Checking Realizability of a Timed Business Processes Choreography

Manuel I. Capel

Software Engineering Department, Granada University,
Periodista Daniel Saucedo Aranda, 18071 Granada, Spain

Keywords: Business Process Management, Business Process as a Service, BPMN, Process Modeling, Temporal Constraints and Dependencies, Temporal Semantics.

Abstract: A business process (BP) can be understood as a set of related, structured, interacting services acting as peers, according to an intended choreography that is capable of giving complex functionality to customers. Several authors have made progress in solving the "choreography realization" problem. The research work carried out in this paper amounts to analyzing and automatically checking the realizability of the defined choreography for services that communicate through messages in a general, distributed, and highly parallel system.

1 INTRODUCTION

Business Process Model and Notation (BPMN) was extended into BPMN 2.0 to provide an interaction model for the business processes (BP) which is based on specification of choreographies as opposed to "interconnected interface models" (Poizat P., 2012) that promote orchestration-oriented specifications, which is a bottom-up and a local approach of service composition.

Realizability of choreographies consists of formally determine if the individual peers obtained from a choreography, independently from the kind of communication or their relative execution order, are capable of interacting as prescribed in the choreography requirement specification. We can say that a choreography is realizable if all the interactions specified in BPMN 2.0 choreography-diagrams are equivalent to those that can be executed by the interacting peers when the BP model is implemented in a Web-services description language. In this way, a formalization of behavioral and temporal aspects of BPMN 2.0 models through a provably-realizable choreography will provide important benefits at analysis and design of complex applications built over distributed and highly parallel platforms.

There are already contributions to the resolution of the choreography realization problem (Rozinat A., 2006), (Dongen B., 2004). The research work up until now can be divided into two categories. The first approach aims at transforming BPMN models into executable environments (Capel M.I., 2014) and performing formal analysis (Aalst, 2009). The second one includes formal methods for BPMN models verification, which are based on Π-calculus (Milner, Hall) or Petri Nets (Cerone, 2002), which can debug grammatical errors and can transform business processes diagrams (BPD) into BP Execution Language (BPEL) code (Arkin A., 2005; OASIS, 2007), or representing BPM system properties with CCTL. In this second group, the central problem to tackle consists of proving the soundness of BPMN model transformation.

BPMN transformation is more than just individually converting the model’s entities. In many cases, a model cannot be verified because its representation in an executable environment does not react to external events in the same way.

We propose here a solution to that problem based on the transformation of the BPMN and the peer-system models into a formal specification, realized in the timed process calculus named CSP+T, thereby allowing their complete verification with state-of-the-art model checking tools.

In Section 2 we introduce the use of BPMN 2.0 for choreography specification. Section 3 presents our formal model transformation from BPMN 2.0 choreographies into CSP+T process calculus. Section 4 shows how this encoding can be used to reify choreographies with LTS, and check their realizability. Model checking of choreographies is addressed in Section 5, and tool support is discussed in section 6. Finally, the conclusions and future leads of our work.
2 BPMN CHOREOGRAPHIES FORMALIZATION

The graphical standardized notation Business Process Model and Notation (BPMN) (OMG, 2011) provides representation elements for business processes with an emphasis on control–flow. BPMN introduces a specific type of flowchart, named Business Process Diagram (BPD), which provides graphical constructs tailored to model BPMN elements, such as: (a) gateways, (b) events start, stop, (c) tasks, and (d) control flows. Non-technical personnel, usually management-oriented, can easily understand a BPD diagram and use it for modeling BPs. The apparent simplicity of BPD, however, offers the necessary expressiveness power to model very complex BPs and it can be mapped to different business execution languages such as BPEL (Arkin A., 2005) or XLANG (Thatte, 2001). Figure 1 shows BPMN elements as they are represented in BPD.

Semi-formal specification of task workflows can be done with BPMN notation.

BPMN 2.0 supports the collaboration between analysis entities in BP models, which brings forward a choreographic model based on peer interactions, instead of following a design model based on services orchestration. Thus, it promotes a collaborative and abstract description of software systems that allows for focusing more on what services do in a composition than on how they do it. Interactions between system’s components or peers should be more precisely described now than in “interconnected interface” models, within which the interactions are defined internally to each peer only. Contrary to interface description–based, “interaction–based” models consider the description of “conversations” between peers as the basic building blocks of any BP system design, whereas the specification of interfaces then becomes secondary to the system’s properties analysis.

Case Study. We will use a Pharmacy Hospital logistic process as a simple running example. The UML sequence-diagram in Figure 2 depicts the message flows between two participants, the Ward and the Pharmacy, which are independent BP and may have been constructed separately. Clearly, the synchronization between both participants is a necessary behavioural property for successful collaboration. The Ward-Pharmacy interaction choreography is realizable if that property is individually proved for each participant.

2.1 Coreography Description Language

BPMN 2.0 has introduced the Choreography Description Language (CDL) for specifying the requirements of a business process choreography. CDL is capable of defining, from an integrated point of view, the common and complementary observable behavior of different services that exchange messages according to some ordered protocol for accomplishing a common business goal (W3C, 2005).

Peers interactions are the basic building blocks of any CDL specification. Peers A and B are represented by the upper and lower bands, respectively, in the round boxes of Figure 3. One–way and two–way interactions between peers are easily described by BPMN 2.0 Choreography Diagrams (BCD).
In a choreography, two-way interactions of peers (tasks) are represented by message exchanging. A is called the initiating peer and is represented by a white band as opposite to the dark filled one for the the receiving peer B.

![Figure 3: BPMN 2.0 choreography diagram tasks.](image)

There are one-way interactions (Figure 3, a,c,d) and two-way interactions (Figure 3, b). Two-way interactions are represented by an initiating message (white envelope) and a return message (black envelope). Tasks can also include internal markers, such as the standard loop (Figure 3, b), in which the interaction is performed several times depending on a boolean condition. In Multi-instance parallel loops (Figure 3, c,d), the interactions are performed by several instances of the choreography task, which can execute the actions in parallel (||) or sequentially (≡). If the message exchange need not be repeated, no marker is used to describe a task.

Business decisions and flow branching are modelled using gateways, which are similar to decision symbols in a flowchart. Gateways describe how sequence flows join (multiple outgoing sequence flows and at most one incoming sequence flow), and fork (multiple incoming sequence flows and at most one outgoing sequence flow). In our description language we take into account (Figure 4) the following symbols for gateways: XOR exclusive gateways (decision, alternative paths), inclusive gateways (inclusive decision, but also parallel paths), parallel gateways (creating and merging parallel flows) and event-based gateways (choices based on events). Diagrams needing both converging and diverging choices can be described by a sequence of single-converging/diverging gateways.

![Figure 4: BPMN Elements Extended Graphical Representation.](image)

**BCD of the Case Study.** From the pharmacy-participant perspective, a deliver drug message is sent to the ward-participant for guaranteeing that a prior demanded drug is now available; and an internal order received prior a purchase order for a specific drug is made or sent. The ward-participant has to send the demand drug message with enough time in advance to guarantee one drug availability to begin a medical treatment.

The specification in the BCD (Figure 5) shows the interaction between the peers (customer, ward, pharmacy, db) of the choreography, translated from the business process model of ‘Logistic Process in Hospitals’ (Baacke L., 2009) that represents the behaviour of tasks within the 2 pools (BPMN notation to define task sequences); Ward and Pharmacy of the example. In this specification, we can firstly see that the customer interacts with the ward (connect), then the ward sends the prescription to the pharmacy (request) and eventually receives the drug when it is available (received). If that drug is not in stock, the pharmacy makes an order (purchase order), communicates (purchase order placed) to the ward and when the drug is available, it delivers the drug to the ward. Finally, the prescription will be prepared and given to the customer (delivered), thereby terminating the complete protocol.

**WS-CDL Description.** The Web Services Choreography Description Language (WS-CDL) is an XML-based language for describing Web services interactions (W3C, 2005). In WS-CDL there is no centralized control to coordinate different services; and therefore, there are no global variables, conditions or workunits either. In order to give the illusion of a global or shared state among the choreographed services, the variables located in one service can be aligned (synchronized) via message passing with other variables located in a distinct service. Message
sequences that do not follow the ordering rules are considered out of sequence messages and the language shows an error of conformance with respect to the description of the intended choreography. The lack of formal semantics of WS-CDL is currently considered an issue since it hinders the development of tool support for verifying message conformance to a specific choreography description. In addition to conformance checking, we need to determine whether a given choreography is realizable, i.e., if its individual peers obtained through projection from a choreography interact as prescribed in the specification of choreography requirements.

3 ENCODING BPMN 2.0 INTO CSP+T

CSP+T (Zic, 1994) is a real–time specification language which extends Communicating Sequential Processes (CSP) allowing the description of complex event timings, within a single sequential process, for use in the behavioural specification of any critical communicating process. A CSP+T process term \( \mathcal{P} \) is defined as a tuple \((\alpha P, P)\), where \( \alpha P = \text{Comm}_\text{act}(P) \cup \text{Interface}(P) \) is the communication alphabet of \( P \).

**BPMN 2.0 into CSP.** There have been several proposals to formalize BPMN 2.0 non–temporal constructs with process algebras, (Puhlnmann, 2007; Ma S., 2008; Mendoza L.E., 2012).

In the CCS–based notation proposed in (Wong P.Y.H., 2008), which we will follow in the sequel, each BPMN entity has associated attributes describing its properties; for example the number of loops of a sequence multiple instance is recorded by the natural number in the constructor \( \text{misseq}. \)

The formal semantics of BPMN abstracts (partial function \( \text{hide} \)) the internal flow of the modeling entity named \( \text{state} \) and only describes the sequence of initializations and terminations with the semantic function \( \text{bpmm} \).

The type of a sequence flow or an exception flow is given by the following schema definition:

\[
\text{Transition} \equiv \{ \text{guard} : \text{Guard}; \text{line} : \text{Line} \}
\]

and the type of message flow:

\[
\text{Message}_\text{flow} \equiv \{ \text{message} : \text{Message}; \text{channel} : \text{Channel} \}
\]

If the sequence flow has no guard or the message flow contains an empty message, then the values of Transition and Message\_flow record the default values "\( tt \)" and empty respectively. There are five sets of message flows (send, receive, reply, accept, break) associated to the \( \text{state} \) entity, which are syntactically defined in the next type and whose function follows from their names.

\[
\text{State} = \{ \text{type} : \text{Type}; \text{in}, \text{out}, \text{error} : \mathbb{P} \text{Transition}; \text{exit} :\mathbb{P} \text{Message}_\text{flow}; \text{link} : \mathbb{P} \text{Transition} \times \mathbb{P} \text{Message}_\text{flow}; \text{depend} : \mathbb{P} \text{Message}_\text{flow} \times \mathbb{P} \text{Message}_\text{flow}; \text{loopMax} : \mathbb{N} \}
\]

Each state also incorporates the variable \( \text{loopMax} \) to limit the number of state instances that each process can invoke. The state’s component \( \text{link} \) pairs the incoming message flow which initiates or interrupts the execution of the state with either an incoming transition or an exception flow. The component \( \text{depend} \) pairs each incoming message that initializes the state’s execution with its corresponding outgoing message flow.

\[
\begin{align*}
\text{[bpmm]} \\
\text{bpmm} : \mathbb{P} \text{Name} \rightarrow \text{Local} \rightarrow \text{Process} \\
\text{hide} : \mathbb{P} \text{Name} \rightarrow \text{Local} \rightarrow \text{Event} \\
\text{function\_name} = \{ Y | (\alpha Y)'X \backslash \text{hide} || S \} \\
\text{Where} X = \{ i : \alpha Y : \text{hide} (\text{fin}, \text{abt}) \bullet (i \rightarrow X \cup \text{fin} \rightarrow \text{null} \square \text{abt} \rightarrow \text{stop}) \} \\
\text{And} Y = \{ i | \text{Process\_set} \bullet (\alpha P) \circ P(i) \}
\end{align*}
\]

The partial function \( \text{bpmm} \) maps a syntactic description of a BPMN diagram encapsulated by a \( \text{pool} \) or BPMN–\( \text{subprocess} \) into a parallel composition of CSP+T subprocesses, which correspond to the diagrams of the basic elements of BPMN shown in Figure 1. The set \( \alpha Y \) represents the communication alphabet that includes all the messages types and events that may affect the execution of processes. These communications represent the events that a process \( P \) receives from its \( \text{environment} \) (made up of all the other processes in the system) or those events that occur internally, i.e., which are not externally visible.

3.1 BPMN 2.0 Temporal Extension

In many models of choreographies, constraints on time and resources appear that may cause the violation of the system’s safety properties.

In BPMN 2.0 the intermediate event (\( \text{timer} \)) admits the definition of a delay period, but the minimum and maximum execution time allowed for any activity cannot be specified as a task element.

We propose to extend BPMN 2.0, so that it includes temporal attributes for specifying temporal constraints for \( \text{task} \) and other modelling elements. An intermediate event (\( \text{timer} \)) and a sub–process (\( \text{timeout} \)) represent a temporal constraint on the task flow. Notice that in BPMN 2.0 we can use the attribute of the \( \text{timer} \) element to define the task delay period.
Start Event. The schema labelled Start event represents the necessary instantiation of an activity prior to the start of its execution. To allow this in CSP+T, the specification of this event is represented by means of the ★ instantiation event according to the following pattern,

\[ P(\text{start}) = \star \parallel v_\star \rightarrow \text{SKIP} \uparrow (P(S) \parallel e_{\text{end}} \rightarrow \text{SKIP}) \]

Minimum and Maximum Duration Time of an Activity. The invocation of activities that make up a BP must be performed timely.

Let bpmn S_1 be the activity that comes before activity bpmn S_2 (Figure 6). Thus, it should be guaranteed that the execution time of bpmn S_1 activity does not overrun its maximum range of duration \( S_1.\text{ran.max} \) and it does not occur before its minimum starting time \( S_1.\text{ran.min} \) elapses. If the above times are not controlled, the activity fails and thus we cannot certify the temporal properties of the business process.

\[ \epsilon.S_1 \leq v.S_1 + S_1.\text{ran.min} \leq v.S_1 + S_1.\text{ran.max} \]

With CSP+T we are allowed to precisely specify the time frame for the execution of an activity. The CSP+T pattern corresponding to two consecutive activities \( S_1 \) and \( S_2 \) is as it follows,

\[ P(S_1) = (S_1.\text{ran.min} \parallel v.S_1 \rightarrow \text{SKIP} \uparrow P(S_1) \parallel \text{end} \rightarrow \text{SKIP}) \]

The measured range values \( S_1.\text{ran.min} \) and \( S_2.\text{ran.max} \) depend on event \( \epsilon.S_2 \) occurrence.

Timed Exception Flow. BPMN proposes two versions of type boundary event to represent timeouts. Timer boundary represents the occurrence of an event \( \epsilon_{\text{exc}} \) (see Figure 7) that either triggers a parallel thread or interrupts the execution of activity bpmn S_1 at \( \text{itime}.\text{ran} \) time units after the inception of S_1. Hence, there are two instances of the modelling entity timer boundary even, depending on whether the activity started at \( \epsilon.S_1 \) is interrupted or not when this event occurs. The \( \text{itime}.\text{ran} \) time period must be therefore constrained by the maximum time limit that we have associated to any activity.

\[ (\text{itime}.\text{ran} < S_1.\text{ran.max}) \land (s(\epsilon_{\text{exc}}) = v.S_1 + \text{itime}.\text{ran}) \land (s(\epsilon_{\text{exc}}) \in [v.S_1, S_1.\text{ran.max}]) \]

To specify the behavior denoted by a interrupting timed exception flow, we use the interrupt operator (\( \triangle \)) in CSP+T, according to the following pattern,

\[ P(S_1) = (S_1.\text{ran.min} \parallel v.S_1 \rightarrow \text{SKIP} \uparrow (I(\text{Time}, v.S_1 + S_1.\text{ran.min}) \parallel \epsilon_{\text{exc}} \rightarrow (\text{SKIP} \uparrow P(S_2) \parallel (\text{itime}.\text{ran} + v.S_1) \rightarrow \text{SKIP} \uparrow \epsilon_{\text{exc}} \rightarrow \text{SKIP}) \parallel \text{end} \rightarrow \text{SKIP})) \]

To represent a non-interrupting exception flow, we only need to substitute the interrupt operator (\( \triangle \)) for the parallel operator (\( \parallel \)) or the interleaving (\( \| \)) of activities in CSP+T.

4 CHOREOGRAPHY ANALYSIS

Any choreography can be considered realizable if all the interactions that we have specified in the BPMN 2.0 diagram (e.g., the one in Figure 5) are equivalent to those that can be executed by the interacting peers when we implement the model in a service-description language, such as WS-CDL.

Labelled transition system models (LTS) (Bezem M., 2003) reify choreographies and then allows the verifier to check its realizability w.r.t. the model of the system composed of interacting peers, which are also individually described as LTS (Figure 8). The reification of the case study choreography is shown in Figure 9.

4.1 Behavioral Equivalence

We encode both, the choreography and the interacting peers, into bpmn process terms that are semantically defined by the CSP+T (Zic, 1994) process calculus. In this way, we can count on a sound and well defined formalization of behavioral aspects of BPMN models in order to allow the designer to carry out an analysis of choreographies.
The proposed analysis procedure for checking choreography realizability consists of the following integrated steps,

1. Generate the CSP+T encoding of a given choreography.
2. The derived choreography reification (LTS) is generated from the CSP+T description.
3. For the extracted interacting peers of the choreography, each peer is encoded as one CSP+T process (using the bpmn pattern).
4. The distributed system model is built as a parallel composition of the structured process terms.
5. The choreography reification is model-checked, to prove behavioral equivalence with the peer-based and distributed system model (in 4).

If the two models mentioned above are behaviorally equivalent then the model-checker response is void, meaning that the peer generation exactly satisfies the BPMN communication requirements. On the contrary, if the peers do not generate the same interactions as the ones specified in the choreography then a counter-example is returned by the model-checker, and we can say that this choreography is unrealizable.

5 CHOREOGRAPHY MODEL CHECKING

For each participant in the BCD, we specify a parallel composition of parallel CSP+T processes, which is given by a syntactically correct bpmn term. The proposed formal specification abstracts the internal interaction between the individual peer states and only represents the sequence of task initializations and terminations that occur in the choreography model. A compact model susceptible of being transformed into an LTS, and then verified by an automatic tool, is obtained.

The states of each two main interacting peers in the BCD of Figure 5 are represented by the following bpmn pattern of a CSP+T process,

\[
\text{name} = (Y \mid \alpha(P_i) \mid X) \text{hide } \{\text{connect.Customer}\} \\
\text{Y} = \bigparallel_{i=1}^{n} \text{Tasks} \bullet \alpha(P_i) \circ P(i) \\
\text{name} = (\text{Customer} \parallel \text{Ward} \parallel \text{Pharmacy} \parallel \text{DB}); (\text{peers}) \\
\text{Tasks} = (\text{Demand drug} \parallel \text{Send internal order} \parallel \text{Prepare order} \parallel \text{Search drug} \parallel \text{Deliver drug} \parallel \text{Purchase order} \parallel \text{Prepare drug});
\]

The parallel composition of CSP+T \(P(i)\) processes is mechanically obtained according with the language operators rules.

The process terms above are reified by the two LTS in Figure 8: (a) for bpmn Ward and (b) for bpmn Pharmacy.

\[
P(\text{System}(\text{LTS})) = \text{Ward} \parallel \text{Pharmacy} \parallel \text{Customer} \parallel \text{DB} \]

CSP+T term is the choreography encoding that should be transformed into a timed LTS as the one shown in Figure 9, and then checked against the peer system-LTS (Figure 8) before starting the implementation of the distributed application system.

5.1 Choreography Reification

\[P(\text{System}(\text{LTS})) = \text{Ward} \parallel \text{Pharmacy} \parallel \text{Customer} \parallel \text{DB}\]

CSP+T term is the choreography encoding that should be transformed into a timed LTS as the one shown in Figure 9, and then checked against the peer system-LTS (Figure 8) before starting the implementation of the distributed application system.

5.2 Verification

In our proposal, the reification of the choreography \(P(\text{System}(\text{LTS}))\) is considered realizable if the set of interactions specified by the process term \(P(\text{System})\) and those executed by the interacting peers in \(P_{\text{BPMN}}(\text{LTS})\), are the same. Thus, according to traces and failures semantics of CSP (Schneider, 2000), it must be ascertained that the following refining assertion is true,

\[P(\text{System}(\text{LTS})) \equiv_F P_{\text{BPMN}}(\text{LTS})\] (1)

The model-checker FDR2 (FormSys, 2005), however,
returned \textit{false} since the trace \texttt{< connect, demand drug, send order, prepare order, delivered, preparedrug >} appears in both models, but the trace \texttt{< connect, search drug, prepare order, send order, purchase order, abort >} is present in the peers-based distributed system and not in the LTS of the choreography.

The solution to this error in the LTS model is to make explicit an extra state in which the LTS is waiting for completing the purchase of the drug and add a timeout to this state. If that time period expires then the LTS will reach an \textit{abort} state signifying that the purchase is cancelled, since probably the distributor's stock has been exhausted. The new LTS can be seen in Figure 10.

6 SUPPORTING TOOLS AND MEASUREMENTS

In order to fully support our approach the BTransformer (González A., 2011) tool has been implemented on top of the Eclipse platform (http://www.eclipse.org) as a plugin. BTransformer allows a business process designer to generate a specification in CSP+T from a BPMN model. The tool allows for the creation, the editing, the deleting and the storing of a set of the transformation rules of BPMN 2.0 into CSP+T. Since BTransformer tool is able to read input/output models written in standard XML files, it can be easily extended to encode BPMN 2.0 models into other notations of which can be defined transformation rules that are semantically sound. For instance, the set of transformations of BPMN 2.0 into LOTOS NT described in (Poizat P., 2012) have been recently used to extend the output models generated by BTransformer into LOTOS NT processes, which is one of the input formats accepted by CADP Toolbox (Garavel A., 2011) for checking BP models realizability and verification (Figure 11).

BTransformer permits to automatically generate formal specifications in several process languages, such as CSP+T and LOTOS NT, and business process orchestration languages as WS-BPEL. BTransformer is based on the \textit{model transformation language ATLAS (ATL)} (Jouault F., 2008) and is implemented as an Eclipse plugin, which is capable of transforming BPMN 2.0 models designed with the editor Intalio (http://www.intalio.com) into output files of different formats. A UML class diagram that shows the components of the plugin that provides the functionality of BTransformer tool is shown in Figure 12. A detailed description of BTransformer implementation is given in (González A., 2011).

The BPMN model and case study used to assess the results on the choreography realization model checking and the operation of the implemented BTransformer tool were taken from the “Logistic Process in Hospitals” (Baacke L., 2009). A series of heuristics or criteria have been defined in order to detect inconsistencies between the notational BPMN elements in the source model and processes generated, according to our proposal, for the target model:

- \textbf{Completeness}: all elements in the source BPMN model diagram appear reflected in their semantically equivalent specifications as process terms.
- \textbf{Number of processes}: there is at least one process activity defined in the diagram.
- \textbf{Completeness of relations}: it is possible to establish relationships between two processes as long as they present a relationship between two or more activities in the model.
- \textbf{Behavioral safety}: the set of execution sequences of the processes in the specification must be included (or coincide with) in the set of sequences of the model.

Table 1 shows the results obtained by applying the \textit{consistency} criteria to the CSP+T encoding obtained with BTranformed tool for the complete case study.
Table 1: Results of tests obtained with BTransformer tool.

<table>
<thead>
<tr>
<th>Model</th>
<th>Completeness of elements</th>
<th>Number of processes</th>
<th>Completeness of relations</th>
<th>Behavioral safety</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hospital Logistics</td>
<td>✓</td>
<td>22</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

7 CONCLUSION

In order to enforce the realization of feasible choreographies within the realm of business processes, we have presented a feasible formalization of a subset of BPMN 2.0 constructs.

Consequently, a possible solution to the choreography realization problem is presented here by encoding BCD (BPM 2.0 Choreography Diagrams) modelling-entities into a proces calculus named CSP+T. In this way, we can analyze and automatically verify the realizability of a defined choreography into services that communicate through messages in a general, distributed, and highly parallel system. In a longer term perspective, we also aspire to develop supporting tools for business analysts and modelers to find a way to improve the quality of their business models.

REFERENCES


