\overline{t})\) can execute at the current state. Formally, an execution \(\langle s_1, s_2, ..., s_n \rangle\), such that \(s_1 \rightarrow s_2\) holds for some task atom \(r(\overline{V})\), satisfies \(\text{suspend}(p(\overline{t}))\) if \(r(\overline{V})\) does not match \(p(\overline{t})\).

The constraint \(\text{next_right_before}(q(\overline{u}) \leftarrow p(\overline{t}) \leftrightarrow q(\overline{u}))\) says that \(q(\overline{u})\) must be immediately preceded by \(p(\overline{t})\) unless it is the first task to be executed. That first task can be a \(q\)-task, if it matches \(q(\overline{u})\). Formally, an execution \(\langle s_1, s_2, ..., s_n \rangle\) satisfies this constraint if and only if either \(\langle s_1, s_2, ..., s_n \rangle\) satisfies \(\text{right_before}(p(\overline{t}) \leftarrow q(\overline{u}))\) or \(s_1 \rightarrow s_2\) holds, where \(r(\overline{V})\) matches \(q(\overline{u})\), and \(\langle s_2, ..., s_n \rangle\) satisfies \(\text{right_before}(p(\overline{t}) \leftarrow q(\overline{u}))\).

The generalization of \(\text{before}\) has the following syntax:

\[
\text{before}(p(\overline{t}) \leftarrow q(\overline{u}) \\setminus \{q(\overline{u}_1), ..., q(\overline{u}_n)\})
\]

where \(p(\overline{t}), q(\overline{u}), q(\overline{u}_1), ..., q(\overline{u}_n), n \geq 0\) (\(n = 0\) means that the sets of exceptions are empty), are task patterns.

The use of "\(\setminus\)" here indicates that before\((p(\overline{t}) \leftarrow q(\overline{u}))\) must hold, except for the tasks that match one of the exceptions \(\{q(\overline{u}_1), ..., q(\overline{u}_n)\}\).

Formally, an execution \(\langle s_1, ..., s_m \rangle\) satisfies the constraint (9) if and only if whenever there is a state \(s_i\) in this execution such that \(s_i \rightarrow s_{i+1}\) and \(q(\overline{u}_i)\) matches \(q(\overline{u})\) but none of the \(q(\overline{u}_1)\)'s then there is \(j, j < i\), such that \(s_j \rightarrow s_{j+1}\) holds and \(p(\overline{t})\) matches \(\text{refine}(p(\overline{t}); q(\overline{u}), q(\overline{u}))\).

Note that, when the set of exceptions is empty, the constraint (9) reduces to the old form of the before-constraint. Therefore, to simplify the language, in the rest of this section we will be using only the generalized form of this constraint.

**Substitutions.** As usual in logic proof theories, we will rely on the notion of substitution, which is a mapping from variables to terms. If \(\sigma\) is a substitution and \(\psi\) is a service process or a term then we write \(\psi\sigma\) for the result of applying the substitution \(\sigma\) to \(\psi\). We call \(\psi\sigma\) an instance of \(\psi\). If \(\psi\sigma\) has no variables left, we say that \(\psi\sigma\) is a ground instance and that \(\sigma\) is a ground substitution.

**Sequents.** Let \(P\) be a set of composite task definitions. The proof theory manipulates expressions of the form \(P, s \rightarrow (\exists)\psi \land C\), called sequents, where \(P\) is a set of task definitions, \(s\) is an identifier for the underlying database state, \(\psi\) is a CTR goal, and \(C\) is a (possibly composite) constraint, which may include the constraints in \(\text{Constr}\) as well as the new constraints (\text{force}, \text{suspend}, etc.) introduced just above. Informally, a sequent is a statement that \((\exists)\psi\), which is defined by the rules in \(P\), can execute along some path that starts at state \(s\) so that all the constraints in \(C\) will be satisfied. Each inference rule has two sequents, one above the other. This is interpreted as: if the upper sequent is inferred, then the lower sequent should also be inferred. As in classical logic, any instance of an answer-substitution is a valid answer to a query.

The inference system presented here extends the inference system for Horn CTR [10] with two additional inference rules (Rules 2 and 3). Other rules from [10] (e.g., Rule 6) are also significantly modified. The new system reduces to the old one when the set \(C\) of constraints is empty. The new system also extends and simplifies the proof theory developed in [28].

All rules and the axioms operate with sets of constraints, which get modified as a result of the rule application. However, some of the rules require the constraint to be a conjunction of the existence, serial, and the additional constraints introduced in this section. We call such constraints conjunctive. A conjunctive constraint can be viewed as a set, so we will often write \(c \in C\) meaning that \(c\) is a conjunct in \(C\).

**Axiom.** \(P, s \rightarrow (\cdot) \land C\), for any database state identifier \(s\), where \(C\) is a conjunctive constraint that does not contain constraints of the form \(\text{force}(\overline{T}), \text{atleast}_k(\overline{T})\), and \(\text{exactly}_k(\overline{T})\), where \(k \geq 1\).

**Inference Rules.** In Rules 1-7 below, \(\sigma\) denotes a substitution, \(\psi\) and \(\psi'\) are service processes, \(C\) and \(C'\) are sets of constraints, \(s, s_1, s_2\) are database state identifiers, and \(p\) is a task.

1. Eliminating disjunctive constraints: Let \(\psi\) be a CTR goal and \(C'\) a disjunct in the disjunctive normal form of \(C\) (i.e., \(C'\) is a conjunctive constraint). Then
Commutativity with respect to builtin tests

Applying composite task definitions

Solving builtin tests

Executing primitive update tasks

Executing query tasks

Note that $C'$ is a conjunction of existence and serial constraints.

2. Solving builtin tests: Let $\chi$ be a conjunction of builtin test tasks. Suppose there is a ground substitution $\sigma$ such that $\chi\sigma$ evaluates to true. Then

$$
\begin{align*}
P, s & \rightarrow \psi \land C' \\
\vdash & \\
P, s & \rightarrow \psi \land C
\end{align*}
$$

3. Commutativity with respect to builtin tests: Let $\chi$ be a conjunction of builtin test tasks and $\psi$ a CTR goal. Then

$$
\begin{align*}
P, s & \rightarrow (\exists) (\psi \land \chi) \land C \\
\vdash & \\
P, s & \rightarrow (\exists) (\chi \land \psi) \land C
\end{align*}
$$

4. Applying composite task definitions: Let $r \leftarrow \beta$ be a rule in $P$, and assume that its variables have been renamed so that none are shared with $\psi$. If $p$ and $r$ unify with the most general unifier $\sigma$ then

$$
\begin{align*}
P, s & \rightarrow (\exists) \left( \psi' \sigma \land C \right) \\
\vdash & \\
P, s & \rightarrow (\exists) \left( \psi \land C \right)
\end{align*}
$$

where $\psi'$ is obtained from $\psi$ by replacing a hot occurrence of $p$ with $\beta$.

5. Executing query tasks: Suppose that $p\sigma$ and $\psi'\sigma$ share no variables and either (i) $p$ is a primitive query task such that $(\exists)p\sigma$ is true in the state $s$; or (ii) $p = \text{state}$ and $\sigma$ is the identity substitution. Then

$$
\begin{align*}
P, s & \rightarrow (\exists) \left( \psi' \sigma \land C \right) \\
\vdash & \\
P, s & \rightarrow (\exists) \left( \psi \land C \right)
\end{align*}
$$

where $\psi'$ is obtained from $\psi$ by deleting a hot occurrence of $p$.

6. Executing primitive update tasks: Let $p\sigma$ be a primitive update task such that $s_1 \xrightarrow{p\sigma} s_2$. $p \in \text{hot}(\psi)$, and $C$ has no constraint of either of the forms below. In the description below, we assume that $\bar{p}$ is a task pattern such that $p\sigma$ matches $\bar{p}$ and that $\bar{r}$ denotes an arbitrary task pattern:

- absence($\bar{p}$)
- suspend($\bar{p}$)
- force($\bar{r}$) such that $p\sigma$ does not match $\bar{r}$
- before($\bar{r} \leadsto \bar{p} \setminus \{\bar{p}_1, \ldots, \bar{p}_k\}$) $k \geq 0$ and $p\sigma$ matches neither $\bar{r}$ nor any of the $\bar{p}_i$'s
- right_before($\bar{r} \leftarrow \bar{p}$)
- next_right before($\bar{r}' \leftarrow \bar{p}$), where $p\sigma$ does not match $\bar{p}'$.

Then the inference rule has the following form:

$$
\begin{align*}
P, s_2 & \rightarrow (\exists) \left( \psi' \sigma \land C' \right) \\
\vdash & \\
P, s_1 & \rightarrow (\exists) \left( \psi \land C \right)
\end{align*}
$$

where $C$ is a conjunctive constraint, $\psi'$ is obtained from $\psi$ by deleting the hot component $p$, and $C'$ is constructed out of $C$ as follows.

Initially, $C'$ is empty. Then constraints are added to it according to the cases below. (Again, in all these cases, we assume that $p\sigma$ matches $\bar{p}$ and $\bar{r}$, $s$ are arbitrary task patterns.)

(a) If $\text{atleast}_n(\bar{p}) \in C$, where $n > 1$, add the following to $C'$:

- $\text{atleast}_{n-1}(\bar{p})$;

(b) If $\text{exactly}_n(\bar{p}) \in C$, where $n > 1$, add the following to $C'$:

- $\text{exactly}_{n-1}(\bar{p})$: 
Theorem 1

The above inference system is sound and complete for proving constrained service processes, if the service processes and the constraints satisfy the service contracts assumptions.
Example. The following example illustrates the inference procedure.

Goal \( G \): \( \exists x ((p(x) \otimes q) \mid r(x)) \)

Rules: \( r(x) \leftarrow (xyz(y,z) \otimes s(x,y,z) \otimes r(y)) \)
\( r(x) \leftarrow \text{state} \)

Constraint \( C \): \( \text{atleast}_2(s(_,_,_)) \lor \text{before}(p(_,x) \leftarrow s(_,1,\_)) \)

Here \( p, q, \) and \( s \) are assumed to be elementary tasks, and in the case of \( p \) and \( s \) they can be executed with any integer argument. Furthermore, in this example we assume that the execution of \( p, s, \) and \( q \) modifies the database, as follows: \( p(1) \) adds \( xyz(3,7) \); \( s(1,3,7) \) adds \( xyz(2,7) \); \( s(3,2,7) \) adds \( abc(1) \); and \( q \) deletes \( xyz(3,7) \). These assumptions were needed in order to show that the constraint is satisfied in this example. If we did not take these assumptions, then the constraint might not be satisfied in which case the inference procedure would not infer the constraint. The goal \( G \) can be executed in several ways such that \( C \) is satisfied. We show one possibility, which corresponds to one derivation of the sequent \( P, s \vdash G \land C \), where \( s \) is an identifier for \( \{ \} \), the empty state. In this derivation, we use the top-down method, i.e., we start with the goal and apply the inference rules backwards. Each sequent is derived from the previous one by an inference rule. The deduction succeeds when the last sequent is an axiom. We start with the sequent

\[
P, \{ \} \vdash (\exists x ((p(x) \otimes q) \mid r(x)) \land \text{atleast}_2(s(_,_,_)) \lor \text{before}(p(_,x) \leftarrow s(_,1,\_))
\]

Here, instead of a state identifier \( s \) we put the corresponding database state \( (\{ \}) \) explicitly. To make the sequent easier to read, we will continue doing this in the rest of this example.

Hot components: \( \{ p(x), r(x) \} \). By inference rule 1 (eliminating disjunctive constraints) we obtain:

\[
P, \{ \} \vdash (\exists x ((p(x) \otimes q) \mid r(x)) \land \text{atleast}_2(s(_,_,_))
\]

Hot components: \( \{ p(x), r(x) \} \). By inference rule 4 (composite task definitions) we obtain:

\[
P, \{ \} \vdash (\exists x, y, z ((p(x) \otimes q) \mid (xyz(y,z) \otimes s(x,y,z) \otimes r(y))) \land \text{atleast}_2(s(_,_,_))
\]

Hot components: \( \{ p(x), xyz(y,z) \} \). By inference rule 6, choosing \( p(x) \) and executing this elementary task with the argument \( x = 1 \), we obtain (recall that the execution of \( p(1) \) adds the fact \( xyz(3,7) \) to the database):

\[
P, \{ xyz(3,7) \} \vdash (xyz(y,z) \mid xyz(y,z) \otimes s(1,y,z) \otimes r(y)) \land \text{atleast}_2(s(_,_,_))
\]

Hot components: \( \{ q, xyz(y,z) \} \). By inference rule 5 (executing query tasks):

\[
P, \{ xyz(3,7) \} \vdash q \mid (s(1,3,7) \otimes r(3)) \land \text{atleast}_2(s(_,_,_))
\]

Hot components: \( \{ q, s(1,3,7) \} \). By inference rule 6 applied to the primitive task \( s(1,3,7) \) and the earlier assumption, this execution adds the fact \( xyz(2,7) \) to the database:

\[
P, \{ xyz(3,7), xyz(2,7) \} \vdash \ldots (q \mid r(3)) \land \text{atleast}_1(s(_,_,_))
\]
Hot components: \{q, r(3)\}. By inference rule 6 applied to the primitive task q (which, as mentioned above, deletes \textit{xyz}(3, 7)):

\[ P, \{ xyz(2, 7) \} \leadsto r(3) \land \text{atleast}_1(s(_, _, _)) \]

Hot components: \{r(3)\}. By inference rule 4 (composite task definitions):

\[ P, \{ xyz(2, 7) \} \leadsto (\exists ?y, ?z \ (xyz(?y, ?z) \otimes s(3, ?y, ?z) \otimes r(?y))) \land \text{atleast}_1(s(_, _, _)) \]

Hot components: \{\textit{xyz}(?y, ?z)\}. By inference rule 5 (executing query tasks) and choosing \textit{xyz}(?y, ?z):

\[ P, \{ xyz(2, 7) \} \leadsto (s(3, 2, 7) \otimes r(2)) \land \text{atleast}_1(s(_, _, _)) \]

Hot components: \{s(3, 2, 7)\}. By inference rule 6 for executing the primitive task \textit{s}(3, 2, 7) and by the earlier assumption, it adds the fact \textit{abc}(1) to the database:

\[ P, \{ xyz(2, 7), abc(1) \} \leadsto r(2) \]

Hot components: \{r(2)\}. By inference rule 4 (composite task definitions), where we use the second rule for r:

\[ P, \{ xyz(2, 7), abc(1) \} \leadsto \text{state} \]

Hot components: \{\text{state}\}. By rule 5 applied to \text{state} we derive:

\[ P, \{ xyz(2, 7), abc(1) \} \leadsto () \]

During the deduction, we executed the following sequence of primitive tasks:

\{p(1), xyz(3, 7), s(1, 3, 7), q, xyz(2, 7), s(3, 2, 7), state\}

It is easy to see that this sequence indeed satisfies the constraint \textit{C}, which required \textit{s}(\_, \_, \_, \_\_\_\_) to be executed at least twice with the same third argument.

**Decidability and Complexity.** In general, query answering in CTR is semi-decidable [10], like in classical logic, so no effective procedure exists to answer all possible queries and terminate. Appendices A and B show that the same is true for \textit{ServLog}. In [7], various subsets of CTR were investigated for their decidability and complexity properties. One important restriction studied there is called \textit{full boundedness}. In terms of \textit{ServLog}, this roughly means that every update task must diminish some bounded from below, discrete measure (e.g., a positive integer measure). In that case, the CTR proof procedure is data-complete for NP [7] (i.e., NP-complete when the rules are fixed but data is allowed to vary). Fortunately, all existing workflow modeling systems are implicitly based on the assumption that any useful workflow must be fully bounded (where the bounds can be defined for different processes). Typically this is manifested by imposing upper bounds on the number of iterations, the number of steps, and other similar restrictions. It is also easy to instrument \textit{ServLog} service processes so that they become fully bounded. For instance, primitive update tasks could be forced to diminish a global discrete and bounded resource.

**6. Related Work**

\textit{ServLog} builds on the service contracting framework that was partially developed in [29,28], greatly expanding and generalizing it, while at the same time simplifying the technical details. The major simplifications and extensions include:
– Removing the unique task occurrence restriction and obviating the need for complex simplification transformations.
– Tasks are no longer limited to propositional constants. This provides the ability to represent complex data flow among tasks as well as general transitional constraints.
– General iterative and even mutually recursive tasks in the specification of service processes.
– Generalization of the proof theory, which now deals with the many new additions to the language, and handles constrained (non-Horn) formulas, which previous CTR proof theory was not able to handle.
– Proofs of soundness and completeness for the generalized proof theory.

The present paper significantly extends this proof theory to formulas that contain the $\land$ connective thus enabling execution of constrained transactions, which are non-Horn. We also deal with a much larger class of constraints than [13,28], including iterative processes.

Declare [35] is a service flow language that is closely related to ServLog. It uses Linear Temporal Logic to formalize service flows and automata theory to enact service specifications. The relations between tasks are described entirely in terms of constraints. Apart from the obvious radical differences in the formalisms, some other important differences are worth noting. First, the constraint algebra $\text{CONSTR}$ of ServLog is more expressive than the one used in Declare. Second, by combining constraints with service processes (conditional control flow and data flow), ServLog incorporates current practices in workflow modeling. Third, data flow and conditional control flow are easily available in ServLog, while they have not been developed in the context of Declare. Declare was also formalized using Abductive Logic Programming in [25]. While this formalization supports different verification tasks, the focus remains on modeling service flow exclusively in terms of constraints and does not deal with control and data flow.

In [39,38] the authors propose a combination of colored Petri nets, Declare, and DCR graphs as a way of modeling procedural processes with data support. This combination can be seen as either “adding declarative control-flow to CP-nets” or as “adding data-flow to declarative formalisms.” From the modeling perspective, the approach requires the user to be familiar with three formalisms, as opposed to our framework where all aspects are represented within a single logical language. The authors do not provide a precise formalization of the combination of the three languages, so it is unclear how certain elements such as atomic transactions, hierarchical definitions of tasks, and constraints over complex tasks can be expressed in the combination of those three languages. From the analysis point of view, the authors address the problem of simulation: checking whether a transition is enabled in the CP-net model, and subsequently whether it is also allowed according to the declarative constraint. This is done using model checking. Our approach differs conceptually in that we rely on a logical proof theory as opposed to model checking.

In the same spirit of extending Declare with data elements, [24] extends the Declare notation by allowing activities to have associated multiple ports (denoting events associated to the activity lifecycle) and constraints to be attached to ports thereby allowing data-aware conditions. These extensions are then formalized in the Event Calculus (EC). In terms of modeling, the focus remains on modeling service flow exclusively in terms of constraints and this does not address the aforesaid limitations with respect to control and data flow (e.g., it is unclear how elements such as hierarchical definitions of tasks can be achieved with the proposed extensions). For reasoning, the paper defers to generic EC reasoners, which are significantly more complicated than those for CTR.

An emerging area related to our work is that of compliance checking between business processes and business contracts. For example, in [19,20] both processes and contracts are represented in a formal language called FCL. FCL is based on a formalism for the representation of contrary-to-duty obligations, i.e., obligations that arise when other obligations are violated as typically happens with penalties embedded in contracts. Based on this, the authors give a semantic definition for compliance, but no practical algorithms. In contrast, ServLog provides a proof theory for verifying feasibility of service contracting as well as for contract execution.

Several other approaches to service contracting and contract execution are relevant to our work [31,2,3,11], but not directly related. Most of these present logical modeling languages for contracts in various settings. Being based on normative deontic notions of obligation, prohibition, and permission, we believe that these works are complementary to ours and the approaches could be combined.

Other popular tools for process modeling are based on Petri nets, process algebras, and temporal logic. A related area of research is data-centric business and service modeling and verification, for which a representative approach is [37]. Approaches in this area primarily combine databases and model checking techniques for the purpose of
automated verification. The advantage of CTR over these approaches is that it is a unifying formalism that integrates a number of process modeling paradigms ranging from conditional control flows to data flows to hierarchical modeling to constraints, and even to game-theoretic aspects of multiagent processes (see, for example, [14]). Moreover, as shown in [13], CTR modeling sometimes leads to algorithms with better known complexity for special cases than general model checking.

Finally, it is worth mentioning our work in the context of general AI techniques on reasoning about actions. A key differentiation of our approach based on CTR is that, while many works in AI focus on reasoning about actions, CTR also has constructs for defining transactions and executing them. These issue and further comparison are discussed in detail in the original and follow-up papers [8,10,26,5].

7. Conclusions

The main contribution of this paper is ServLog, a unifying logical framework for semantics-aware services. In particular, this work adds a few new building blocks to the theoretical foundations for service contracting. It offers an expressive set of modeling primitives and addresses a wide variety of issues in service contracts. These primitives range from complex process description (including iteration and conditional transitions between tasks) to temporal and data constraints. Despite its expressive power, ServLog still provides a reasoning procedure for automated service contracting.

In the context of established business process languages (e.g., BPMN, WS-BPEL), ServLog not only captures typical procedural constructs found in such languages, but also greatly extends them, enables declarative specification and reasoning, and opens the way for automatic generation of business processes from service contracts.

The approach presented in this paper has been implemented for the propositional case (when tasks do not have parameters) in the GEECoP (Graphical Editor and Engine for Constrained Processes) tool with promising results (further described in [34]).

Future work may include support for modeling of and reasoning about quality of service (QoS) and non-functional properties in service contracts—see, e.g., [12] for an example of a framework for QoS-based Web service contracting, which could be combined with ServLog, and a more complete implementation of the reasoning framework based on recent results on efficient implementation of Transaction Logic [16]. Another challenging issue is generation of service contracts in ServLog from informal, often ambiguous, natural language contracts. It would be also interesting to do a methodological comparison of the expressive power of ServLog with workflow specification languages based on automata, pre/post conditions, and temporal constraints. In this respect, the recent framework introduced in [1] for comparing distinct workflow models by means of views might be relevant. Explicitly treating violations and reasoning about the effects of violations in contracts is also an interesting direction for further research in ServLog. We are also investigating applicability of ServLog and CTR to other areas such as smart contracts in the context of cryptocurrencies. Here, ServLog could prove to be a useful approach for modeling and reasoning about distributed cryptocurrencies.

Acknowledgements. Michael Kifer was partially supported by the NSF grant 0964196.
Appendix

A. Soundness of the Inference System

**Theorem 2** (Soundness of the Inference System). Suppose $P$ is a set of composite task definitions, $s$ a database identifier, $\psi$ a service process, $C$ a conjunction of primitive or serial constraints. We also assume that $\psi$ and $C$ satisfy the service contracts assumption. Then:

$$P, s \vdash \psi \land C \Rightarrow P, s \models \psi \land C$$

**Proof:** It suffices to prove that the axiom and each inference rule are sound. The only nontrivial case here is soundness of the inference rule 6, execution of primitive update tasks. This proof is given in Proposition 1. □

**Proposition 1** (Soundness of the executing primitive tasks inference rule)

Let $\sigma \rho$ be a primitive update task such that $s_1 \xrightarrow{\rho} s_2$ holds, $\rho \in \text{hot}(\psi)$, and $C$ has no constraint of either of the forms below. We assume that $\bar{p}$ is a task pattern such that $\mapsto_{\bar{p}} \sigma \rho$ matches $\bar{p}$ and that $\bar{r}$ denotes an arbitrary task pattern:

- absence($\bar{p}$)
- suspend($\bar{p}$)
- force($\bar{r}$) such that $\rho \sigma \rho$ does not match $\bar{r}$
- before($\bar{r} \leftarrow \bar{p} \setminus \{\bar{p}_1, ..., \bar{p}_k\}$) $k \geq 0$ and $\rho \sigma \rho$ matches neither $\bar{r}$ nor any of the $\bar{p}_i$’s
- right_before($\bar{r} \leftarrow \bar{p}$)
- next_right_before($\bar{p}' \leftarrow \bar{r} \leftarrow \bar{p}$), where $\rho \sigma \rho$ does not match $\bar{p}'$.

Then the inference rule has the following form:

$$\frac{P, s_2 \vdash (\exists) \psi' \sigma \land C'}{P, s_1 \vdash (\exists) \psi \land C}$$

where $\psi'$ is obtained from $\psi$ by deleting the hot component $p$ and $C'$ is constructed out of $C$ as follows.

Initially, $C'$ is empty. Then constraints are added according to the cases below. As before, in all these cases we assume that $\rho \sigma \rho$ matches $\bar{p}$ and $\bar{r}$, $\bar{s}$ represent arbitrary task patterns.

(a) If $\text{atleast}_n(\bar{p}) \in C$, where $n > 1$, add the following to $C'$:

- $\text{atleast}_{n-1}(\bar{p})$;

(b) If $\text{exactly}_n(\bar{p}) \in C$, where $n > 1$, add the following to $C'$:

- $\text{exactly}_{n-1}(\bar{p})$;

(c) If $\text{exactly}_1(\bar{p}) \in C$, add the following to $C'$:

- absence($\bar{p}$);

(d) If $\text{after}(\bar{p} \rightarrow \bar{r}) \in C$, add the following to $C'$:

- $\text{after}(\bar{p} \rightarrow \bar{r})$
- $\text{atleast}_1(\bar{r}')$, where $\bar{r}' = \text{refine}(\bar{r} gfloor p \sigma, \bar{p})$

(e) If $\text{blocks}(\bar{p} \Rightarrow \bar{r}) \in C$, add the following to $C'$:

- $\text{blocks}(\bar{p} \rightarrow \bar{r})$
- absence($\bar{r}'$), where $\bar{r}' = \text{refine}(\bar{r} gfloor p \sigma, \bar{p})$

(f) If $\text{right_after}(\bar{p} \rightarrow \bar{r}) \in C$, add the following to $C'$:
Proof of (g) If \( \text{force}(\bar{r}') \), where \( \bar{r}' = \text{refine}(\bar{r}; p_{\sigma}, \bar{p}) \).

(h) If \( \text{not_right_after}(\bar{p} \rightarrow \bar{r}) \in C \), add the following to \( C' \):
- \( \text{not_right_after}(\bar{p} \rightarrow \bar{r}) \)
- \( \text{suspend}(\bar{r}') \), where \( \bar{r}' = \text{refine}(\bar{r}; p_{\sigma}, \bar{p}) \)

(i) If \( \text{right_before}(\bar{p} \leftrightarrow \bar{r}) \in C \), add the following to \( C' \):
- \( \text{next_right_before}(\bar{r}' \leftrightarrow \bar{p} \leftrightarrow \bar{r}) \), where \( \bar{r}' = \text{refine}(\bar{r}; p_{\sigma}, \bar{p}) \)

(j) If \( \text{next_right_before}(\bar{r}' \leftrightarrow \bar{s} \leftrightarrow \bar{r}) \in C \), add the following to \( C' \):
- \( \text{right_before}(\bar{s} \leftrightarrow \bar{r}) \), where \( \bar{s}, \bar{r} \) are arbitrary task patterns.

(k) If \( \text{between}(\bar{p} \rightarrow \bar{r} \leftrightarrow \bar{s}) \in C \), add the following to \( C' \):
- \( \text{between}(\bar{p} \rightarrow \bar{r} \leftrightarrow \bar{s}) \)
- \( \text{before}(\bar{r}' \leftrightarrow \bar{s}') \), where \( \bar{r}' = \text{refine}(\bar{r}; p_{\sigma}, \bar{p}) \) and \( \bar{s}' = \text{refine}(\bar{s}; p_{\sigma}, \bar{p}) \)

(l) If \( \text{not_between}(\bar{p} \rightarrow \bar{r} \leftrightarrow \bar{s}) \in C \), add the following to \( C' \):
- \( \text{not_between}(\bar{p} \rightarrow \bar{r} \leftrightarrow \bar{s}) \)
- \( \text{blocks}(\bar{r}' \rightarrow \bar{s}') \) where \( \bar{r}' = \text{refine}(\bar{r}; p_{\sigma}, \bar{p}) \) and \( \bar{s}' = \text{refine}(\bar{s}; p_{\sigma}, \bar{p}) \)

(m) For all other constraints in \( C \), copy them over to \( C' \), but leave out the constraints of the form:
- \( \text{atleast}_1(\bar{p}) \)
- \( \text{force}(\bar{p}) \)
- \( \text{suspend}(\bar{r}) \), for any task pattern \( \bar{r} \)

Proof: Suppose \( \mathbf{P}, s_1 ... s_n \models (\exists) \psi' \sigma \land C' \). By soundness of the CTR proof theory, it follows that \( \mathbf{P}, s_1 ... s_n \models (\exists) \psi \land C \) and, as a special case, that:

\[
\mathbf{P}, (s_1 ... s_n) \models C
\]

In the proof, we will rely on the premise

\[
\mathbf{P}, (s_2 ... s_n) \models C' \tag{11}
\]

which is a special case of the assumption \( \mathbf{P}, s_2 ... s_n \models (\exists) \psi' \sigma \land C' \).

Proof of (10) To prove (10), we need to consider each case in the inference rule 6 and, for each constraint that is not copied directly from \( C \) to \( C' \) via Case (m), we need to show that it holds over the execution \( (s_1 ... s_n) \).

Case (a): In this case, \( C \) has \( \text{atleast}_n(\bar{p}) \) and \( C' \) has \( \text{atleast}_{n-1}(\bar{p}) \). By (11), it follows that \( \mathbf{P}, s_2 ... s_n \models \text{atleast}_{n-1}(\bar{p}) \).

If \( \mathbf{P}, s_2 ... s_n \models \text{atleast}_{n-1}(\bar{p}) \) then since \( s_1 : \xrightarrow{p_{\sigma}} s_2 \) and \( p_{\sigma} \) matches \( \bar{p} \), it follows that \( \mathbf{P}, s_1 ... s_n \models \text{atleast}_n(\bar{p}) \).

Case (b): similar to Case (a).

Case (c): In this instance, \( C' \) contains \( \text{absence}(\bar{p}) \), so \( s_i \xrightarrow{r} s_{i+1}, 1 < i < n \), is not possible for any \( r \) that matches \( \bar{p} \). Therefore, \( (s_1 ... s_n) \) satisfies \( \text{exactly}_1(\bar{p}) \), i.e., \( \mathbf{P}, s_1 ... s_n \models \text{exactly}_1(p) \) holds.

Case (d): In this case, \( C' \) contains \( \text{after}(\bar{p} \rightarrow \bar{r}) \) and \( \text{atleast}_1(\bar{r}') \), and \( C \) contains \( \text{after}(\bar{p} \rightarrow \bar{r}) \). By (11), it follows that \( \mathbf{P}, s_2 ... s_n \models \text{atleast}_1(\bar{r}') \). Suppose \( \text{after}(\bar{p} \rightarrow \bar{r}) \) does not hold on \( (s_1, s_2, ..., s_n) \). Since the
transition from $s_1$ to $s_2$ can be made only by $p\sigma$ (by the primitive update task independence assumption), it follows that $\text{absence}(\vec{r}')$ must hold on $(s_2, \ldots, s_n)$. But this contradicts the fact that $\text{atleast}_1(\vec{r}')$ must hold on the same path. Thus, $\text{after}(\vec{p} \rightarrow \vec{r})$ must hold on $(s_1, \ldots, s_n)$, i.e., $P, s_1 \ldots s_n \models \text{after}(\vec{p} \rightarrow \vec{r})$ holds.

**Case (e):** Here $C'$ contains blocks($\vec{p} \rightarrow \vec{r}$) and absence($\vec{r}'$), and $C$ contains blocks($\vec{p} \rightarrow \vec{r}$). By (11), $P, s_2 \ldots s_n \models \text{absence}(\vec{r}')$. Suppose blocks($\vec{p} \rightarrow \vec{r}$) does not hold on $(s_1, s_2, \ldots, s_n)$. Since the transition from $s_1$ to $s_2$ can be made only by $p\sigma$ (by the primitive update task independence assumption), it follows that $\text{atleast}_1(\vec{r}')$ must hold on $(s_2, \ldots, s_n)$. But this contradicts $P, s_2 \ldots s_n \models \text{absence}(\vec{r}')$. Thus $P, s_1 \ldots s_n \models \text{blocks}(\vec{p} \rightarrow \vec{r})$ must hold.

**Case (f):** Here $C'$ contains both $\text{right_after}(\vec{p} \rightarrow \vec{r})$ and $\text{force}(\vec{r}')$, while $C$ contains $\text{right_after}(\vec{p} \rightarrow \vec{r})$. By (11), it follows that $P, s_2 \ldots s_n \models \text{force}(\vec{r}')$, and thus $s_2 \rightarrow s_3$ must hold for some $r'$ that matches $\vec{r}'$, a refinement pattern of $\vec{r}$. Since $s_1 \xrightarrow{p\sigma} s_2$ and $p\sigma$ matches $\overline{p}$, it follows that $P, s_1 \ldots s_n \models \text{right_after}(\vec{p} \rightarrow \vec{r})$ holds.

**Case (g):** In this instance, $C'$ contains $\text{before}(\vec{p} \leftrightarrow \vec{r} \setminus \{\vec{r'}, \vec{r_1}, \ldots, \vec{r}_k\})$ while $C$ contains $\text{before}(\vec{p} \leftrightarrow \vec{r} \setminus \{\vec{r_1}, \ldots, \vec{r}_k\})$. By (11), it follows that $P, s_2 \ldots s_n \models \text{before}(\vec{p} \leftrightarrow \vec{r} \setminus \{\vec{r'}, \vec{r_1}, \ldots, \vec{r}_k\})$. Since $s_1 \xrightarrow{p\sigma} s_2$ holds, $p\sigma$ matches $\overline{p}$, and $\vec{r}' = \text{refine}(\vec{r}; p\sigma, \overline{p})$, it follows that before($p\sigma \rightarrow \vec{r}'$) (and, as a special case, before($\vec{p} \rightarrow \vec{r}'$)) is satisfied by $(s_1, s_2, \ldots, s_n)$. Since $(s_2, \ldots, s_n)$ satisfies before($\vec{p} \rightarrow \vec{r} \setminus \{\vec{r'}, \vec{r_1}, \ldots, \vec{r}_k\}$), we conclude that it also satisfies before($\vec{p} \leftrightarrow \vec{r} \setminus \{\vec{r_1}, \ldots, \vec{r}_k\}$). But the transition from $s_1$ to $s_2$ was made by $p\sigma$, which matches $\overline{p}$, so $(s_1, s_2, \ldots, s_n)$ cannot violate before($\vec{p} \leftrightarrow \vec{r} \setminus \{\vec{r_1}, \ldots, \vec{r}_k\}$).

**Case (h):** In this case, $C'$ contains $\text{not_right_after}(\vec{p} \rightarrow \vec{r})$ and $\text{suspend}(\vec{r}')$, and $C$ contains $\text{not_right_after}(\vec{p} \rightarrow \vec{r})$. By (11), it follows that $P, s_2 \ldots s_n \models \text{suspend}(\vec{r}') \land \text{not_right_after}(\vec{p} \rightarrow \vec{r})$. Suppose $\text{not_right_after}(\vec{p} \rightarrow \vec{r})$ does not hold on $(s_1, s_2, \ldots, s_n)$. Since the transition from $s_1$ to $s_2$ can be made only by $p\sigma$ (by the primitive update task independence assumption), it follows that the only way to violate $\text{not_right_after}(\vec{p} \rightarrow \vec{r})$ over $(s_1, s_2, \ldots, s_n)$ is to cause a transition $s_2 \rightarrow s_3$ using some $r'$ that matches $\vec{r}' = \text{refine}(\vec{r}; p\sigma, \overline{p})$. But this contradicts the assumption that $P, s_2 \ldots s_n \models \text{suspend}(\vec{r}')$ holds. Therefore $P, s_1 \ldots s_n \models \text{not_right_after}(\vec{p} \rightarrow \vec{r})$ holds.

**Case (i):** In this case, $\text{next_right_before}(\vec{r}' \leftarrow \vec{p} \leftarrow \vec{r}) \in C'$, where $\vec{r}' = \text{refine}(\vec{r}; p\sigma, \overline{p})$, and $\text{right_before}(\vec{p} \leftarrow \vec{r}) \in C$. Since the execution path $(s_2, \ldots, s_n)$ satisfies the constraint $\text{next_right_before}(\vec{r}' \leftarrow \vec{p} \leftarrow \vec{r})$, we have that either (i) $s_2 \rightarrow s_3$ holds and $(s_3, \ldots, s_n)$ satisfies $\text{right_before}(\vec{p} \leftarrow \vec{r})$; or (ii) $(s_2, \ldots, s_n)$ satisfies $\text{right_before}(\vec{p} \leftarrow \vec{r})$. In either case, since $s_1 \xrightarrow{p\sigma} s_2$ holds by the assumption, it follows that $\text{right_before}(\vec{p} \leftarrow \vec{r})$ holds on the entire execution path $(s_1, \ldots, s_n)$.

**Case (j):** Here $\text{right_before}(\vec{s} \leftarrow \vec{r}) \in C'$ and $\text{next_right_before}(\vec{r}' \leftarrow \vec{s} \leftarrow \vec{r}) \in C$, where $\vec{s}, \vec{r}$ are arbitrary task patterns. We thus have that $P, (s_2 \ldots s_n) \models \text{right_before}(\vec{s} \leftarrow \vec{r})$ and $s_1 \xrightarrow{p\sigma} s_2$ both hold. By the premise of rule 6, either

- $p\sigma$ does not match $\vec{r}$; or
- $p\sigma$ matches $\vec{r}'$.

Since $\text{right_before}(\vec{s} \leftarrow \vec{r})$ is satisfied by $(s_2, \ldots, s_n)$, in either of these cases the constraint $\text{next_right_before}(\vec{r}' \leftarrow \vec{s} \leftarrow \vec{r}) \in C$ holds on $(s_1, \ldots, s_n)$.

**Case (k):** In this case, $C'$ contains between($\vec{p} \rightarrow \vec{r} \leftarrow \vec{s}$) and $\text{before}(\vec{r}' \leftarrow \vec{s}')$, where $\vec{r}' = \text{refine}(\vec{r}; p\sigma, \overline{p})$ and $\vec{s}' = \text{refine}(\vec{s}; p\sigma, \overline{p})$, while $C$ contains between($\vec{p} \rightarrow \vec{r} \leftarrow \vec{s}$) instead. In addition, $s_1 \xrightarrow{p\sigma} s_2$ holds. Since $(s_2, \ldots, s_n)$ satisfies $\text{between}(\vec{r}' \leftarrow \vec{s}')$ and between($\vec{p} \rightarrow \vec{r} \leftarrow \vec{s}$), it follows that $(s_1, s_2, \ldots, s_n)$ satisfies between($p\sigma \rightarrow \vec{r}' \leftarrow \vec{s}'$). Therefore, between($\vec{p} \rightarrow \vec{r} \leftarrow \vec{s}$) cannot be violated by $(s_1, s_2, \ldots, s_n)$.

**Case (l):** Here $C'$ contains $\text{not_between}(\vec{p} \rightarrow \vec{r} \leftrightarrow \vec{s})$ and $\text{blocks}(\vec{r}' \rightarrow \vec{s}')$ while $C$ contains $\text{not_between}(\vec{p} \rightarrow \vec{r} \leftrightarrow \vec{s})$. In addition, $s_1 \xrightarrow{p\sigma} s_2$ holds. Since $(s_2, \ldots, s_n)$ satisfies $\text{blocks}(\vec{r}' \rightarrow \vec{s}')$, and $\text{not_between}(\vec{p} \rightarrow \vec{r} \leftrightarrow \vec{s})$, it follows that $(s_1, s_2, \ldots, s_n)$ satisfies $\text{not_between}(p\sigma \rightarrow \vec{r}' \leftrightarrow \vec{s}')$. Therefore, $\text{not_between}(\vec{p} \rightarrow \vec{r} \leftrightarrow \vec{s})$ cannot be violated by $(s_1, s_2, \ldots, s_n)$.

**Case (m):** In this catch-all case, all constraints are copied over from $C$ to $C'$ except the following:
Theorem 3 (Completeness of the Inference System). Suppose \( \mathbf{P} \) is a set of composite task definitions, \( s \) a database identifier, \( \psi \) a service process, \( C \) a conjunction of primitive or serial constraints, and \( \psi \) and \( C \) satisfy the service contracts assumption. Then:

\[
\mathbf{P}, s \ldots \vdash (\psi \land C) \quad \text{implies} \quad \mathbf{P}, s \ldots \vdash (\psi \land C)
\]

Proof: The proof of completeness is based on the following facts:

- The original CTR proof theory is complete.
- The current proof theory restricts the original CTR proof theory by eliminating certain derivation paths. Soundness implies that all the remaining paths satisfy \( C \). Completeness means that none of the eliminated path satisfies \( C \).

The key step in the proof is to show that the eliminated derivation paths correspond to executions that violate some of the constraints.

First, by the inference rule 1, it is always possible to ensure that \( C \) in rule 6 and in the axiom is a conjunctive constraint.

One reason why a derivation path may be eliminated by the proof theory is that in the inference rule 6 (executing primitive update tasks) a hot task \( p \sigma \) is blocked because \( C \) contains:

- \( \text{absence}(\bar{p}) \)
- \( \text{force}(\bar{r}) \) such that \( p \sigma \) does not match \( \bar{r} \)
- \( \text{before}(\bar{r} \leftarrow s \leftarrow \{\bar{p}_1, ..., \bar{p}_k\}) \) \( k \geq 0 \) and \( p \sigma \) does not match any of the \( \bar{p}_i \)’s
- \( \text{right_before}(\bar{r} \leftarrow \bar{r} \leftarrow \bar{q}) \)
- \( \text{next_right_before}(\bar{r} \leftarrow \bar{r} \leftarrow \bar{q}) \), where \( p \sigma \) does not match \( \bar{p} \).

where \( \bar{r} \) is an arbitrary task pattern and \( p \sigma \) matches \( \bar{p} \).

If such a hot task \( p \sigma \) is executed at some state, \( s_1 \), and causes a transition to state \( s_2 \) then either \( \text{absence}(\bar{p}) \), or \( \text{before}(\bar{r} \leftarrow \bar{r} \leftarrow \{\bar{p}_1, ..., \bar{p}_k\}), \text{right_before}(\bar{r} \leftarrow \bar{r}), \text{next_right_before}(\bar{r} \leftarrow \bar{r} \leftarrow \bar{q}) \), or \( \text{force}(\bar{r}) \) are not satisfied on any path of the form \( (s_1 s_2 ... s_n) \). Therefore, such an eliminated path is not a valid execution.

The other case when our proof theory may eliminate an execution path that would have been otherwise found by the plain CTR proof procedure is when we derive \( \mathbf{P}, s \ldots \vdash (\psi \land C) \) but cannot declare success because the axiom does not apply due to the fact that \( C \) contains a constraint of the form \( \text{force}(\bar{r}), \text{atleast}_k(\bar{r}), \text{or exactly}_k(\bar{r}) \), where \( k \geq 1 \). But then, if the eliminated execution path satisfied the original constraint \( C \) then, by soundness of the inference rules, the path \( (s) \) would have to satisfy \( C \) and thus also one of the aforesaid constraints: \( \text{force}, \text{atleast}, \text{or exactly} \), which is not the case.

Since the remaining inference rules do not restrict execution of transactions and are essentially identical to the corresponding rules in CTR (except that the constraints are tacked on to them), the result follows. \( \square \)
C. Representing Constraints of ServLog as CTR Formulas

This appendix provides direct definitions of the constraints introduced CONSTR (Section 4.2) and the additional ones from Section 5.

The followings are CTR representations for the constraints in CONSTR that are equivalent to the semantic definitions for these constraints given in Sections 4.2 and 5. In the formulas below we use $\text{vars}(\overline{t})$ to represent the set of variables appearing in $\overline{t}$,\(^\text{17}\) and $\nabla p(\overline{t})$ denotes\(^\text{18}\)

$$\text{path} \otimes p(\overline{t}) \otimes \text{path}$$

(12)

The constraints in CONSTR are now represented in CTR as follows:

- **atleast\(_n\)(p(\overline{t}))**: \(\exists \text{vars}(\bigcup_{i=1}^{n} \overline{t}_i) \ (\nabla p(\overline{t}) \otimes \ldots \otimes \nabla p(\overline{t}_n))\) \hspace{1cm} (13)

Where \(\overline{t}_i\) is \(\overline{t}\) in which all underscores (unnamed placeholders) are replaced with new variables. (Recall that named placeholders are regular variables.)

- **absence(p(\overline{t}))**: \(\forall \text{vars}(\overline{t}) \neg \nabla p(\overline{t})\) \hspace{1cm} (14)

- **exactly\(_n\)(p(\overline{t}))**: \(\text{atleast}_n(p(\overline{t})) \land \neg \text{atleast}_{n+1}(p(\overline{t}))\) \hspace{1cm} (15)

- **after(p(\overline{t}) \rightarrow q(\overline{\pi}))**: \(\forall \text{vars}(\overline{t}) \exists \text{vars}(\overline{\pi}) \land \text{vars}(\overline{t}) \land \text{path} \otimes p(\overline{t}) \Rightarrow \nabla q(\overline{\pi})\) \hspace{1cm} (16)

Here \(\text{vars}(\overline{\pi}) \land \text{vars}(\overline{t})\) is the set of variables in \(\overline{\pi}\) minus those that appear in \(\overline{t}\).

- **before(p(\overline{t}) \leftrightarrow q(\overline{\pi}))**: \(\forall \text{vars}(\overline{\pi}) \exists \text{vars}(\overline{\pi}) \land \text{vars}(\overline{\pi}) \land \nabla p(\overline{t}) \equiv q(\overline{\pi}) \otimes \text{path}\) \hspace{1cm} (17)

- **before(p(\overline{t}) \leftrightarrow q(\overline{\pi}) \land \{q(\overline{\pi}_1), \ldots, q(\overline{\pi}_n)\})**:

\(\forall \text{vars}(\overline{\pi}) \land \forall \text{vars}(\bigcup_{i=1}^{n} \overline{\pi}_i) \land \exists \text{vars}(\overline{t}) \land \text{vars}(\overline{\pi}) \land \bigvee_{i=1}^{n} (\overline{\pi} = \overline{\pi}_i) \land (\nabla p(\overline{t}) \Leftarrow q(\overline{\pi}) \otimes \text{path})\) \hspace{1cm} (18)

- **blocks(p(\overline{t}) \rightarrow q(\overline{\pi}))**: \(\forall \text{vars}(\overline{t}, \overline{\pi}) \land \text{path} \otimes p(\overline{t}) \Rightarrow \neg \nabla q(\overline{\pi})\) \hspace{1cm} (19)

\(^\text{17}\)Note that all underscores (unnamed placeholders) are replaced with new variables, e.g. \(\text{vars}(p(\ldots)) = \{?X, ?Y\}\).

\(^\text{18}\)Informally, \(\nabla p(\overline{t})\) means that \(p(\overline{t})\) eventually executes.
- \( \text{between}(p(\overline{t}) \rightarrow q(\overline{v}) \leftarrow r(\overline{v})) \):

\[
\forall \text{vars}(\overline{t}, \overline{v}) \exists \text{vars}(\overline{v}) \setminus \text{vars}(\overline{t}, \overline{v}) \\
\text{path} \odot p(\overline{t}) \Rightarrow \lor q(\overline{v}) \leftarrow r(\overline{v}) \odot \text{path}
\]  

(20)

- \( \text{not\_between}(p(\overline{t}) \rightarrow q(\overline{v}) \leftarrow r(\overline{v})) \):

\[
\forall \text{vars}(\overline{t}, \overline{v}, \overline{w}) \\
\text{path} \odot p(\overline{t}) \Rightarrow \neg \lor q(\overline{v}) \leftarrow r(\overline{v}) \odot \text{path}
\]  

(21)

- \( \text{right\_after}(p(\overline{t}) \rightarrow q(\overline{v})) \):

\[
\forall \text{vars}(\overline{t}) \exists \text{vars}(\overline{v}) \setminus \text{vars}(\overline{t}) \\
\text{path} \odot p(\overline{t}) \Rightarrow q(\overline{v}) \odot \text{path}
\]  

(22)

- \( \text{right\_before}(p(\overline{t}) \leftarrow q(\overline{v})) \):

\[
\forall \text{vars}(\overline{v}) \exists \text{vars}(\overline{t}) \setminus \text{vars}(\overline{v}) \\
\text{path} \odot p(\overline{t}) \Leftarrow q(\overline{v}) \odot \text{path}
\]  

(23)

- \( \text{next\_right\_before}(q(\overline{v}) \Leftrightarrow p(\overline{t}) \leftarrow q(\overline{v})) \):

\[
\text{right\_before}(p(\overline{t}) \leftarrow q(\overline{v})) \vee \\
(\exists \text{vars}(\overline{v}) q(\overline{v}) \odot \text{right\_before}(p(\overline{t}) \leftarrow q(\overline{v}))
\]  

(24)

- \( \text{not\_right\_after}(p(\overline{t}) \rightarrow q(\overline{v})) \):

\[
\forall \text{vars}(\overline{t}, \overline{v}) \\
\text{path} \odot p(\overline{t}) \Rightarrow (\lor \neg q(\overline{v}) \Rightarrow q(\overline{v}) \odot \text{path}
\]  

(25)

where \( \lor \) is a CTR proposition such that \( s_1 \rightarrow^\lor s_2 \) is true for any pair of states \( s_1, s_2 \) and \( \lor \) is not true on any other path.

- \( \text{force}(p(\overline{t})) \):

\[
\exists \text{vars}(\overline{t}) \ p(\overline{t}) \odot \text{path}
\]  

(26)

- \( \text{suspend}(p(\overline{t})) \):

\[
\forall \text{vars}(\overline{t}) \ (\lor \neg p(\overline{t}) \odot \text{path}
\]  

(27)

- **Composite constraints**: If \( C_1, C_2 \in \text{CONSTR} \) then so are \( C_1 \land C_2 \) (a conjunctive constraint) and \( C_1 \lor C_2 \) (a disjunctive constraint).

References


