Choreographies in the Wild

Massimo Bartoletti, Julien Lange, Alceste Scalas, Roberto Zunino

\textit{Dipartimento di Matematica e Informatica, Università degli Studi di Cagliari, Italy}
\textit{Department of Computing, Imperial College London, UK}
\textit{Dipartimento di Matematica, Università degli Studi di Trento, Italy}

\textbf{Abstract}

We investigate the use of choreographies in distributed scenarios where, as in the real world, mutually distrusting (and possibly dishonest) participants may be unfaithful to their expected behaviour. In our model, each participant advertises its promised behaviour as a \textit{contract}. Participants may interact through multiparty sessions, created when their contracts allow to synthesise a \textit{choreography}. We show that systems of honest participants (which always adhere to their contracts) enjoy \textit{progress} and \textit{session fidelity}.

\textbf{Keywords:} contracts, choreographies, session types

1. Introduction

Designing reliable large-scale distributed systems is a difficult challenge. Several intrinsic issues (e.g. concurrency, physical distribution, non-reproducibility of bugs, etc.) make it extremely unlikely that any non-trivial distributed application actually works as its designers intended. To cope with this grand challenge, a lot of foundational and applied research has been carried over in the last 15 years.

One of the most notable approaches is that which studies session types \cite{1, 2, 3} as an abstraction of the communication behaviour of distributed applications. In the top-down approach to session types, the designer of a distributed application specifies its overall communication behaviour through a \textit{choreography}, which validates some global properties of the application (e.g. safety and deadlock-freedom). However, a distributed application is nothing more than the composition of a number services, running on different nodes in the network. In order to globally enjoy the properties validated by the choreography, all the components forming the application have to be verified; this can be done e.g. by projecting the choreography to \textit{end-point types}, against which these services are actually type-checked.

The above-mentioned top-down approach is quite suitable in case the designers have complete control over the application, e.g., when they develop all the needed services. However, in many realistic situations this is not the case: for instance, one might need to integrate with already-existing third-party services, e.g., e-commerce facilities, maps, etc. In this case, the top-down approach might not work as expected, because it could be impossible to adapt the end-point type projected from the choreography to the actual interface of the third-party service. In the last few years, approaches mixing top-down and \textit{bottom-up} composition have been studied to cope with this kind of situation \cite{4, 5, 6}.

Dropping the (unrealistic) hypothesis that a distributed application is built of services coming from a single provider has a further consequence. We can no longer suppose that services live in a “civil society” where they smoothly collaborate with each other; rather, they live in a “wild” environment (like Hobbes’
condition of mere nature [7]) where they must compete for resources, and possibly diverge from the intended behaviour in case they find it convenient to do so. Indeed, this situation is quite reasonable in inter-organisational scenarios, where mistrusted providers compete to achieve conflicting goals. Although some proposals like OASIS [8] try to deal with this issue by prescribing runtime monitoring and logging, no general techniques exist for checking that services actually adhere to their specifications, unless one controls the “bottom layer” of the software stack [9].

To overcome the state of nature, and move towards the “social contract” (as in Hobbes’ Leviathan), some recent works have proposed to discipline the interaction of services through contracts [10, 11]. The idea is a contract-oriented, bottom-up composition, where only those services with compliant contracts can establish sessions through which they interact. To give some intuition about the underlying dynamics, let us assume three participants A, B and K, where the latter plays the role of a contract broker. A possible interaction among these participants could be the following:

1. A advertises its contract \( c_A \) to K;
2. B advertises its contract \( c_B \) to K;
3. K checks if \( c_A \) and \( c_B \) are compliant; if so, it creates a fresh session \( s \) between A and B;
4. at this point, A and B can start interacting through \( s \), by performing the communications prescribed by their contracts. While doing that, they can advertise other contracts, so establishing other sessions (possibly, with different participants).

One key aspect of this contract-oriented approach is that the runtime behaviour of a service may diverge from the advertised one. For instance, participant B above may fail to perform some of the actions in \( c_B \) (either maliciously or accidentally); if A assumes that B respects its contract, then A may get stuck on session \( s \), and consequently it may fail to fulfil its contracts in some other sessions (so possibly damaging itself or other participants). A notable question then arises: how can designers guarantee that their services are honest, in that they always respect the advertised contracts, despite the possible dishonesty of the services they may interact with?

In this paper we address the issue of reconciling the top-down and the bottom-up approaches to service composition in wild environments. To this purpose, we propose a combination of two approaches. Like in the contract-oriented approach [10], we assume that applications are built bottom-up, by composing services whose advertised contracts admit agreements. However, unlike in previous contract-oriented calculi, service contracts are local session types, in a form similar to those used in [12, 13]. A set of contracts admits an agreement whenever it is possible to synthesise from them a choreography — i.e. a global type — whose projections are the contracts themselves [6]. The existence of such a choreography ensures that, if all services behave honestly, the overall application enjoys progress and session fidelity. However, by the wilderness assumption, the actual behaviour of services may deviate from the promised one. In this case, progress no longer holds, but at least we can single out the services which are responsible for contract violations.

Contributions. We have combined notions and techniques from the session types literature to devise a model of distributed systems where services may be composed bottom-up according to their contracts, and they may deviate from the expected behaviour at runtime. Contracts are expressed as local session types, and multiparty agreements are constructed by adapting the synthesis technique for choreographies of [6]. This allows us to import some results from the session types literature — in particular, the existence of a choreography ensures that contractual agreements enjoy safety and progress (Theorem 3.9). Using choreographies as the basis for agreements is quite flexible: indeed, it allows us to impose further constraints before establishing sessions, e.g. on the number of involved participants, whether or not the session may terminate, etc.

We model services in the \( \text{CO}_2 \) calculus [14], which is adapted here to use (multiparty) local session types as contracts. We define when a service is culpable (i.e., responsible for the next move) in a session, and we show that at least one culpable participant exists in each non-terminated session (Theorem 4.10). If
a system gets stuck, it is always possible to identify which participants violated their contracts. We then adapt the notion of honesty of \cite{11}. We show that an honest service is always able to satisfy its contracts by firing some of its actions, even when the other services do not cooperate (Theorem 6.1). Furthermore, we show that systems of honest services preserve the properties enjoyed by choreographies: in particular, session fidelity (Theorem 6.2) ensures that no extra-contractual interactions are possible in a session, and that honest participants will eventually perform the actions required by their contracts.

We present a case study where the key features of our framework are put into practice. Multiparty sessions, honest participants, and choreography-based contractual agreements are used to model distributed systems where brokers let participants start a new session only if they are advertising contracts such that, as a whole, they implement a class of gossip protocols \cite{13}.

**Synopsis.** The rest of the paper is structured as follows. In Section 2 we sketch an example that we will use through the paper to illustrate our framework. In Section 3 we introduce a contract model based on multiparty session types, and we define contractual agreement as choreography synthesis. In Section 4 we present CO2, suitably adapted to deal with the contracts of Section 3, and we highlight some of its main features. In Section 5 we formalise the notion of honesty. Our main technical results are presented in Section 6 where we show how global properties of choreographies are projected to (honest) CO2 systems \cite{14}. In Section 7 we test the suitability of our framework through a large case study. Finally, in Section 8 we discuss some related work, and in Section 9 we draw some conclusions.

### 2. A Motivating Example

We introduce a running example which we will use through the paper to illustrate our framework. To make symbols lookup easier, we have summarised the syntactic categories and some notation in Table 1 some of the symbols defined therein will only be used in later sections.

An online store \( \mathcal{A} \) allows two (still unknown) buyers \( \mathcal{B}_1 \) and \( \mathcal{B}_2 \) to make a joint purchase through a simple protocol: after they both request the same item, a quote is sent to \( \mathcal{B}_1 \), who is then expected to either place an order (order) or end the session (bye); the store also promises to notify \( \mathcal{B}_2 \) about whether the order was placed (ok) or cancelled (bye). The behaviour promised by \( \mathcal{A} \) is described by the following contract:

\[
c_a = b_1\text{?req}; b_2\text{?req}; b_1\text{!quote}; (b_1\text{?order}; b_2\text{?ok + b}_2\text{?bye}; b_2\text{!bye})
\]

What kind of contracts would be compliant with \( c_a \)? One answer consists in the following contracts, advertised by buyers \( \mathcal{B}_1 \) and \( \mathcal{B}_2 \):

\[
c_{b_1} = a\text{!req}; a\text{!quote}; (b'_1\text{!ok}; a\text{!order} \oplus b'_2\text{!bye}; a\text{!bye})
\]

\[
c_{b_2} = a'\text{!req}; (b'_1\text{?ok}; a'\text{?ok} + b'_1\text{?bye}; a'\text{?bye})
\]

Here, \( \mathcal{B}_1 \) promises to send the request to a (still unknown) store \( \mathcal{A} \), wait for the quote, and then notify the other buyer \( b'_2 \) before accepting or rejecting the store offer; symmetrically, \( \mathcal{B}_2 \) promises to send the request to the store \( a' \), and then expects to receive the same notification (either ok or bye) from both the other buyer \( b'_1 \) and the store itself. Each contract represents the local viewpoint of the participant who advertises
it: in particular, \( c_A \) represents the local viewpoint of the store, and thus it does not (and indeed, it cannot) capture the communications between \( B_1 \) and \( B_2 \).

An agreement among \( c_A, c_{B_1}, \) and \( c_{B_2} \) may be found after instantiating the participant variables (in each contract) to actual names, through the substitutions \( \{y/\text{a}, x\}' \), \( \{b_{z}/b_{v}, y\}' \), and \( \{b_{2}/b_{v}, y\}' \). The contracts admit an agreement because a choreography exists which can be projected back to \( c_A, c_{B_1}, \) and \( c_{B_2} \). The existence of such a choreography guarantees progress and safety of the contractual agreement (Theorem 3.9). We can exploit the technique in \([6]\) to synthesise a choreography from a set of contracts. In our example, we obtain:

\[
S_{A B_1 B_2} = \quad B_1 \rightarrow A : \text{req} ; B_2 \rightarrow A : \text{req} ; A \rightarrow B_1 : \text{quote} ; \{B_1 \rightarrow B_2 : \text{ok} ; B_1 \rightarrow A : \text{order} ; A \rightarrow B_2 : \text{ok} + B_1 \rightarrow B_2 : \text{bye} ; B_1 \rightarrow A : \text{bye} ; A \rightarrow B_2 : \text{bye} \}
\]

Global type \( S_{A B_1 B_2} \) says that \( B_1 \) and \( B_2 \) send a request to \( A \), consecutively; then \( A \) replies to \( B_1 \) with a quote. Subsequently, \( B_1 \) chooses whether to accept or reject the quote. In the former case, \( B_1 \) notifies \( B_2 \) (ok message) then \( A \) (order message), who sends a confirmation to \( B_2 \) (ok message). In the latter case, \( B_1 \) sends bye to both \( B_2 \) and \( A \), and \( A \) confirms the cancellation to \( B_2 \) by sending bye as well.

However, in a distributed system, the existence of a contractual agreement among participants does not guarantee that progress and safety will also hold at runtime: in fact, a participant may advertise a contract promising some behaviour, and then fail to respect it — either maliciously or accidentally. Such failures may then cascade on other participants, e.g., if they remain stuck waiting for a promised message that is never sent.

This sort of situations can be modelled using the CO2 calculus \([14]\). For instance, a CO2 system for the above scenario will have the following form:

\[
A[(x) (b_1, b_2) \text{tell}_{A \downarrow x} c_A . \text{fuse}. P_A] \quad | \quad B_1[(y)(a, b')_2 \text{tell}_{A \downarrow y} c_{B_1} . P_{B_1}] \quad | \quad B_2[(z)(a, b')_1 \text{tell}_{A \downarrow z} c_{B_2} . P_{B_2}]
\]

Here, participant \( A \) advertises the contract \( c_A \) to itself via the primitive \( \text{tell}_{A \downarrow x} c_A \), where \( x \) is used as a session handle for interacting with other participants. \( B_1 \) and \( B_2 \) advertise their respective contracts to \( A \) with a similar invocation.

In this example, \( A \) also plays the role of contract broker, which collects the advertised contracts, while looking for an agreement. Once an agreement is found, the fuse prefix is fired, and a new session is established among all the involved participants. At this point, the participants may realise the agreed contracts — or they may even choose not to. In fact, we will see that when the contracts are violated, the calculus allows for responsible participants to be always ruled out (Theorem 4.10). In systems where participants are honest (i.e. they always fulfil their contracts) the progress and safety of the contractual agreement are also reflected in the runtime behaviour of the CO2 system (Theorem 6.1 and Theorem 6.2).

Other possible agreements. Our framework allows for modelling a wide variety of scenarios. For instance, a participant \( B_{12} \) may impersonate both buyers, and promise to always accept the store offer, by advertising the following contract:

\[
c_{B_{12}} = a'' \text{req} a'' \text{req} a'' \text{quote} a'' \text{order} a'' \text{ok}
\]

where the request to the store \( a'' \) is sent twice (i.e., once for each impersonated customer). In this case, if we instantiate the participant variables in \( c_A \) and \( c_{B_{12}} \) with substitutions \( \{b_{1}/a, \} \), \( \{b_{2}/b_{1, b_{2}} \} \), we can find an agreement by synthesising the following choreography:

\[
S_{A B_{12}} = \quad B_{12} \rightarrow A : \text{req} ; B_{12} \rightarrow A : \text{req} ; A \rightarrow B_{12} : \text{quote} ; B_{12} \rightarrow A : \text{order} ; A \rightarrow B_{12} : \text{ok}
\]

This scenario may be modelled by a CO2 system of the form:

\[
S_2 = A[(x)(b_1, b_2) \text{tell}_{A \downarrow x} c_A . \text{fuse}. P_A] \quad | \quad B_{12}[(w)(a'') \text{tell}_{A \downarrow w} c_{B_{12}} . P_{B_{12}}]
\]

where the fuse prefix can now create a session involving \( A \) and \( B_{12} \).

The participants in the CO2 systems \( S_1 \) and \( S_2 \) may also be combined, so to obtain:

\[
S_{12} = A[(x)(b_1, b_2) \text{tell}_{A \downarrow x} c_A . \text{fuse}. P_A] \quad | \quad B_1[(y)(a, b')_2 \text{tell}_{A \downarrow y} c_{B_1} . P_{B_1}]
\]

\[
\quad | \quad B_2[(z)(a, b')_1 \text{tell}_{A \downarrow z} c_{B_2} . P_{B_2}] \quad | \quad B_{12}[(w)(a'') \text{tell}_{A \downarrow w} c_{B_{12}} . P_{B_{12}}]
\]


In this case, after all contracts have been advertised to A, either a session corresponding to $S_{AB:B_2}$, or to $S_{AB:2}$ may take place, thus involving a different number of participants depending on which contracts are fused. As mentioned above, the honesty of A can not depend on which of the two agreements is chosen: to be honest, A must respect its contracts whatever agreement is found with the other participants.

3. A Choreography-Based Contract Model

In this section we exploit concepts and results from the session types literature to devise a multiparty contract model. In Sections 3.1 and 3.2, we present the syntax and semantics of contracts. In Section 3.3, we define the crucial notion of agreement. Intuitively, a set of contracts admits an agreement when there exists a choreography which can be projected back to such contracts. This choreography can be inferred through a type system, adapted from [16], which guarantees safety properties (Theorem 3.9). In Section 3.4, we introduce a notion of culpability. Theorem 3.12 ensures that, if a system of contracts admits an agreement, then at each step of its execution some participants will be “culpable”, i.e. responsible for making the next step in a non-terminated system of contract.

3.1. Local session types as Contracts

We express contracts using the syntax of local session types, in the style of [12, 13]. Let $\mathcal{P}$ and $\mathcal{P}$ be disjoint sets of, respectively, participant names (ranged over by $A, B, \ldots$) and participant variables (ranged over by $a, b, \ldots$), and let $\mathcal{P}_A, \mathcal{P}_B$ range over $\mathcal{P} \cup \mathcal{P}$. We abstract from the actual data exchanged between participants by using sorts $e, e', \ldots$. We denote (individual) contracts with $c, c', \ldots$, and systems of contracts with $T, T', \ldots$. We write $\mathcal{P}(T)$ for the set of participant names in $T$.

Definition 3.1 (Contracts). The syntax of contracts is given below (we use the colour blue for contracts and green for systems of contracts).

\[
\begin{align*}
  c &::= \bigoplus_{i \in I} a_i e_i; c_i \quad \text{where } i \neq j \in I: (a_i, e_i) \neq (a_j, e_j) \quad \text{(Internal choice)} \\
  &\quad \mid \sum_{i \in I} a_i e_i; c_i \quad \text{where } i \neq j \in I: e_i \neq e_j \quad \text{(External choice)} \\
  &\quad \mid \mu x.c \quad \text{(Recursion)} \\
  &\quad \mid x \quad \text{(Recursion variable)} \\
  &\quad \mid 0 \quad \text{(End)} \\
  T &::= T \mid T' \quad \text{where } \mathcal{P}(T) \cap \mathcal{P}(T') = \emptyset \quad \text{(System composition)} \\
  &\quad \mid A(c) \quad \text{(Participant’s contract)} \\
  &\quad \mid (AB) : \rho \quad \text{where } A \neq B \text{ and } \rho \text{ is a sequence of sorts} \quad \text{(Queue)} \\
  &\quad \mid 0 \quad \text{(End)}
\end{align*}
\]

Intuitively, in an internal choice, after sending the message $e_i$ to participant $a_i$, behaviour $c_i$ takes place; whereas in an external choice if a message of sort $e_i$ is received from $a_i$, then behaviour $c_i$ takes place. We write $fv(c)$ for the (free) participant variables in $c$. We identify empty internal and external sums, i.e. $0 = \bigoplus_{i \in \mathcal{P}} a_i e_i; c_i = \sum_{i \in \mathcal{P}} a_i e_i; c_i$.

A system of contracts $T$ may be either: (i) a parallel composition of systems $T \mid T'$; or (ii) a named contract $A(c)$, saying that participant A promises to behave according to $c$; (iii) a queue $(AB) : \rho$ of messages from A to B; we write $[]$ for the empty sequence of message. In a system $T$, we assume that there are exactly two uni-directional queues per pair of participants; that is, two participants A and B can only communicate via queues named $AB$ (from A to B) and $BA$ (from B to A). Whenever we refer to a system $T$, we assume that it respects the conditions in Definition 3.1 as well as those above.
Hereafter, we will use the following notation. We write:

\[ A(B; c_0 \oplus c_1) \mid (AB): \rho \quad \overset{\Delta=\rightarrow}{\longrightarrow} \quad A(c_0) \mid (AB): \rho \cdot e \]

\[ A(B; c_0 + c_1) \mid (BA): e \cdot \rho \quad \overset{\Delta=\rightarrow}{\longrightarrow} \quad A(c_0) \mid (BA): \rho \]

\[ \mu \times e \equiv c \{ \mu \times e / x \} \quad \overset{\Delta=\rightarrow}{\longrightarrow} \quad \mu \]

\[ A(0) \equiv 0 \]

\[ T \lambda \overset{\Delta=\rightarrow}{\longrightarrow} T' \]

commutative monoidal laws for |, \oplus and +

\[ T_1 | T_2 \lambda \overset{\Delta=\rightarrow}{\longrightarrow} T'_1 | T'_2 \]

Figure 1: Semantics of systems of contracts.

3.2. Semantics

The semantics of systems of contracts is given in Figure 1, where labels \( \lambda \) have either the form \( A \rightarrow B : e \) or \( A \leftarrow B : e \). The label \( A \rightarrow B : e \) means that a message of sort \( e \) is sent by participant \( A \) to \( B \); instead, the label \( A \leftarrow B : e \) means that a message of sort \( e \) is received by \( A \) from \( B \). The first rule in Figure 1 says that, after an internal choice, \( A \) puts a message \( e \) on its queue for participant \( B \). The second rule says that, in an external choice, \( A \) can receive a message (of the expected sort) from an input queue \( BA \). The semantics is assumed up-to the congruence relation \( \equiv \) defined in Figure 1.

**Notation 3.2.** Hereafter, we will use the following notation. We write:

- \( T \overset{\Delta=\rightarrow}{\longrightarrow} T' \) when either \( T \overset{\Delta=\rightarrow}{\longrightarrow} T' \) or \( T \overset{\Delta=\rightarrow}{\longrightarrow} T' \).
- \( T \overset{\Delta=\rightarrow}{\longrightarrow} T' \) when \( T \overset{\Delta=\rightarrow}{\longrightarrow} T' \), and the label is unimportant.
- \( \mathcal{P}(T) \) (resp. \( \mathcal{C}(T) \)) for the set of participant (resp. queue) names in \( T \).
- \( \mathcal{Q}(T) \) for the parallel composition of the empty queues connecting all distinct pairs of participants appearing in \( T \).
- \( T(A) \) for the contract of \( A \) in \( T \), and \( T[AB] \) for the content of the queue \( AB \), i.e.:

\[
T(A) = \begin{cases} 
\text{c} & \text{if } T \equiv A(c) \mid T' \\
\perp & \text{otherwise}
\end{cases} \quad T[AB] = \begin{cases} 
\rho & \text{if } T \equiv (AB): \rho \mid T' \\
\perp & \text{otherwise}
\end{cases}
\]

**Example 3.3.** From the example in Section 2, consider the instantiated contracts of the store \( A \) and its customer \( B_{12} \). We illustrate the system, and how it progresses (hereafter, for the sake of readability, we often omit empty queues).

\[
T_{AB_{12}Q} = A(c_4 \{ A/\rho' \} \{ B_{12}/b_{1,2} \}) \mid B_{12}(c_{B_{12}} \{ A/\rho' \} \{ B_{12}/b_{1,2} \}) \mid (B_{12}A):[]
\]

\[
\overset{B_{12} \rightarrow A/req}{\longrightarrow} A(B_{12}/req; B_{12}/req; \ldots) \mid B_{12}(A/req; A/req; \ldots) \mid (B_{12}A):[]
\]

\[
\overset{\Delta=\rightarrow}{\longrightarrow} A(B_{12}/req; \ldots) \mid B_{12}(A/req; \ldots) \mid (B_{12}A):[]
\]

3.3. Choreography Synthesis as Agreement

We introduce in Definition 3.5 below the agreement relation, that tells whether some contracts can be combined to describe a correct interaction. Roughly, a set of contracts admits an agreement if it can be assigned a choreography. To do that, we borrow some results from [6][10]; in particular, we exploit a typing system which assigns a (unique) choreography to a set of contracts.

**Definition 3.4.** The syntax of choreographies \( \mathcal{G} \) is defined as follows:

\[
\mathcal{G} ::= A \rightarrow B; e; \mathcal{G} \mid \mathcal{G} + \mathcal{G}' \mid \mathcal{G} \mid \mathcal{G}' \mid \mu \times \mathcal{G} \mid x \mid 0
\]
The first production means that \( A \) sends a message of sort \( e \) to \( B \), then interactions in \( \mathcal{G} \) take place; \( \mathcal{G} \uplus \mathcal{G}' \) means that either interactions in \( \mathcal{G} \) or in \( \mathcal{G}' \) take place; \( \mathcal{G} \uplus \mathcal{G}' \) means that interactions in \( \mathcal{G} \) and \( \mathcal{G}' \) are executed concurrently; the rest of the productions are for recursive interactions, and termination \( (0) \).

To synthesise a choreography \( \mathcal{G} \) from a system of contracts \( T \), we analyse what actions can occur at each possible state of \( T \). These actions are contained in the set \( \text{TRS}(T) \) defined below. The auxiliary predicate \( T \updownarrow \) holds whenever two participants are able to synchronise (i.e., \( B \) may receive the message sent by \( A \) as soon as it appears on the queue).

**Definition 3.5 (System of contracts Ready Set).** We define the ready set of a system as follows:

\[
\text{TRS}(T) = \begin{cases}
\{ A \leftarrow B \} \cup \text{TRS}(T') & \text{if } T \equiv A \langle \sum_{i \in I} B^i e_i; c_i \rangle | T' \\
\{ A \rightarrow B_i \mid i \in I \} \cup \text{TRS}(T') & \text{if } T \equiv A \langle \bigoplus_{i \in I} B^i_1 e_i; c_i \rangle | T' \\
\emptyset & \text{if } T \equiv 0
\end{cases}
\]

We write \( T \updownarrow \) when \( \exists A, B : A \rightarrow B \in \text{TRS}(T) \land B \leftarrow A \in \text{TRS}(T) \). We write \( T \nmid \) if \( T \updownarrow \) does not hold.

In Figure 2 we present the synthesis of choreographies from systems of contracts (this is a slight adaptation of the synthesis in [16]). Typing judgements have the form \( \Gamma \vdash \text{hypothesis } \Gamma, \text{system } T \rightarrow \mathcal{G} \), where \( \Gamma \) is an environment which keeps track of recursion variables. Roughly, \( \Gamma \) maps pairs \( (A, x) \) to \( x' \), where \( x \) is the local recursion variable of participant \( A \) and \( x' \) is a global recursion variable. A typing judgement \( \Gamma \vdash T \rightarrow \mathcal{G} \) says that, under hypothesis \( \Gamma \), the system \( T \) is assigned to global type \( \mathcal{G} \).

We write \( \mathcal{P}(\mathcal{G}) \) (resp. \( \mathcal{E}(\mathcal{G}) \)) for the set of participant (resp. queue) names in \( \mathcal{G} \).

**Definition 3.6 (Agreement).** We say that a system of contracts \( T \) admits an agreement when \( \vdash T \rightarrow \mathcal{G} \) can be inferred from the rules in Figure 2.

Essentially, the synthesis rules allow to execute a set of contracts step-by-step, while keeping track of the structure of the interactions in a global type. The rules are driven by the ready set of \( T \) and the structure of its processes. We briefly describe them below.

Sequential interactions are dealt with by rule \([.]\). The rule validates prefixes provided that the continuation is typable and that no other interactions are possible in \( T \). For instance, rule \([.]\) does not apply to

\[ A_1 \langle B_1^1 e_1; c_1 \rangle \mid B_2 \langle A_2^1 e_1; \nu_2 \rangle \mid A_2 \langle B_2^2 e_2; \nu_2 \rangle \mid B_2 \langle A_2^2 e_2; c_2 \rangle \quad \times \]

because there is no ordering relation between the actions of \( A_1 \) and \( B_1 \) on one hand, and the actions of \( A_2 \) and \( B_2 \) on the other hand. In such cases, the rule \([.]\) must be used. In fact, the premise \( T \updownarrow \) guarantees that a synthesised global type is unique (up-to structural congruence), cf. [16 Theorem 3.3].

Concurrent branches are introduced by rule \([+].\). The rule validates concurrent branches when they can be validated using a partition of the system being considered (recall that \( \mathcal{P}(\mathcal{G}) \cap \mathcal{P}(\mathcal{T}^\prime) = \emptyset \) by assumption).

Choice in a global type is dealt with by rules \([\#]\) and \([+].\). Rule \([\#]\) introduces the global type choice operator, it requires both branches to be typable and that no other interactions are possible in \( T \) — for the same reason as in rule \([.]\). Rule \([+]\) discharges a branch of an external choice. This allows, e.g., to assign a type to systems of the form:

\[ A \langle B^1 e \uplus B^2 e_1 \rangle \mid B \langle A^1 e + A^2 e_2 \rangle + A^2 e_2 \quad \checkmark \]

Recursion is handled by rules \([\mu]\) and \([\kappa]\). The former rule “guesses” the participants involved in a recursive behaviour. If two of them interact, \([\mu]\) validates the recursion provided that the system can be typed when such participants are associated to the global recursion variable \( x \) — assuming \( x \) is fresh. Rule \([\kappa]\) checks that all the participants in the recursion have reached a local recursion variable corresponding to the global recursion. The termination \( 0 \) is introduced by rule \([0]\), which only applies when all the participants in \( T \) are terminated. Rule \([\equiv]\) validates a system up-to structural congruence, so allowing recursive behaviour to be unfolded.
The main properties that we are interested in — and that are guaranteed by the synthesis — are (i) that the inferred global type is projectable back to the original contracts and (ii) that if a system of contract $T$ is typable, then $T$ is “safe”, cf. Theorem 3.9 below. The synthesised global types also satisfy other well-formedness properties which we discuss briefly in Section 8.2.

**Example 3.7.** Recall the buyer-seller scenario from Section 2. By combining the contract of store $A$ with those of customers $B_1$ and $B_2$, we obtain the system:

$$T_{AB,B_1,B_2} = A(x_1) \mid B_1(x_1) \mid B_2(x_1)$$

The system $T_{AB,B_1,B_2}$ admits an agreement: indeed, we have that $\vdash T_{AB,B_1,B_2} \rightarrow \mathcal{G}_{AB,B_1,B_2}$ (where the choreography $\mathcal{G}_{AB,B_1,B_2}$ is that defined at page 7). Similarly, if we combine the store $A$ with $B_{12}$, we have that $T_{AB,B_{12}}$ admits an agreement, since $\vdash T_{AB,B_{12}} \rightarrow \mathcal{G}_{AB,B_{12}}$ (where $\mathcal{G}_{AB,B_{12}}$ is the choreography at page 7).

Theorem 3.9 below establishes some properties of systems of contracts which do admit an agreement. We first define some undesirable properties of systems. A system is a **deadlock** if all its queues are empty, there is one participant expecting a message, and all participants are either terminated or waiting for an input. An **unspecified reception configuration** is a system for which there is a participant $B$ who is permanently unable to read any messages from its queues, i.e., there are **unexpected** messages in $B$’s buffers. A system has a **self-interaction** if some participant is expected to send a message to himself (or to receive a message from himself). Definition 3.8 formalises these concepts.

**Definition 3.8 (Deadlock/unspecified reception).** For a system of contracts $T$, we say that:

- $T$ is a **deadlock** if the following conditions hold:
  
  $$\forall AB \in \mathcal{C}(T): T[AB] = []$$
  
  $$\exists C \in \mathcal{P}(T): T(C) \equiv A?ec + c'$$
  
  $$\forall C \in \mathcal{P}(T): T(C) \equiv A?ec + c' \lor T(C) \equiv 0$$

- $T$ is an **unspecified reception configuration** if $\exists B \in \mathcal{P}(T)$:

  $$T(B) \equiv A?ec' + c_2 \land (\forall e, A \in \mathcal{P}(T): T(B) \equiv A?ec + c' \rightarrow T[AB] \neq [] \land T[AB] \neq e \cdot \rho)$$

- $T$ has a **self-interaction** if there exists $A \in \mathcal{P}(T)$ such that $T'(A) \equiv A?ec + c'$. 

Figure 2: Inference of global types.
Theorem 3.9 (Agreement). If $T$ admits an agreement, then for all $T''$ such that $T \trianglerighteq q(T) \rightarrow^*$ $T'':$

1. $T'$ is not a deadlock.
2. $T'$ is not an unspecified reception configuration.
3. $T'$ has no self-interactions.
4. Either there is $T'''$ such that $T' \rightarrow^* T'''$, or
   $$\forall AB \in C(T'): T'[AB] = [] \quad \text{and} \quad \forall C \in P(T'): T'(C) \equiv 0$$

Proof. Items (1), (2) and (4) follow from Theorem 3.1 in [16]. Item (3) is a direct consequence of Theorem 3.9 and of the assumption that for all $AB \in C(T)$: $A \neq B$. \qed

3.4. Culpability

We introduce a notion of culpability in systems of contracts. Intuitively, a participant is culpable in a system if it is able to send or receive some message. This condition is not necessarily permanent: a culpable participant may exculpate itself by sending/receiving the required messages.

Definition 3.10 (Contract culpability). $A$ is culpable in $T$ when $T \overset{A=\text{send } e}{\Rightarrow} B$, for some $B$ and $e$.

As expected, a participant $A$ is not culpable in a system of contracts not including any contracts of $A$.

Lemma 3.11. For all system $T$, if $A \notin P(T)$, then $A$ is not culpable in $T$.

The following theorem ensures that, if a system of contracts admits an agreement, then at each step of its execution (until reaching the final state $0$) some participant is culpable. This guarantees that, unless some participants want to stay culpable forever, the system of contracts enjoys progress.

Theorem 3.12 (Presence of culpability in contracts). If $T$ admits an agreement and $T \trianglerighteq q(T) \rightarrow^*$ $T' \neq 0$, then there exists at least one culpable participant in $T'$.

Proof. Since $T' \neq 0$ then, for some $A, B, e, c, c'$, either (a) $T' \equiv A(B|e;c \circ c')$ or (b) $T' \equiv A(B|e;c + c')$. Case (a) follows directly, since a send action can always be fired. Case (b) follows from Theorem 3.9, i.e., there must be a message for $A$ in a queue $BA$, otherwise we would have a deadlock or an unspecified reception configuration. \qed

4. Introducing Choreographies to the Wild

We present a variant of the CO₂ calculus (for COntract-Oriented computing) [14] which we adapt to deal with the multiparty contracts and sessions introduced in Section 3. In Sections 4.1 and 4.2 we give its syntax and semantics, while in Section 4.3 we explain in detail how new sessions are established, providing several examples. In Section 4.4 we formalise when participants in CO₂ systems are “culpable”, and show that a non-terminated session always implies the (possibly transient) culpability of some participant (Theorem 4.10). Finally, in Section 4.5 we discuss how the session creation mechanics can be adjusted in order to enforce some desired properties on the resulting interactions.
commutative monoidal laws for $+$ and $|$, with $0$ as identity element:

\[
\begin{align*}
A((u)P) &= (u)A[P] \\
A((u)vZ) &= (v)(u)Z \\
A((u)Z) &= Z \iff u \notin \text{fnv}(Z)
\end{align*}
\]

Figure 3: Structural congruence rules for CO$_2$.

4.1. Syntax

Let $S$ and $S'$ be disjoint sets of, respectively, session names (ranged over by $s,s',t,...$) and session variables (ranged over by $x,y,z,...$). Let $u,v,...$ range over $S \cup S'$. The syntax of CO$_2$ is given as follows:

\[
\begin{align*}
\text{(Processes)} & \quad P,Q ::= \sum_{i \in I} p_i . P_i \mid P \mid Q \mid (\vec{a})P \mid \text{do}^e_x \mid \text{fnv}(\vec{a})S \\
\text{(Prefixes)} & \quad p ::= \tau \mid \text{tell}_{\vec{a}} \mid \text{fuse} \mid \text{do}_x^e \\
\text{(Latent contracts)} & \quad K ::= \downarrow_a \text{ says } c \mid K \mid K \\
\text{(Systems)} & \quad S ::= A[P] \mid A[K] \mid s[T] \mid S \mid S \mid (\vec{a},\vec{a})S \mid 0
\end{align*}
\]

CO$_2$ features CCS-style processes, equipped with branching $\sum$ (not to be confused with the choice operator used in contracts), parallel composition $|$, restrictions of session and participant variables, and named process invocation. As usual, we denote the empty sum with $0$, and omit its trailing occurrences. We assume that each process identifier $X$ has a unique defining equation $X(y_1,...,y_n,a_1,...,a_m) := P$ such that $\text{fv}(P) \subseteq \{y_1,...,y_n,a_1,...,a_m\}$ and each occurrence of process identifiers in $P$ is prefix-guarded. We shall take the liberty of omitting the arguments of $X(\vec{y},\vec{a})$ when they are clear from the context. The prefixes are for internal action ($\tau$), contract advertisement ($\text{tell}$), session creation upon contractual agreement ($\text{fuse}$), and execution of contractual actions ($\text{do}$).

A latent contract of the form $\downarrow_a \text{ says } c$ represents the promise of participant $A$ to fulfil $c$ by executing $\text{do}$-actions, once $x$ has been instantiated with a session name.

CO$_2$ systems may be parallel compositions of participants $A[P]$ (where $P$ is the process being executed by $A$), collections of latent contracts $A[K]$ (where $A$ is the participant to which the contracts in $K$ have been advertised), and sessions $s[T]$ (where $s$ is a session name, and $T$ is a system of contracts as in Section 3). We assume well-formed systems where each participant name $A$ can appear with at most one process — i.e., we rule out systems like $A[P] \mid A[Q]$.

Note that CO$_2$ process and system productions allow to delimit both session names/variables (i.e., but not participant names: this is because participant names are used to uniquely (and statically) identify participants in a network, and therefore we do not want these names to be a-converted.

4.2. Semantics

The semantics of CO$_2$ (Figure 3) is defined up-to a congruence relation between CO$_2$ systems (Figure 3). The rules defining $\equiv$ are mostly standard. In Figure 3, $Z,Z',Z''$ range over processes, systems, or latent contracts; all rules involving delimited session names/variables $u,v$ also hold for delimited variable participants $a,b$ (e.g., $A[(a)P] \equiv (a)A[P]$). Rule $A[(u)P] \equiv A[u]P$ allows to move delimitations between CO$_2$ systems and processes; the last rule $A[A[K] \mid A[K']] \equiv A[K \mid K']$ will be defined in Section 4.3 and in Example 4.4.

**Definition 4.1 (Semantics of CO$_2$).** $\ell$ is the smallest relation between CO$_2$ systems closed under the rules of Figure 4, which are to be considered up-to the structural congruence relation $\equiv$ in Figure 3. The labels $p$ have the form $A : p$, where $A$ is the name of the participant firing prefix $p$.

We briefly discuss the rules in Figure 4. [CO$_2$-TAU] allows a participant $A$ to fire an internal action. [CO$_2$-TELL] allows $A$ to advertise a contract $c$ to $B$; as a result, a new latent contract is created, recording the fact that...
it was promised by $A$. [CO₂-Fuse] establishes a new session, when an agreement is found among the latent contracts held in $A[K]$. Such an agreement is reached only if the relation $K 	riangleright γ _e T$ holds (cf. Definition 4.3). Intuitively, the relation holds if the contracts in $K$ can be synthesised into a well-formed choreography. When an agreement exists, a fresh session name $s$ and the names of the involved participants are shared (via substitutions $σ$ and $π$, respectively); the initial state of the new session consists in the system of contracts $T$ (constructed from $K$) with the communication queues connecting its participants. Rule [CO₂-Do] allows $A$ to perform an input/output action $e$ towards $B$ on session $s$, provided that $T$ (i.e., the current state of the contracts at $s$) allows for it. The other CO₂ rules are standard. In rule [CO₂-Del], the prefix $p$ fired in the premise becomes $τ$ in the conclusions, when $p$ contains a delimited name/variable. For instance, consider the following reduction: $(x)A[\text{tell}_{+x} c . P] \xrightarrow{A[\text{tell}_{+x} c]} (x)(A[P] \mid B[↓x A \text{ says } c])$; one would not use the label $A : \text{tell}_{+x} c$ instead of $A : \tau$, since the variable $x$ is not visible outside the process of $A$.

**Example 4.2.** Consider the CO₂ system:

$$S = A[do^*_s \text{ int} + do^*_b \text{ bool}] \mid s[A(B\text{ int}) \mid B(A?\text{ int}) \mid (AB) : []] \mid B[do^*_b \text{ int}]$$

Participant $A$ is willing to perform an action towards $B$ on session $s$, with either a message of sort $\text{int}$ or $\text{bool}$. However, $A$’s contract in $s$ only specifies that $A$ should send a message of sort $\text{int}$ to $B$: therefore, according to rule [CO₂-Do], only the first branch of $A$ may be chosen, and the system reduces as follows.

$$S \xrightarrow{A[do^*_b \text{ int}]} A[0] \mid s[A(0) \mid B(A?\text{ int}) \mid (AB) : \text{int}] \mid B[do^*_b \text{ int}]$$

$$S \xrightarrow{B[do^*_b \text{ int}]} A[0] \mid s[A(0) \mid B(0) \mid (AB) : []] \mid B[0]$$

4.3. Reaching Agreements

We now discuss the mechanics underlying the [CO₂-Fuse] rule, with a focus on the agreement relation which regulates session establishment.
Definition 4.3 (Agreement relation $K \triangleright^T_s T$). Let $K \equiv _{i \in I} \downarrow x_i \ A_i \ says \ e_i$, and let $\pi : \mathcal{P} \to \mathcal{P}$ and $\sigma : \mathcal{S} \to \mathcal{S}$ be two substitutions mapping participant variables to participant names, and session variables to session names, respectively. Also, let $T \equiv _{i \in I} A_i(e_i)\pi$. Then, $K \triangleright^T_s T$ holds iff:

1. $\text{dom}(\sigma) = \bigcup_{i \in I} \{x_i\}$;
2. $\text{dom}(\pi) = \bigcup_{i \in I} \text{fv}(e_i)$;
3. $T$ admits an agreement — i.e., $\exists \mathcal{S} : \vdash T \triangleright \mathcal{S}$.

Conditions 1 and 2 on $\sigma$ and $\pi$, guarantee that all the session and participant variables are indeed instantiated. Condition 3 holds whenever a judgment $\vdash T \triangleright \mathcal{S}$ can be inferred, for some $\mathcal{S}$ (see Definition 3.6). We discuss two safety issues that Definition 4.3 rules out:

1. our assumptions on a system $T$ imply that each participant may have at most one contract in $K$, cf. Section 3.2. This restriction allows us to avoid problematic scenarios arising during session establishment. For instance, consider a process $A[P]$ advertising two contracts $c$ and $c'$ using session variables $x$ and $y$, respectively. Were we not preventing both contracts to appear in a same session $s$, $c$ and $c'$ could be fused together in $s$. Therefore, by rule $[\text{CO-Fuse}]$, both $x$ and $y$ would be replaced by $s$. Hence, each prefixes of the form $do^s$ and $do^y$ in $P$ would be turned into $do^s$, possibly mixing the actions originally intended for one contract with those intended for the other. In such situations, one cannot enforce safety properties such as deadlock freedom, since contracts $c$ and $c'$ are independent in $T$, while there might be causal dependencies in $P$;

2. by item (3) of Theorem 3.9, it cannot be the case that, within a contract $e_i$ belonging to $A_i$, a free participant variable is substituted by $A_i$ itself.

When searching for possible agreements, the congruence rule $A[K] \equiv A[K'] \equiv A[K \mid K']$ in Figure 3 allows for rearranging collections of latent contracts, in order to select the compliant ones that satisfy the premises of $[\text{CO-Fuse}]$ rule, cf. Example 4.4 below.

Example 4.4 (Agreements and structural congruence). Consider the system:

$$S = (x, y, z)(B[\text{fuse} . P \mid Q] \mid B[K] \mid A_1[\ldots] \mid A_2[\ldots] \mid A_3[\ldots])$$

where $K = \downarrow x \ A_1 \ says \ A_2[\text{int}] \mid \downarrow y \ A_2 \ says \ A_3[\text{int}] \mid \downarrow z \ A_3 \ says \ B[\text{bool}]$.

The $\text{fuse}$ prefix of participant $B$ cannot be immediately fired: no contract matches $A_3$’s, and thus the three latent contracts cannot be assigned a global type. Therefore, there is no $T, \pi$ such that $K \triangleright^T_s T$, as required by rule $[\text{CO-Fuse}]$. However, we can rearrange the latent contracts via structural congruence:

$$S \equiv S' = (x, y, z)(B[\text{fuse} . P \mid Q] \mid B[K'] \mid B[\downarrow x \ A_3 \ says \ B[\text{bool}] \mid A_1[\ldots] \mid A_2[\ldots] \mid A_3[\ldots])]$$

where $K' = \downarrow x \ A_1 \ says \ A_2[\text{int}] \mid \downarrow y \ A_2 \ says \ A_1[\text{int}]$.

If we take $\sigma = \{s/x, y\}$ (with $s$ fresh) and $T = A_1(A_2[\text{int}] \mid A_2(A_3[\text{int}])$, we now have $K' \triangleright^T_s T$. Indeed, $T$ is composed by $A_1$’s and $A_2$’s contracts, and $\vdash T \triangleright A_1 \rightarrow A_2 : \text{int}$; furthermore, the domains of $\sigma$ and $\pi = \emptyset$ comply with the conditions in Definition 4.3.

Since all the premises of $[\text{CO-Fuse}]$ are now satisfied, it is possible to create a new session $s$, involving $A_1$ and $A_2$, whose initial state is $T \mid Q(T)$, and we obtain the system:

$$S' = B[\text{fuse} \rightarrow (z)(s)(B[P \mid Q] \mid s[T \mid Q(T)] \mid B[\downarrow x \ A_3 \ says \ B[\text{bool}] \mid A_1[\ldots] \mid A_2[\ldots] \mid A_3[\ldots])]$$

Note that $A_3$’s contract remains latent, and thus it may be fused later on.
Example 4.5. Going back to our running example, take $c_A$, $c_B$, $c_S$ from Example 3.7. Consider the following CO₂ system, where $P_A$, $P_B$, $P_S$ are given in Example 5.6:

$$S = \{A[(x)(b_1, b_2)] \text{tell} \downarrow x c_A \cdot \text{fuse} \cdot P_A\} \mid \{B_1[(y)(a, b'_2)] \text{tell} \downarrow y c_B \cdot P_B\} \mid \{B_2[(z)(a, b'_1)] \text{tell} \downarrow z c_B \cdot P_B\}\rightarrow\rightarrow\rightarrow\rightarrow$$

$$(x, y, z)[(a, a', b_1, b'_1, b_2, b'_2)](A[\text{fuse} \cdot P_A] \mid B_1[P_B] \mid B_2[P_B]A[\downarrow z A \text{says} c_A \mid \downarrow y B_1 \text{says} c_B \mid \downarrow z B_2 \text{says} c_B])$$

$$(s)S' = (s)[A[P_A \sigma \pi] \mid s[T_{A'B_2}] \mid Q(T_{A'B_2})] \mid B_1[P_B \sigma \pi] \mid B_2[P_B \sigma \pi]$$

where $\sigma = \{y/x, y, z\}$ (s fresh) and $\pi = \{a/a', b_1/b_1, b_2/b_2, b'_2\}$.

System $S$ is the one already considered in Section 3, where all the participants are willing to advertise their respective contracts to the store $A$, by using a $\text{tell}$ primitive. Firing those prefixes (first 3 steps of the computation above) has the effect of creating corresponding-latent contracts, collected by $A$. At the 4th step, these contracts are fused. Given $\sigma$ and $\pi$ as above, Definition 4.3 is indeed applicable: the domains of $\sigma$ and $\pi$ comply with the definition premises, and as shown in Example 3.7, a system $T_{A'B_2}$ consisting of contracts $c_A \pi$, $c_B \pi$, and $c_B \pi$ may be assigned a global type $S_{A'B_2}$. Hence, a new session $s$ is created, based on $T_{A'B_2}$, plus the queues connecting all pairs of participants. The session variables of the latent contracts being fused (i.e., $x$ for participant $A$, $y$ for $B_1$, and $z$ for $B_2$) are replaced by the fresh session name $s$ in the processes $P_A$, $P_B$, and $P_S$, via $\sigma$. Similarly for participant variables which are replaced by participant names, via $\pi$.

The next example illustrates a case where advertising a contract with both participant names and variables allows for enforcing specific properties on new sessions — in particular, making sure that a certain participant will be involved.

Example 4.6. A seller $S$ wants to sell an item to a specific buyer $B$, via any shipping company that provides a package tracking system. The contract of $S$ is:

$$B!\text{price; }\langle B'?\text{pay}\rangle; c!\text{request; }c'?\text{tracking.B!tracking} + B'?\text{cancel; }c!\text{cancel}$$

saying that the seller $S$ promises to send a price to the buyer $B$. Then, $B$ can either pay, or cancel the transaction: in the first case, $S$ must send a shipping request to a (still unknown) shipping company $c$, who in turn must send back a tracking number, which is then forwarded to $B$; otherwise, if $B$ cancels, then $S$ will also cancel the transaction with $c$. This contract admits an agreement only if the role of the buyer is played by $B$; the role of shipping company $c$, instead, may be played by any participant offering a compliant contract (such as $A'?\text{request; }A!\text{tracking} + A'?\text{cancel}$).

Example 4.7 shows how a participant can use participant variables to enforce some property along different sessions — in particular, to ensure the presence of some participant that was involved in a previous agreement.

Example 4.7. A seller $S$ wants to find a shipping company; then, $S$ wants to use that very same shipping company for every new transaction with some buyer.

The shipping company can be found with a “service discovery” contract such as $c_D = c!\text{find\_ship}$: when a session is established on such a contract, $c$ will be replaced with the name of the participant playing the role of the shipper. For each selling transaction, $S$ advertises a contract of the form $c_S = b!\text{price; }c!\text{address}$, where the seller sends a price to a (still unknown) buyer $b$, and after some interaction (here omitted) it sends a shipping address to the shipping company $c$.

The seller can be implemented with the following CO₂ process, where $K$ is a contract broker:

$$P = (x)(c)\text{tell} \downarrow x c_D \cdot \text{do}^x c_D \cdot \text{find\_ship} \cdot X(c)$$

where $X(c) := (y)(b)\text{[tell} \downarrow y c_S \ldots X(c)]$. (c-related actions omitted)

Note that $c_S$ may be advertised multiple times, using different session variables, inside the recursive process $X$.

An extended version of this example is available in the appendix [17].
4.4. On Culpability

In this section we formalise the notion of culpability in CO₂ systems. Intuitively, a CO₂ participant is “culpable” when it is expected to send or receive some message in a session — and may exculpate itself by performing the required interaction.

We begin by noticing that, according to the CO₂ semantics, each do prefix exposed by a participant, say \( A[P] \), is driven by the contract that \( A \) promised to abide by. In a sense, CO₂ is “culpability-driven”, according to Definition 4.8 below: when a participant is culpable, it has the duty of making the session progress according to its contract. The notion of culpability in CO₂ systems is an extension of the one on contracts (Definition 3.10), covering established sessions.

**Definition 4.8 (Process culpability).** We say that \( A \) is culpable at \( s \) in \( S \) iff \( S \equiv s[T] \mid S' \) and \( A \) is culpable in \( T \).

We say that \( A \) is culpable in \( S \) iff, for some \( s \) and \( u \), \( S \equiv (u)S_0 \) and \( A \) is culpable at \( s \) in \( S_0 \).

According to Definition 4.3 when e.g. \( A \) is culpable at \( s \) in \( S \), then necessarily \( s \in \mathrm{fv}(S) \); furthermore, \( A \) is also culpable in \( (s)S \) (i.e., culpability is preserved when \( s \) is closed and the corresponding session is hidden).

**Example 4.9.** Consider the system \( S' \) from Example 4.5, having the form \( S' \equiv s[T] \mid \cdots \) where \( T = T_{A/B_1B_2} \mid q(T_{A/B_1B_2}) \). We have that, by Definition 3.10, both \( B_1 \) and \( B_2 \) are culpable in \( T \): in fact, \( T \xrightarrow{B_1 \text{ req}} T_{B_1 \equiv A \text{ req}} \). Therefore, by Definition 4.8, we also have that \( B_1 \) and \( B_2 \) are culpable at \( s \) in \( S' \); also, we have that \( B_1 \) and \( B_2 \) are culpable both in \( S' \) and in \( (s)S' \).

A culpable participant can overcome its culpability status by firing its do prefixes, according to \([\text{CO}_2, \text{do}]\), until another participant becomes culpable or session \( s \) terminates. Hence, as long as a culpable participant \( A \) does not enable a do-prefix matching a contractual obligation, \( A \) will remain culpable. Note that when a participant is involved in multiple sessions, it may result culpable in more than one of them.

**Theorem 4.10 (Presence of culpability in CO₂ systems).** Given a CO₂ system \( S \), whenever \( S \rightarrow^* (s)S' \) such that \( S' \equiv s[T] \mid \cdots \) with \( T \neq 0 \), then there exists at least one culpable participant at \( s \) in \( S' \).

**Proof.** By Theorem 3.12 we have that there exists at least one culpable participant in \( T \) — and by Definition 4.8 such a participant is also culpable at \( s \) in \( S' \).

Theorem 4.10 says that in an active session established through a fuse reduction, there is always at least one participant \( A[P] \) who leads the next interaction. Thus, if a corresponding do prefix is not in \( P \), \( S \) may get stuck, and \( A \) is culpable.

4.5. Finer-grained Forms of Agreement

In our framework, the participants firing fuse prefixes are playing the role of contract brokers. In some situations, they may want to guarantee some additional requirements before establishing a new session, in order to obtain finer levels of agreement.

In this section, we discuss some scenarios which can be addressed with simple fuse variations. The first one addresses the (possible) issue of participants playing multiple roles in a contract. The other examples add constraints to the synthesised choreography, in order to enforce global properties on sessions.

**Example 4.11 (Single-role enforcement).** Consider the last example of Section 2, it shows a participant \( B_{12} \) that, by advertising the contract \( c_{B_{12}} \), plays the roles of two buyers in the contract advertised by the seller \( A \). This is allowed since Definition 4.3 is quite liberal with respect to the substitution \( \pi \) mapping participant variables to participant names. However, \( A \) may want to ensure that the two buyers are actually distinct participants: i.e., \( \pi \) must not map two participant variables to the same participant name. Since \( A \) is also a contract broker, it could enforce this requirement through its fuse prefixes, e.g., by adding the following condition to Definition 4.3:

\[
\forall i \in I: \forall a, b \in \mathrm{fv}(c_i): a \neq b \implies \pi(a) \neq \pi(b)
\]

With this restriction in place, \( A \) disregards the possible agreement involving \( c_{B_{12}} \).
Using the same approach as in Example 4.11 (i.e., extending the requirements of Definition 4.3), we can parameterise fuse with a predicate \( \phi \) on global types. Such a predicate is checked on the global type \( S \) (synthesised from the set of latent contracts): when \( \phi(S) \) holds for all \( S \), we have the standard fuse behaviour described in Section 4.3, in general, fuse\(^{\phi} \) can enforce global properties emerging from the composition of individual contracts, as illustrated in Examples 4.12 and 4.13 below.

**Example 4.12 (Avoiding infinite sessions).** A contract broker wants to start new sessions only if they are guaranteed to terminate within a finite number of interactions. This is useful e.g. when a system has a short life expectancy due to an approaching scheduled maintenance, and the brokers are instructed to put limitations on new sessions.

We obtain such a behaviour by letting \( \phi(S) \iff \text{bv}(S) = \emptyset \), where \( \text{bv}(S) \) is the set of bound variables in \( S \). When \( \phi(S) \) holds, we have that new sessions will be created only if there are no recursion variables in \( S \) — and therefore, the choreography does not admit recursive behaviours.

Note that this property cannot be decided by looking at the individual contracts, only. For instance, consider:

\[
A(B e) \mid B(\mu x. (A?e + A?e'x))
\]

Here, B’s contract is recursive; however, A’s contract always chooses its non-recursive branch. Indeed, when such contracts are composed, their global type is \( A \rightarrow B : e \), i.e., a non-recursive choreography.

**Example 4.13 (Requiring infinite sessions).** A contract broker wants to start only non-terminating sessions, where the participants can perpetually interact. To do that, we define \( \phi(S) \) so that it holds only if \( 0 \) is not a sub-term of \( S \). This ensures that the choreography underlying the session does not admit terminating behaviours.

As in Example 4.12, this property needs to be checked at the global (choreography) level; otherwise, contracts with a terminating branch would always be discarded, even when composed with other contracts that always select recursive branches.

In Section 7, we present a case study which takes advantage of the parameterised fuse primitive.

5. Domesticating Wild Choreographies via Honesty

In this section, we discuss the notion of honesty \[^{11}\] providing several examples.

Intuitively, a participant is honest when it is always able to fulfil its contractual obligations, in any context, no matter what other participants do. This is a desirable property in a distributed contract-oriented scenario: in the simplest case, honesty ensures that a participant will never violate its own contracts due to “trivial” bugs, e.g. missing do prefixes (see Example 5.6). In more complex cases, i.e., when a participant A handles multiple sessions simultaneously, honesty ensures that the erratic behaviour of some participants in a session will not cause A to remain stuck, and unable to fulfil its obligations in other sessions (see Example 6.4).

Honesty is especially important when contract-based interactions have some economic relevance: a non-honest (but not necessarily malicious) participant may be unable to reach its goals (i.e., make its contract progress), or may damage other participants if its session obligations remain unfulfilled. Since the state of CO2 sessions allows to single out which participants are responsible for session advancement (by Definition 5.10 and Theorem 4.10), such a state could be used as “proof” for sanctioning contractual violations, possibly in an automatic way. In this scenario, a non-honest participant A may be vulnerable to attacks: other participants may induce a situation in which A cannot fulfil its contract, causing it to be punished. For all these reasons, before deploying a service, its developers might want to ensure that its implementation is honest.

Formally, the definition of honesty relies on that of contract ready sets (Definition 5.1) and process ready sets (Definition 5.3).
Definition 5.1 (Contract Ready Sets). For a contract $c$, we define the set $\text{CRS}(c)$ as follows:

$$
\text{CRS}(c) = \begin{cases} 
\text{CRS}(c') & \text{if } c = \mu x . c' \\
\{ (A, e_i) \mid i \in I \} & \text{if } c = \sum_{i \in I} A? e_i; c_i \\
\{ (A_1, e_i) \mid i \in I \} & \text{if } c = \bigoplus_{i \in I} A_1! e_i; c_i \text{ and } I \neq \emptyset
\end{cases}
$$

and we call each element of $\text{CRS}(c)$ a contract ready set of $c$.

Intuitively, contract ready sets represent the immediate obligations in a contract: if $A$ advertises $c$, then $A$’s process should also offer the interaction capabilities of one of the ready sets in $\text{CRS}(c)$. Each interaction is a pair $(A, e)$, where $A$ is the other party involved in the interaction, and $e$ is the sort of the exchanged message. An external choice results in a single ready set, meaning that $A$ is expected to offer all the interactions therein; instead, an internal choice results in one singleton ready set for each branch, meaning that $A$ should pick one of its choice and offer the corresponding interaction (cf. Example 5.2 below). Note that Definition 5.1 above extends the notion of ready set in [11] (where only bilateral sessions are considered) to suit our multiparty contract model.

Example 5.2. Consider the system of contracts $T_{AB_1B_2}$ from Example 4.7 and in particular, the stipulated contracts therein, using the substitution $\pi = \{a/b, a'/b, a''/b, b'/b, b''/b \}$ from Example 4.3:

\[
\begin{align*}
\tilde{c}_A &= c_A \pi = B_1?q; B_2?q; B_2!quote; (B_1?q; order; B_2!ok + B_1?q; bye; B_2!bye) \\
\tilde{c}_{B_1} &= c_{B_1} \pi = A?quote; (B_2!ok; A?order \oplus B_2!bye; A!bye) \\
\tilde{c}_{B_2} &= c_{B_2} \pi = A?quote; (B_1?q; A?ok + B_1?q; bye; A!bye)
\end{align*}
\]

We have $\text{CRS}(\tilde{c}_A) = \{ (B_1, \text{req}) \}$: in other words, at this point of the contract, an interaction is expected between $A$ and $B_1$ (since $A$ is waiting for $\text{req}$), while no interaction is expected between $A$ and $B_2$. Furthermore, since we just have a single ready set, $A$ does not have other interactions to choose from.

Let us now equip $T_{AB_1B_2}$ with one queue between each pair of participants, and let it perform the request exchange between $B_1$ and $A$, with the transitions:

$$
T_{AB_1B_2} \mid Q(T_{AB_1B_2}) \xrightarrow{B_1\to A \text{ req}} A\leftarrow B_1 \text{ req} \xrightarrow{T'_{AB_1B_2}} T_{AB_1B_2}' \mid Q(T_{AB_1B_2})
$$

We have that $\tilde{c}_A$ in $T'_{AB_1B_2}$ is now reduced to:

$$
\tilde{c}'_A = B_2?q; B_2!quote; (B_1?q; order; B_2!ok + B_1?q; bye; B_2!bye)
$$

and thus we have $\text{CRS}(\tilde{c}'_A) = \{ (B_2, \text{req}) \}$, i.e., $A$ is now waiting for a request from $B_2$.

If we let the system reduce further, $\tilde{c}'_A$ reaches its internal choice:

$$
\tilde{c}''_A = B_1?q; order; B_2!ok + B_1?q; bye; B_2!bye
$$

We now have a single ready set with two elements: $\text{CRS}(\tilde{c}''_A) = \{ (B_1, \text{order}) \}$, i.e., $A$ must handle both answers from $B_2$. Instead, when $\tilde{c}'_{B_1}$ reduces to its internal choice, we have:

$$
\tilde{c}'_{B_1} = B_2!ok; A\text{order}\oplus B_2!bye; A!bye
$$

We now have two ready sets, each one containing one interaction: $\text{CRS}(\tilde{c}'_{B_1}) = \{ (B_2, \text{ok}) \}, \{ (B_2, \text{bye}) \}$. $B_1$ is expected to fulfill either of them, and is free to choose which one.

Example 5.2 shows that, when a contract $c$ of a participant $A$ evolves within a system $T$, its ready sets change accordingly. Now we need to define the counterpart of contract ready sets for CO$_2$ processes, i.e., the process ready sets.
Definition 5.3 (Process Ready Set). For all CO₂ systems $S$, all participants $A$ and sessions $s$, we define the set of pairs 

$$\text{PRS}_A^s(S) = \{(B, e) \mid \exists s', a : S \equiv (s', a) (A|do^*_e \cdot \cdots + \cdot \cdots) \land s \notin \vec{v}\}$$

Intuitively, Definition 5.3 says that the process ready set of $A$ over a session $s$ in a system $S$ contains the interactions that $A$ is immediately willing to perform with other participants through its $do^*_e$ prefixes. Just like contract ready sets, the interactions are represented by participant/sort pairs. Definition 5.3 is adapted from [18] to suit our multiparty contract model.

Next, we want to characterise a weaker notion of the process ready set, so that it only takes into account interactions that, since Definition 5.4 builds upon Definition 5.3, if $s \in \text{WPRS}_A^s(S)$ then necessarily $\text{WPRS}_A^s(S) = \emptyset$: i.e., observing the weak process ready set of a session $s$ is only meaningful after $s$ has been established, and it is not delimited (and thus it cannot be $a$-converted).

We now introduce the final ingredient required to define honesty, that is the readiness of a participant.

Definition 5.4 (Weak Process Ready Set). We write $S \not\equiv (A \cdot do^*_e) \rightarrow S'$ iff:

$$\exists B, P : S \xrightarrow{B, P} S' \land (A \neq B \lor (\forall v : P \equiv do^*_e \iff s \neq t))$$

We then define the set of pairs $\text{WPRS}_A^s(S)$ as:

$$\text{WPRS}_A^s(S) = \left\{(B, e) \mid \exists S' : S \not\equiv (A \cdot do^*_e) \rightarrow S' \land (B, e) \in \text{PRS}_A^s(S')\right\}$$

In Definition 5.4, the transition relation $\not\equiv (A \cdot do^*_e) \rightarrow$ abstracts from the reductions of $S$ which are not due to $A$’s actions on session $s$. In other words, such a relation allows a system $S$ with a participant $A[P]$ to reduce freely, until a reducible $S'$ containing $A[P]$ is reached, such that a $do^*_e$ prefix is at top-level in $P'$. Note that, since Definition 5.4 builds upon Definition 5.3 if $s \notin \text{fv}(S)$, then necessarily $\text{WPRS}_A^s(S) = \emptyset$: i.e., observing the weak process ready set of a session $s$ is only meaningful after $s$ has been established, and it is not delimited (and thus it cannot be $a$-converted).

Definition 5.5 (Readiness). We say that $A$ is ready in $S$ iff, whenever $S \equiv (\bar{u}(\bar{v})|S_0$ for some $\bar{u}, \bar{v}$ and $S_0 \equiv s[A(c) | \cdots | \cdots$, the following holds:

$$\exists X \in \text{CRS}(e) : X \subseteq \text{WPRS}_A^s(S_0)$$

Definition 5.5 says that a participant $A$ is ready in a system $S$ when, for all its established sessions $s$, $S$ can reduce in such a way that $A$’s process will provide at $s$ all the interactions expected by one of the ready sets of $A$’s contract at $s$. In other words, when $A$ is “ready”, then, for all its contracts $c$ already involved in a session, $A$ has a way to fulfil at least its immediate obligations in $c$.

Example 5.6. Consider the system $S$ in Example 4.5. We have seen that, after its latent contracts have been fused, we obtain:

$$(s)S' \equiv (s)(A[P_A \sigma] | S[T_{AB1:B2} | Q(T_{AB1:B2})] | B_1[P_B \sigma] | B_2[P_B \sigma])$$

where the substitutions $\sigma$ and $\pi$ are also from Example 4.5. Let us define the processes (after substitutions):

- $P_A \sigma \pi = do^*_a \cdot req \cdot do^*_b \cdot req \cdot do^*_c \cdot \text{quote} \cdot (do^*_o \cdot \text{order} \cdot do^*_o \cdot \text{ok} + do^*_o \cdot \text{bye} \cdot do^*_b \cdot \text{bye})$
- $P_B \sigma \pi = \tau \cdot do^*_a \cdot \text{req} \cdot do^*_c \cdot \text{quote} \cdot do^*_c \cdot \text{order}$
- $P_B \sigma \pi = do^*_a \cdot \text{req} \cdot (do^*_o \cdot \text{ok} + do^*_b \cdot \text{bye} \cdot do^*_a \cdot \text{bye})$

The side condition “$s \notin \vec{v}$” of Definition 5.3 deals with cases like $S_0 = (s)(A[do^*_a \cdot \text{int}])$ and $S = S_0 | s[A[B \cdot \text{int}] | \cdots | \cdots$; without the side condition, $\text{PRS}_A^s(S_0) = \{(B, \vec{v})\}$ — hence, by Definition 5.3, $A$ would result to be ready in $S$. 

17
We have:
\[
\begin{align*}
\text{PRS}_A^*(S') & = \{\text{B}_1, \text{req}\} = \text{WPRS}_A^*(S') \\
\text{PRS}_B^*(S') & = \varnothing \neq \{\text{A}, \text{req}\} = \text{WPRS}_B^*(S') \\
\text{PRS}_B^*(S') & = \{\text{A}, \text{req}\} = \text{WPRS}_B^*(S')
\end{align*}
\]

Note that the \(\tau\) prefix in \(P_{B_1}\) prevents \(B_1\) from interacting immediately with \(A\) on session \(s\), although it is “weakly ready” to do so. Hence, considering that the weak process ready sets of each participant in \(S'_1\) match their respective contract ready sets in \(T_{AB,B_2}\) (Example 8.3) according to Definition 5.5 we have that participants \(A, B_1\) and \(B_2\) are all ready in \((s)S'_1\).

Before defining honesty formally, we need to characterise the class of systems for which this concept is meaningful, — i.e., those systems where a participant is not (yet) involved in latent contracts nor active sessions, and whose process has no free variables.

**Definition 5.7 (Initial System).** A \(CO_2\) system \(S\) is \(A\)-initial if \(S\) has no sub-term of the form \(\downarrow A\) says \(c\) or \(A(c)\) with \(c \neq 0\), and whenever \(S\) has a sub-term of the form \(A[P]\), then \(P\) is closed. A \(CO_2\) system \(S\) is initial when it is \(A\)-initial for all participants \(A\) in \(S\).

**Definition 5.8 (Honesty).** We say that \(A[P]\) is honest iff, for all \(A\)-initial \(S \equiv A[P] \upharpoonright S_0\) s.t. \(S \rightarrow^* S'\), \(A\) is ready in \(S'\).

A process \(A[P]\) is said to be honest when, for all contexts and reductions that \(A[P]\) may be engaged in, \(A\) is persistently ready. In other words, for all context \(S_0\) with which \(A[P]\) may be composed, and for all systems \(S'\) resulting from the continuation of such a composition, there must be a correspondence between the interactions required by \(A\)'s stipulated contracts and \(A\)'s (weak) process ready set. The definition requires an \(A\)-initial system to rule out contexts with latent/stipulated contracts of \(A\) — otherwise, \(A\) could be made trivially dishonest, e.g. by inserting a latent contract \(\downarrow_A\) says \(c\) that \(A\) cannot fulfi. Notice that \(A[P]\) is trivially honest when \(P\) advertises no contracts.

**Example 5.9.** Consider the process \(B_1[\text{tell} \downarrow y \ c_{B_1} \ . \ P_{B_1}]\) of system \(S\), as defined in Examples 4.5 and 5.6. We show that this process is not honest. In fact, \(S\) can reduce as \(S \rightarrow^* (s)S' \rightarrow^* (s)S''\), where:
\[
(s)S'' = (s)\left(A[\text{do}^*_B \ . \ \text{ok} \ . \ \text{do}^*_B \ . \ \text{bye} \ . \ \text{do}^*_B \ . \ \text{bye}\right] \\
\quad \mid s[A(B_1[\text{order:B}_2\text{ok} + B_1?\text{bye:B}_2\text{bye})] \\\n\quad \quad \mid B_1(B_2[\text{ok:A}\text{order} \oplus B_2[\text{bye:A}\text{bye}]) \\\n\quad \mid B_2(B_1[\text{ok:A}\text{ok} + B_1?\text{bye:A}\text{bye}]) \]
\]
At this point, we see that there is a problem in the implementation of \(B_1\): it does not notify the other buyer before making an order. In fact, \(B_1\)'s process is trying to perform \(\text{do}^*_B\text{order}\), but its contract requires that \(\text{do}^*_B\text{ok}\) is performed first (or \(\text{do}^*_B\text{bye}\), if the quote is rejected). This is reflected by the mismatch between \(B_1\)'s process ready set in \(S''\), and its contract ready sets, in session \(s\):
\[
\text{PRS}_B^*(S'') = \{\{A, \text{order}\}\} \\
\text{CRS}(B_2[\text{ok:A}\text{order} \oplus B_2[\text{bye:A}\text{bye}]) = \{\{B_2, \text{ok}\}, \{\{B_2, \text{bye}\}\}
\]
In terms of the definitions above, we can say that there exists a system \(S\) — containing \(B_1[\text{tell} \downarrow y \ c_{B_1} \ . \ P_{B_1}]\) — that reduces to a \((s)S''\) where \(B_1\) is not ready (Definition 5.5). Therefore, \(B_1[\text{tell} \downarrow y \ c_{B_1} \ . \ P_{B_1}]\) is not honest. In fact, \(B_1\) is culpable in \((s)S''\), according to Definition 4.8.

As in [18], the definition of honesty subsumes a fair scheduler, eventually allowing participants to fire persistently (weakly) enabled \(\text{do}\) actions.
6. Properties of CO₂ Systems

We now give the main properties of our framework. We start by proving that honest participants can always perform their contractual obligations without relying on the context (Theorem 6.1). Then, we formalise a link between the honesty of participants, and two key properties borrowed from the session types literature: Theorem 6.2 introduces session fidelity in CO₂, while Corollary 6.3 (descending from Theorem 6.1) introduces a notion of progress of CO₂ systems.

Theorem 6.1 (Session progress). Let $S_0$ be an k-initial CO₂ system with $\mathbb{A}[P_0]$ honest. If $S_0 \rightarrow^* (s)S$, where $S \equiv s[T] \ | \ \mathbb{A}[P] \ | \cdots \ | \mathbb{A}$ is culpable at $s$ in $S$, then:

$$S \xrightarrow{\mathbb{A}[p_1]} \cdots \xrightarrow{\mathbb{A}[p_n]} \mathbb{A}[\mathsf{do}^* e]$$

for some $\mathbb{B}$, $e$, and $p_1, \ldots, p_n$ such that $n \geq 0$, and $p_i \neq \mathsf{fuse}$ for all $i \in 1..n$.

Proof. By assumption, $\mathbb{A}$ is culpable at $s$ in $S$. By Definition 5.8, $\mathbb{A}$ is ready in any reduction of $S_0$ — in particular in $(s)S$. Therefore, by Definition 5.5, there exists $X \in \mathsf{CRS}(e)$ such that $X \subseteq \mathsf{WPRS}_2^*(S)$. From these premises, we have that there must exist some $\mathbb{B}, e$ such that $T \xrightarrow{\mathbb{A}[\mathbb{B}, e]}$ (due to culpability) and $(\mathbb{B}, e) \in \mathsf{WPRS}_2^*(S)$ (due to readiness). By Definition 5.4, it means:

$$\exists S' : S \not\equiv (\mathbb{A}[\mathsf{do}^* e \cdot \mathsf{P}^* + \mathsf{P}' | \mathsf{Q}]) \land (\mathbb{B}, e) \in \mathsf{PRS}_2^*(S')$$

and therefore, by Definition 5.3, we must have that $\mathbb{A}[\mathsf{do}^* e \cdot \mathsf{P}'] \in \mathsf{PRS}_2^*(S')$. Thus, we have to show that $\mathbb{A}[\mathsf{P}]$ in $S$ may reduce to such a process, with $\mathsf{do}^* e$ enabled, by firing its own prefixes $p_1, \ldots, p_n$ only — except $\mathsf{fuse}$.

By contradiction, assume that none of the traces from $S$ to $S'$ (following (6.1)) satisfies the thesis. We show that, in this case, $\mathbb{A}[P_0]$ would not be honest: in fact, there would exist a system $\bar{S}_0 \equiv \mathbb{A}[P_0] \ | \cdots$ with a reduct $(s)\bar{S}$ such that $\bar{S} \equiv s[T] \ | \mathbb{A}[P] \ | \cdots$, where $\mathbb{A}$ would not be ready. Our goal is to create such a system $\bar{S}_0$.

Before proceeding, we define some auxiliary functions.

Given two CO₂ systems $S_A, S_B$ such that there exists a transition $S_A \xrightarrow{\ell} S_B$, $\mathsf{inLab}(S_A, S_B)$ gives the “innermost” label in the derivation inducing such a transition. Intuitively, it pierces through delimitations, and returns the underlying CO₂ prefix before it is turned into $\tau$ by an application of rule $[\mathsf{CO₂-Del}]$:

$$\mathsf{inLab}(S_A, S_B) = \begin{cases} \ell_0 & \text{if } [\mathsf{R}] \quad S_A \xrightarrow{\ell_0} S_B \quad \text{and } [\mathsf{R}] \not\in \{[\mathsf{CO₂-Del}], [\mathsf{CO₂-Par}]\} \\ \mathsf{inLab}(S_A', S_B') & \text{if } [\mathsf{R}] \quad S_A \xrightarrow{\ell_0} S_B \quad \text{and } [\mathsf{R}] \in \{[\mathsf{CO₂-Del}], [\mathsf{CO₂-Par}]\} \end{cases}$$

Given a sequence of CO₂ systems $\vec{S}_0$ such that each element is a reduct of the preceding one, $\mathsf{inTr}(\vec{S}_0)$ gives the trace composed by the “innermost” labels of each reduction:

$$\mathsf{inTr}(S_A \cdot \epsilon) = \epsilon \quad \text{(where } \epsilon \text{ is the empty sequence)}$$

$$\mathsf{inTr}(S_A \cdot S_B \cdot \vec{S}_0) = \mathsf{inLab}(S_A, S_B) \cdot \mathsf{inTr}(S_B \cdot \vec{S}_0)$$

19
Let us now consider an execution going from the A-initial $S_0$ to $(s)S$ in the hypotheses, and let $\tilde{S}$ be the sequence of CO$_2$ system configurations composing such an execution; furthermore, let us define the “inner” trace $\tilde{T} = \text{inTr}(\tilde{S})$.

With this information at hand, we can build $\tilde{S}_0$ so that:

1. $\tilde{S}_0 = A[P_0] | \tilde{S}_0'$;
2. $\tilde{S}_0'$ contains:
   
   (a) all the sessions appearing in $S$;
   
   (b) all the latent contracts appearing in $S_0$, except those that are not fused along $\tilde{S}$ and thus appear in $(s)S$;
   
   (c) participants $B_1[P_1], \ldots, B_n[P_n]$, where $B_1, \ldots, B_n \neq A$ also appear in $S$, and $P_1, \ldots, P_n$ are written so that:

   i. all the “innermost” prefixes in $\tilde{T}$ fired by $B_1, \ldots, B_n$ can still be fired in the same order — with the exception of the tell prefixes corresponding to latent contracts in $(s)S$, which are neglected;

   ii. after the prefixes in $\tilde{T}$ are fired, all $B_1[P_1], \ldots, B_n[P_n]$ reduce to $B_1[0], \ldots, B_n[0]$.

As a result, we have that $\tilde{S}_0$ defined above can reduce to a system $(s)\tilde{S}$ such that $\tilde{S} \equiv s[T] | A[P] | \cdots$ through a “inner” trace $\tilde{T}$ whose prefixes match those in $\tilde{T}$, except for the lack of tell prefixes fired by any $B \neq A$ (by item 2(c)i above). Furthermore, we have that:

1. no latent contracts appear in $\tilde{S}$ (unless they have been advertised by $A$);
2. all established sessions in $S$ also appear in $\tilde{S}$, and vice versa;
3. except for $A[P]$, all participants in $\tilde{S}$ have the form $B[0]$.

Hence, no participant in $\tilde{S}$ can fire a fuse prefix, and no participant $B \neq A$ in $\tilde{S}$ has any prefix enabled. Now, recall that we assumed (by contradiction) that Equation (6.1) cannot be satisfied with a trace consisting only of $A[P]$’s prefixes (except fuse). Therefore:

\[
\not\!\tilde{S}': \tilde{S} \not\!{(A : \text{do}^*)^*} \tilde{S}' \land (B, e) \in \text{PRS}_\pi^A(\tilde{S}')
\]

But then, there exists an A-initial system $\tilde{S}_0 \equiv A[P_0] | \cdots$ such that $\tilde{S}_0 \rightarrow^* (s)\tilde{S}$, and $A$ is not ready in $(s)\tilde{S}$. Therefore, we must conclude that $A[P]$ is not honest — thus violating the theorem hypothesis.

Theorem 6.2 below says that each (honest) participant will strictly adhere to its contracts, once they have been fused in a session $s$: no extra-contractual interactions are possible on $s$, and honest participants will eventually perform do$^*$ actions, as long as it is required by their contracts.

**Theorem 6.2 (Session fidelity).** Given a CO$_2$ system $S_0$ containing $A[P_0]$, whenever $S \rightarrow^* (s)S$ with $S \equiv s[T] | \cdots$, then:

1. $S \not\!{(A : \text{do}^*)} \not\!{(A : \text{do} e)} \Longrightarrow \exists n \geq 0: T \lambda_1 \rightarrow \cdots \lambda_n \rightarrow^{A = B : e} \text{ where } \forall C, e', i = 1..n: \lambda_i \neq A \Rightarrow C : e';$

2. If $S$ is A-initial, $A[P_0]$ honest, and $\mathcal{Y} = \{ (B, e) | T \lambda_i \rightarrow^{A = B : e} \} \neq \emptyset$, then

\[
\exists (B, e) \in \mathcal{Y} \text{ such that } S \not\!{(A : \text{do}^*)} \not\!{(A : \text{do} e)}
\]

Recall that, as in Definition 5.4, \not\!{(A : \text{do}^*)} intuitively stands for any reduction not involving $A$ on session $s$. 20
Proof. \textit{Item 1} follows directly from the semantics of CO$_2$ (in particular, rule [CO$_2$:do]) if A is able to fire a do$_B$e prefix, then its contract must necessarily allow for sending/receiving e to/from B.

\textit{Item 2} follows from the honesty hypothesis for A[P], and from Definition 5.5 (readiness). Let us fix $n \geq 1$ enumerating all enabled interactions in T driven by A — i.e., let n be the cardinality of Y in the theorem statement. Let us consider $e = T(A)$, i.e., A’s contract in T. Due to A’s readiness, e is such that, either:

1. $\text{CRS}(e) = \{((B_j,e_j))\}_{j=1..n}$; or,

2. $\text{CRS}(e) = \{((B_j,e_j))\}_{j=1..n}$.

where $\{((B_j,e_j))\}_{j=1..n} = Y$ — i.e., Y contains exactly the (B,e) pairs appearing in CRS(e). Recall that, since $S_0$ is $\mathcal{A}$-initial and A is honest, then A must be \textit{ready} in all possible reducts of $S_0$ — and, in particular, it must be ready in (s)(S). Therefore, we have that $\exists X \in \text{CRS}(e)$ such that $S' \xrightarrow{\text{do}_B \text{e}} \xrightarrow{\text{A:do}_B \text{e}} (S')$ (6.2)

Since the (B,e) pair satisfying Equation 6.2 also belongs to Y, the thesis holds. \hfill \Box

Note that, by Theorem 6.1, the conclusion of Item 2 of Theorem 6.2 above may be stricter: i.e., the do$_B$e transition may be reached by firing A’s prefixes only (excluding fuse).

Corollary 6.3 below introduces a notion of “global progress” in CO$_2$. This notion is only meaningful after a session has been established, and thus a culpable participant exists. In fact, a system without sessions may not progress — either because a set of compliant contracts cannot be found, or because a fuse prefix is not enabled. In both cases, no participant may be deemed culpable, and thus responsible for the next move. However, the system could progress again if it is composed with another system consisting of (honest) participants, allowing for a new session to be established.

**Corollary 6.3.** Given an initial CO$_2$ system $S_0$ with only honest participants, if $S_0 \rightarrow^* S \equiv (s(T)|\cdots)$ with $T \neq 0$, then $S \rightarrow^*$.

\textit{Proof.} By contradiction. From [6], we know that T is deadlock free, thus if the CO$_2$ system cannot make further reductions, it must be because one of the participant A cannot meet its obligations. This is in contradiction with Theorem 6.1 since all participants are honest in $S_0$. \hfill \Box

This result also holds for systems where a process takes part in multiple sessions: the honesty of all participants guarantees that all sessions will (eventually) progress.

**Example 6.4.** We now give a simple example of a system with multiple sessions. We show how a seemingly honest process B could be deemed culpable due to the unexpected behaviour of other participants, and how honest participants guarantee progress of the whole system. Consider:

\[ S = (x,y,z,w) \begin{array}{l}
A[\text{tell}_A \downarrow_{x} (B\text{int}) \cdot \text{fuse} \cdot \text{fuse}] \\
| B[\text{tell}_B \downarrow_{y} (A?\text{int}) \cdot \text{tell}_C \downarrow_{z} (C\text{bool}) \cdot \text{do}_A^y \text{int} \cdot \text{do}_C^z \text{bool}] \\
| C[\text{tell}_C \downarrow_{w} (B?\text{bool}) \cdot \text{do}_B^w \text{bool}] 
\end{array} \]

After all four contracts have been advertised to A and fused, the system reduces to:

\[ S' = (s_1,s_2) \begin{array}{l}
A[0] \\
| B[\text{do}_A^x \text{int} \cdot \text{do}_C^z \text{bool}] \\
| s_1(A(B?\text{int}) \cdot B(A?\text{int})) \\
| s_2(B(C?\text{bool}) \cdot C(B?\text{bool})) 
\end{array} \]

Even if both sessions $s_1$ and $s_2$ enjoy contractual progress, $S'$ is stuck: A does not perform the promised action, thus remaining culpable in $s_1$; B is stuck waiting in $s_1$, thus remaining culpable in $s_2$. Indeed,

\footnote{In this case, B is deemed culpable in $s_2$ because its implementation did not expect A to misbehave.}
neither A nor B are ready in $S'$, and thus they are not honest in $S$. Hence, global progress is not guaranteed.

Let us now consider the following variant of $S$, where all participants are honest:

$$\hat{S} = (x, y, z, w) \left( A[tell_{x, z} \langle B!\text{int} \rangle \cdot do^x_{\text{int}}] \mid \text{fuse} \mid \text{fuse} \mid \text{fuse} \right)$$

$$\mid C[tell_{y, w} \langle B?'\text{bool} \rangle \cdot do^w_{\text{bool}}]$$

$$\mid B[tell_{y, x} \langle A?'\text{int} \rangle \cdot tell_{x, z} \langle C\text{bool} \rangle \cdot (do^x_{\text{int}} \cdot do^y_{\text{bool}} \cdot fuse \cdot \tau \cdot (do^x_{\text{int}} \cdot do^y_{\text{bool}})) \right)$$

In this case, A will respect its contractual duties, while B will be ready to fulfil its contracts on both sessions — even if one is not activated, or remains stuck (here, $\tau$ represents an internal action, e.g., a timeout: if the first $do^y_{\text{int}}$ cannot reduce, B falls back to running the two sessions in parallel). The honesty of all participants in $\hat{S}$ guarantees that, once a session is active, it will reach its completion.

### 7. Case Study: Sharing Knowledge

We illustrate the suitability of our framework to reason about distributed systems where contract brokers are willing to offer specific guarantees to their clients. In particular, we illustrate how three key components of our framework, i.e., (i) multiparty sessions, (ii) the notion of honesty, and (iii) contractual agreements based on choreographies, address complex challenges encountered in the design and specification of distributed systems.

We present a case study where a contract broker wants to provide a specific service to its clients: guaranteeing that participants will interact according to a class of gossip protocols [15]. Gossip protocols are used to spread a piece of information among a group of participants, ensuring that each one of them will eventually acquire such information. In particular, we focus on "knowledge sharing" protocols, where each participant owns some initial knowledge, which must be shared across all participants before the participants terminate, i.e., everyone must eventually receive each other’s initial knowledge.

For instance, assuming that $k$ is a message sort representing the current knowledge of a participant, the system of contracts on the left below implements a sharing knowledge protocol, while the one on the right does not. Indeed, on the right-hand side, A does not acquire B’s knowledge.

$$A(Bk; B\text{?}k) \mid B(A\text{?}k; A\text{!}k) \quad \checkmark \quad A(Bk) \mid B(A\text{!}k) \quad \times$$

We distinguish three challenges that one needs to address to ensure that a broker provides effectively a knowledge sharing service. (a) Intuitively, for a set of participants to share knowledge, they may interact either via a single (multiparty) session or via several binary sessions. However, guaranteeing safety and liveness properties across several, possibly interleaved, sessions is a notoriously difficult problem (see [19] for instance) and it is even harder to check for semantic properties such as knowledge sharing. Instead, in our framework, i.e., (i) multiparty sessions, (ii) the notion of honesty, and (iii) contractual agreements based on choreographies, address complex challenges encountered in the design and specification of distributed systems. Finally, in Section 7.3, we introduce a variant of the parameterised fuse primitive (cf. Section 4.5) which guarantees that new sessions satisfy the knowledge sharing property. For the sake of readability, some auxiliary technical details are relegated to the appendix [17].
7.1. Semantic knowledge sharing

We introduce a formalization of knowledge transmission and acquisition among a set of participants through labelled transition systems (LTS), which may be induced by both system of contracts and CO2 systems.

Definition 7.1 (Knowledge of a participant). The knowledge of a participant is a pair \((A, Q)\), where \(A\) (acquired knowledge) is a set of participant names, and \(Q\) (pending knowledge) is a map from participant names to queues of the form \(A' \times \ldots \times A''\), where \(A', A''\) are sets of participant names. We write \(\emptyset\) when \(Q\) maps each participant to the empty queue.

If the knowledge of \(B\) is \((\{B\}, \emptyset)\), then \(B\) has only its own initial knowledge. The knowledge of \(B\) may increase as it interacts with other participant. For instance, if the knowledge of \(B\) is \((\{B, E\}, \{A: \{A\}, C: \{C\}, D: \{C, D\}\})\); then \(B\) has acquired \(E\)’s initial knowledge, while some knowledge is pending, i.e., the initial knowledge of \(A\) (transmitted by \(A\) itself) and the initial knowledge of \(C\), followed by \(C\)’s initial knowledge and \(D\)’s initial knowledge.

We introduce an LTS for knowledge transmission/acquisition in Definition 7.2 below. We use two operators: \(\text{push}_B((A, Q), (A', Q'))\) which augments the knowledge of a participant with the knowledge transmitted by \(B\), i.e., the knowledge \(A'\) sent by \(B\) is appended to \(Q\); and \(\text{pop}_B(A, Q)\) moves the first pending knowledge sent by \(B\) to the acquired knowledge \(A\). For instance, we have:

\[
\text{push}_B(\{\{A\}, \{C: \{C\}, B: \{B\}\}\}, \{\{A, B\}, \{D: \{D\}\}\}) = \{\{A\}, \{C: \{C\}, B: \{B\}, \{A, B\}\}\}
\]

\[
\text{pop}_B(\{\{A\}, \{B: \{B\}\}, \{C: \{C\}\}\}) = \{\{A, B\}, \{C: \{C\}\}\}
\]

Observe that push ignores the pending knowledge \(Q'\) of the second parameter. The formal definition of these operators is given in the appendix [17].

Definition 7.2 (Knowledge LTS). An LTS \(L = (\Sigma \times \Delta, \Lambda, \rightarrow, k_0^\delta)\) is a knowledge LTS (abbreviated KLTS) iff:

1. \(\Sigma\) is a set of total mappings from participants names to their respective knowledge (i.e., \((A, Q)\) pairs by Definition 7.1);
2. \(\Delta\) is any set, called decorating set, with elements ranged over by \(\delta, \delta'\); etc. For brevity, we write \(k^\delta\) instead of \((k, \delta) \in \Sigma \times \Delta;\)
3. \(k_0^\delta\) is the initial state, such that for all \(A \in P: k_0(A) = (\{A\}, \emptyset);\)
4. \(\Lambda\) contains labels of the form \(B!A, A?B, \tau;\)
5. the labelled transition relation \(\rightarrow \subseteq (\Sigma \times \Delta) \times \Lambda \times (\Sigma \times \Delta)\) has the following properties, for all \(k_1^{\delta_1}, k_2^{\delta_2} \in \Sigma \times \Delta:\)
   (a) \(k_1^{\delta_1} \xrightarrow{B!A} k_2^{\delta_2} \implies k_2 = k_1\{\text{push}(k_1(A), k_1(B))/A\};\)
   (b) \(k_1^{\delta_1} \xrightarrow{A?B} k_2^{\delta_2} \implies k_2 = k_1\{\text{pop}(k_1(A))/A\};\)
   (c) \(k_1^{\delta} \xrightarrow{\tau} k_2^{\delta_2} \implies k_1 = k_2.\)

Item 1 says that each state of a KLTS maps a participant to its current knowledge, e.g., \(k^\delta(A) = (\{A, B\}, \{D: \{C, D\}\})\). The decorations in item 2 allows to distinguish states with a same knowledge mapping. By item 3 each participant has only its own knowledge in the initial state. The transition labels in 4 represent either knowledge transmission \((B!A)\), knowledge acquisition \((A?B)\), or and silent moves \(\tau\). Transitions and knowledge mapping are related as follows

- by \(\text{push}_B\), when \(B\) sends its knowledge to \(A\), \(A\)’s knowledge is updated by pushing \(B\)’s knowledge to the corresponding (pending) knowledge queue.
• by \([5b]\), when \(A\) acquires the knowledge from \(B\), \(A\)'s knowledge is updated by dequeuing the first element of the knowledge transmitted by \(B\);

• by \([5c]\), silent moves have no effects on \(k(\cdot)\).

Hereafter, we often write \(k, k_1, \ldots\) for KLTS states (omitting decorations \(\delta\) when unimportant).

**Definition 7.3 (Knowledge sharing).** Let \(L = (\Sigma, A, \rightarrow, k_0)\) be a KLTS. We say that a state \(k^\delta \in \Sigma \times \Delta\) shares knowledge among participants \(A \subseteq \mathcal{P}(T)\) iff, for all \(B \in A, \exists A'\ s.t. k(B) = (A', \Omega)\) and \(A \subseteq A'\).

We say that \(L\) shares knowledge among \(A\) iff, whenever \(k_0^\delta \rightarrow^* k_1^\delta\), then there exists \(k_2^\delta\) such that \(k_1^\delta \rightarrow^* k_2^\delta\) and \(k_2^\delta\) shares knowledge among \(A\).

A KLTS state \(k^\delta\) shares knowledge among a set of participants \(A\) when, in this state, each participant has acquired the knowledge of everybody else in \(A\). \(L\) shares knowledge among \(A\) iff all traces in a KLTS \(L\) eventually reach a state that shares knowledge among \(A\).

**Example 7.4.** Let \(L_1\) be a KLTS with a unique trace \(k_0 \xrightarrow{A!k} k_1 \xrightarrow{B?k} k_2\): \(L_1\) does not share knowledge among \(A\) and \(B\). Let \(L_2\) be a KLTS with a unique trace \(k_0 \xrightarrow{A!k} k_1 \xrightarrow{B?k} k_2 \xrightarrow{B!k} k_3 \xrightarrow{A!k} k_4\): \(L_2\) shares knowledge among \(A\) and \(B\).

### 7.2 Knowledge sharing in contracts and CO2

A KLTS \(L^T\) can be easily induced from the semantics of a systems of contracts \(T\). Given a system of contracts \(T\), whenever \(A \in \mathcal{P}(T)\) receives the special “knowledge message” \(k\) from \(B \in \mathcal{P}(T)\), the knowledge of \(A\) is augmented with that of \(B\). Subsequently, if \(A\) sends \(k\) to \(C \in \mathcal{P}(T)\), then \(C\) will receive the knowledge of both \(A\) and \(B\). Technically, each time \(k\) is sent (resp. received) by a participant in \(T\), a corresponding event is induced in \(L^T\), so that knowledge is transmitted (resp. acquired). Instead, if \(T\) reduces by exchanging a message sort \(e \neq k\), then a \(\tau\) transition is induced in \(L^T\). Each configuration of the system of contracts is used as decoration in \(L^T\), so that the LTS of \(T\) and \(L^T\) are isomorphic.

**Example 7.5.** Let \(T_1 = A(B!k) / B(A?k)\). \(L^{T_1}\) does not share knowledge among \(A\) and \(B\), and thus \(T_1\) does not share knowledge among \(A\) and \(B\). Instead, \(T_2 = A(B!k; B?k) / B(A?k; A!k)\) shares knowledge among \(A\) and \(B\).

Analogously, a KLTS \(L^S\) can be induced from the semantics of a CO2 system \(S\). Intuitively, a participant \(A\) in \(S\) transmits (resp. acquires) its knowledge to (resp. from) \(B\) only when \(A\) fires a \(do^n_k\) prefix on some session \(s[T]\), where \(A\)'s contract in \(T\) sends (resp. receives) \(k\) accordingly. In this case, each state of \(L^S\) is decorated with its corresponding CO2 system so to keep both the KLTS and the semantic LTS of \(S\) isomorphic.

In the example below, we show how two participants \(A\) and \(B\) share knowledge by interacting via two sessions — even though none of systems of contracts in these sessions share knowledge among \(\{A, B\}\).

**Example 7.6 (Sharing knowledge via two sessions).** Consider the CO2 system:

\[
S_1 = \begin{align*}
\text{with contracts } & c_{A1} = \Pi_k, \ c_{A2} = \Pi_k, \ c_{B1} = \Pi_k, \ c_{B2} = \Pi_k. \\
\end{align*}
\]

The system \(S_1\) reduces as follows:

\[
S_1 \xrightarrow{K[\text{fuse} K[\text{fuse}]} S' = (s, t)A([do^n_k k . do^n_k k] | do^n_k k . do^n_k k] | K[\Pi_x A \text{ says } c_{A1}] | do^n_k B \text{ says } c_{B1}] | K[\Pi_y A \text{ says } c_{A2} | do^n_k B \text{ says } c_{B2}] | K[\text{fuse} . \text{fuse}])
\]

Participants \(A\) and \(B\) share knowledge in \(S'\) since they both exchange two \(k\) messages: in session \(s\), \(A\) sends its knowledge to \(B\), while in session \(t\), \(B\) sends its knowledge to \(A\).
Consider the CO sessions, $n$ representing the participants from whom with contracts $c$ sessions is not trivial. Started together, and, as discussed above, guaranteeing safety and liveness properties for inter-dependent knowledge protocol. In fact, this implies that these two sessions are not independent since they can only be to be ready before starting them, thus ensuring that both session, as a whole, will implement a sharing knowledge via, e.g., two sessions as above, it would have to wait for both sessions were to guarantee knowledge sharing through multiple sessions. If a broker example 7.6 shows that it is possible to achieve knowledge sharing through multiple sessions. If a broker to guarantee knowledge sharing via, e.g., two sessions as above, it would have to wait for both sessions to be ready before starting them, thus ensuring that both session, as a whole, will implement a sharing knowledge protocol. In fact, this implies that these two sessions are not independent since they can only be started together, and, as discussed above, guaranteeing safety and liveness properties for inter-dependent sessions is not trivial.

Additionally, for a set of $n$ honest participants to implement a sharing knowledge protocol via binary sessions, $n \times (n - 1)$ sessions would be necessary. Multiparty sessions offer a more elegant solution here, as we show in Example 7.7 below.

**Example 7.7.** Consider the CO2 system:

\[
S_3 = A[(x,y)telk_{x} c_k | do^w_k, do^w_k] \mid B[(w,z)telk_{w} c_k | do^w_k, do^w_k]\]

with contracts $c_k = B?k:B!k$ and $c_k = A!k:A?k$. In this case, both A and B are honest and share knowledge in a single session.

### 7.3. Knowledge sharing choreographies

Our contract model allows for system of contracts that have infinite traces (due to the use of FIFO queues, notably). It is therefore crucial to have an effective and sound procedure to decide whether a given system $T$ shares knowledge among $\mathcal{P}(T)$. Definition 7.8 below gives such a procedure, by analysing a global type $\mathcal{G}$.

We first introduce a map $K : \mathcal{P} \to 2^\mathcal{P}$ that, given a participant name $C$, returns the set of names representing the participants from whom $C$ acquired knowledge. Initially, we have that $K(C) = \{C\}$ for each participant $C$, i.e., a participant has only access to its own knowledge. Given a global type $\mathcal{G}$, we want that

\[
\forall C \in \mathcal{P}(\mathcal{G}) : K(C) = \mathcal{P}(\mathcal{G})
\]

holds at the end of the session embodied by $\mathcal{G}$.

We define the boolean function $K(\mathcal{G},N,K)$ which returns true if all the participants in $N \supseteq \mathcal{P}(\mathcal{G})$ share the same knowledge (with respect to $K$) at the end of the session.

**Definition 7.8.** Let $K(\mathcal{G}) \stackrel{\text{def}}{=} K(\mathcal{G},\mathcal{P}(\mathcal{G}),K_0)$, where $K_0$ is the initial knowledge map, i.e. $\forall C \in \mathcal{P}(\mathcal{G}) : K_0(C) = \{C\}$ and

\[
K(\mathcal{G},N,K) \stackrel{\text{def}}{=} \begin{cases}
K(\mathcal{G}',N,K') & \text{if } \mathcal{G} = A \to B : k ; \mathcal{G}' \land K' = K^B_{K(K)}/k \\
K(\mathcal{G},N,K) \cup K(S_2,N,K) & \text{if } \mathcal{G} = S_1 + S_2 \\
\mathcal{G}(\mathcal{G}_1,N,K_1) \cup K(S_2,N,K_2) & \text{if } \mathcal{G} = S_1 \mid S_2 \land K = K_1 \cup K_2 \\
\mathcal{G}(\mathcal{G}_1,N,K_1) \cup K(S_2,N,K_2) & \mathcal{P}(\mathcal{G}_1) \subseteq \text{dom}(K_i) \text{ with } i \in \{1,2\} \\
\text{false} & \text{otherwise}
\end{cases}
\]

We write $\cup$ for the disjoint union of maps.
The first case of Definition 7.8 handles knowledge sharing interactions: B acquires all the current knowledge of A, while the second case deals with interactions that do not represent a transmission of knowledge. The case of choice simply requires that the condition is satisfied in both branches. In the case of concurrent branches, we need to partition the knowledge map in two – since the sets of participants in the concurrent branches of $G$ are disjoint. In the case of $G = 0$, we check that all participants share the same knowledge. For the sake of simplicity, the function returns false if the global type is recursive.

We write $\text{fuse}_K$ for the variant of $\text{fuse}^\phi$ (as introduced in Section 4.5) where $\phi(G) \iff \hat{r}(G)$. We illustrate Definition 7.8 in Example 7.10 below.

The proposition below (whose proof is in the appendix [17]) formalises the relationship between local and global knowledge sharing.

**Proposition 7.9.** Let $T$ by a system of contracts such that $\vdash T \triangleright G$. If $\hat{r}(G)$, then $T$ shares knowledge among $\Pi(T)$.

We conclude this case study with an example that combine: (i) multiparty sessions, (ii) honest participants, and (iii) contractual agreements based on $\text{fuse}_K$.

**Example 7.10.** Consider the following contracts

\[
\begin{align*}
\text{c}_A &= B_1 \downarrow k; B_2 \uparrow k; B_3 \uparrow k; B_1 \! : k; B_2 \! : k; B_3 \! : k \\
\text{c}_{b_i} &= A \! : k; A \! : k \quad i \in \{1, 2, 3\}
\end{align*}
\]

Let $T = A(c_A) \mid B_1(c_{b_1}) \mid B_2(c_{b_2}) \mid B_3(c_{b_3})$, we have $T \vdash G \triangleright$, with

\[
G = B_1 \rightarrow A : k; B_2 \rightarrow A : k; B_3 \rightarrow A : k; \\
A \rightarrow B_1 : k; A \rightarrow B_2 : k; A \rightarrow B_3 : k; 0
\]

$\hat{r}(G)$ holds since at the end of the first line of $G$, we have $K(A) = \{A, B_1, B_2, B_3\}$ and, in the second line, A shares its acquired knowledge with all the other participants.

Consider the CO$_2$ system below which uses the contracts above, where $K$ is a “knowledge sharing” broker, using primitive $\text{fuse}_K$:

\[
S = A[P_k] \mid B_1[P_{b_1}] \mid B_2[P_{b_2}] \mid B_3[P_{b_3}] \mid K[\text{fuse}_K]
\]

with

\[
\begin{align*}
P_k &= (w)(\text{tell}_{k \rightarrow w} \text{c}_A \cdot \text{do}_{w}^{\text{c}_A} k \cdot \text{do}_{w}^{\text{c}_A} k) \\
P_{b_i} &= (x)(\text{tell}_{k \rightarrow x} \text{c}_{b_i} \cdot \text{do}_{x}^{\text{c}_{b_i}} k \cdot \text{do}_{x}^{\text{c}_{b_i}} k) \quad i \in \{1, 2, 3\}
\end{align*}
\]

Each of the participants above is honest and $K$ will only start knowledge sharing sessions.

8. Related Work

8.1. CO$_2$ and Verification of Honesty

The origin of CO$_2$ goes back to [10], where a first calculus of contracting processes was introduced. The calculus was then generalised in [22] [14] to suit different contract models (e.g., contracts as processes or logic formulæ). In [11], it has been instantiated to a theory of bilateral contracts inspired by [23]. The main differences between our adaptation of CO$_2$ and the original presentation in [22] lie in the specification of session establishment and the introduction of participant variables.

The notion of honesty has been introduced in [11], where a participant is considered honest when, in all possible contexts, it can always exculpate itself by a sequence of its own moves. Here, instead, we require that the participant is ready (Definition 5.5) in all possible contexts, similarly to [13]. The two notions are equivalent in the contract model based on binary session types considered in [18]. In the present paper, a similar correspondence is highlighted by Theorem 6.1 which states that an honest participant $A$ who is
culpable at some session $s$ can always fire some $\text{do}_a$ after a sequence of its own moves. Note however that $A$ may still be culpable after such $\text{do}_a$, e.g. when its contract requires to perform more outputs.

A negative result of [11] is that honesty is not decidable: hence, verification techniques can only aim at safely approximating it. Two different approaches have been proposed in [18,24], where CO\textsubscript{2} is instantiated with binary session types. In [18], a type system for CO\textsubscript{2} processes is presented, associating behavioural types to all session variables used in a process; honesty is then verified by model checking these behavioural types. The technique introduced in [24] is rather different, and more closely based on the semantic definition of honesty: it first devises an abstraction of the contexts wherein $A[P]$ may run, and then safely approximates the honesty of $A$ by model checking such an abstraction. Basically, the abstraction only maintains the process of $A$ and its contracts; the rest of the system is discarded, by over-approximating its moves; then, model checking is used to search the abstracted state space of $A[P]$ for states where $A$ is not ready. If the search fails, then $A$ is honest. Such a search for non-ready states is decidable for finite state processes, i.e. those without delimitation/parallel under process definitions.

We expect that both techniques above can be adapted to cope with the contracts and the calculus studied in this paper. In particular, we now sketch one way to adapt the model checking technique of [24] to the present work.

- First, we develop a safe approximation of honesty, called 1-bounded honesty. We begin by observing that for all synthesizable $T$, the culpability states of participants along the reductions of $T$ can be faithfully captured by only considering “1-bounded” configurations (i.e., where each queue contains at most one message). Furthermore, the local and global types in the present work are a subset of those in [4]. Therefore, they enjoy a key property from [4]: for all synthesizable $T$, each reachable 1-bounded configurations is reachable via a “1-bounded” computation (i.e., an execution such that each step has 1-bounded queues). We can then define “1-bounded” honesty as in Definition 5.8, but using the 1-bounded contract semantics, instead of the unbounded one; the observations above can be used to prove that 1-bounded honesty implies honesty. Note that the 1-bounded contract semantics is finite-state.

- Second, we devise an abstraction of 1-bounded systems of contracts wrt. a participant $A$, whose honesty we want to verify. This abstraction must preserve all the moves of $A$, while over-approximating the moves of other possible participants. This is done similarly to [24]. Such an abstraction also has a finite-state semantics.

- Finally, we design an abstraction for CO\textsubscript{2} systems which preserves the moves of a participant $A$ while over-approximating those of the context (similarly to the abstraction for contracts). Again, this abstraction can be obtained by slightly adapting the one in [24], using the abstract contracts from the previous item. Given any $A[P]$ with no delimitation/parallel under process definitions, such an abstraction has a finite-state semantics.

Summing up, we have a finite-state abstraction of $A[P]$ that, as in [24], can be searched for states where $A$ is not ready. If the search fails, we say that $A[P]$ is abstractly honest. From the observations above, abstract honesty implies 1-bounded honesty, which in turn implies honesty.

8.2. Choreographies and Compliance among Contracts

Several papers have addressed the problem of defining notions of compliance (also called agreement, or compatibility) among contracts, both in the biparty and in the multiparty case. Very roughly, compliance relates contracts that, when composed, give rise to “correct” interactions.

Our notion of compliance (Definition 3.6) exploits the approach introduced in [6,16], which gives a typing system allowing to synthesise a choreography (i.e., a global type) from a set of local types, and gives a characterisation of the global types that can be synthesised. To the best of our knowledge, our approach is the first formal model of distributed systems where contractual agreement is based on the existence of a choreography. Deniélou and Yoshida [4] have given a characterisation of finite-state communicating machines from which it is possible to synthesise a global type, as well as an algorithm to build global types.
from such machines. Their results could also be used as a foundation for an agreement relation similar the one presented here. However, we remark that the global type obtained from the algorithm in [4] is *not* unique in general, notably due to the lack of special treatment of independent interactions (i.e., $\mathcal{G} | \mathcal{G}'$, in our notation). Practically, this means that specific fuse relations (such as the ones discussed in Section 4.3) based on this algorithm may become non-deterministic.

In [25] contracts are considered compliant when their composition may always reach success (so guaranteeing progress until success). In our work, progress is provided by the synthesis of a global type. In [26], a contract language has been proposed for processes, where, as in the $\pi$-calculus, communications are synchronous and channel names can be passed around. The notion of compliance in [27] is based on a testing approach: contracts are compliant if, essentially, their composition has progress. In the same line of work, [28] proposes to “project” processes into session types (akin to the contracts of [27] but without name passing capabilities). The author gives an adapted notion of compliance, called completeness, which permits to characterise when a set of session types is able to reduce further. The semantic correspondence between global specifications (choreographies) and local specification (local types) has been investigated in [29, 12]. In the context of binary session types, several notions of duality (i.e., compliance between two complementary session types) are studied and compared in [30], including the higher-order case. A syntax-independent notion of compliance, defined over generic LTSs with input/output actions, and encompassing first-order binary session types with either synchronous or asynchronous semantics, is introduced in [31].

Choreographies have been studied with different types of syntax and several well-formedness conditions (ensuring properties of their projections). The syntax of our choreographies is close to the ones used in [12, 3, 32] and satisfy similar properties such as knowledge of choice (or local choice); other formalisms include automata-based syntax (e.g., [33]), Petri nets [34] and graphical representation such as BPMN [35]. Each formalism generally comes with a set of conditions ensuring good properties on the system specified by the choreography — such as realisability [35] or connectedness [29] which ensure a correspondence between the order of messages exchanged by the projections and the order specified by the choreography. In the present work, we only require that a choreography specifies a safe system of contracts (which is guaranteed by the synthesis).

8.3. Conformance between Processes and Contracts

The seminal top-down approach of multiparty session types was first presented in [3]. The methodology it advocates is as follows: a team of programmers designs a global view of the interactions to be implemented (i.e., a choreography or global type), then the choreography is projected onto each role, or participant, so to obtain local types; finally each program implementing one or more roles in the choreography is type-checked against its corresponding projection(s). The approach permits, if the choreography is “well-formed”, to guarantee safety properties such as deadlock freedom, and to ensure that an implementation abides by the specified protocol (session fidelity).

As in most works on session types (e.g., [3, 32, 12]), our contract model does not feature concurrency at the local level, i.e., it does not allow for a multi-threaded single contract. However, our results could be easily extended to support local concurrency, by using a choreography synthesis framework which does cater for it: for instance, the one in [37], which offers the same safety guarantees as the one in this work.

Honesty ensures that a CO$_2$ process will interact in a session according to some specific contract. A similar problem is tackled in the session types literature. However, session type based systems do not cater for a notion of “culpability”, and are generally unable to guarantee progress for processes interacting on multiple interleaved sessions without strong assumptions on their execution context. The problem of guaranteeing progress in a composition of processes interacting through several interleaved sessions is tackled in [19, 32] via a complex typing system. Their results include that a well-typed compositions of processes can interact on multiple interleaved sessions without deadlocking — and thus, in our setting, without remaining persistently culpable. In the present paper, a similar progress result is provided for systems of honest participants (Corollary 6.3): a main difference wrt. [19, 32] is that the honesty of each participant is checked “in isolation”, without assumptions on the runtime context.
Another related approach in the context of multiparty session types is that of run-time monitoring [38, 39]. Basically, each monitor is constructed from a local type, and mediates the inputs/outputs of a process in a multiparty session; run-time safety and session fidelity properties are enforced by forbidding interactions not specified in the originating local type. This approach is similar to that of CO$_2$ in the sense that local types are used to “guard” the processes implementing the protocols, as required by the [CO$_2$-Do] rule premises (Figure 4). The framework in [38], however, requires a global type to be available beforehand in order to project process monitors — while in our case, global types are synthesised dynamically from the local types advertised by each participant.

9. Concluding Remarks

In this work, we investigated the combination of the contract-oriented calculus CO$_2$ with a contract model that fulfils two basic design requirements: (i) it supports multiparty agreements, and (ii) it provides an explicit description of the choreography that embodies each agreement. These requirements prompted us towards the well-established results from the session types setting — in particular, regarding the interplay between a well-formed global type and its corresponding local behaviours. We introduced the concepts of global progress and session fidelity in CO$_2$, also inspired from the analogous concepts in theories based on session types. We built our framework upon a simple version of session types, and yet it turns out to be quite flexible, allowing for sessions where the number of participants is not known beforehand.

This article is the full version of the extended abstract published in [40]. Here we streamlined several aspects of the theory, introduced additional results, and extended definitions, explanations and proofs. In addition, we present a large case study whereby our approach is put into practice.

Future work. We plan to extend our work so to offer even more flexibility — for example, by analysing further parameterised fuse variants implementing different session establishment criteria. We also plan to study the possibility for a participant to be involved in a session under multiple contracts, e.g., a bank advertising two services, and a customer publishing a contract which uses both of them in a well-formed choreography.

Another research direction is the concept of “group honesty”. In fact, the current definition of honesty is quite strict: it basically verifies each participant in isolation, thus providing a sufficient (but not necessary) condition for progress. Consider, for example, a CO$_2$ system like:

\[
S = (x,y)(A[tell_x \downarrow x (B[int \oplus B!bool] . do_x \int) | B[tell_y \downarrow y (A?int + A?bool) . do_y \int])
\]

B is not honest, since it is not ready for the bool branch of its contract. However, the system S has progress: when B establishes a session with A, the latter will never take the bool branch; hence, B will not remain culpable. This kind of “group honesty” may be used to validate (sub-)systems of participants developed by the same organization: it would ensure that they never “cheat each other”, and are collectively honest when deployed in any context. Furthermore, the group honesty of all participants in a system S may turn out to be a necessary condition for the global progress of S.

Acknowledgements. We would like to thank Emilio Tuosto for his (in)valuable advice, and the anonymous reviewers of ICE 2013 for their comments and suggestions.

References
