Optimistic Concurrency Semantics for Transactions in Coordination Languages

Suresh Jagannathan and Jan Vitek

Department of Computer Sciences
Purdue University
West Lafayette, IN 47906
{suresh,jv}@cs.purdue.edu

Abstract. There has been significant recent interest in exploring the role of coordination languages as middleware for distributed systems. These languages provide operations that allow processes to dynamically and atomically access and manipulate collections of shared data. The need to impose discipline on the manner in which these operations occur becomes paramount if we wish to reason about correctness in the presence of increased program complexity. Transactions provide strong serialization guarantees that allow us to reason about programs in terms of higher-level units of abstraction rather than lower-level data structures. In this paper, we explore the role of an optimistic transactional facility for a Linda-like coordination language. We provide a semantics for a transactional coordination calculus and state a soundness result for this semantics. Our use of an optimistic concurrency protocol distinguishes this work from previous efforts such as Javaspaces, and enables scalable, deadlock-free implementations.

1 Introduction

A transaction defines a locus of computation that adheres to well-known safety and failure properties. To ensure these properties hold, the semantics of a transactional facility must guarantee that the effects of a transaction are not observable until a commit occurs; when a commit does occur, all effects are propagated instantaneously to the enclosing parent transaction; once a transaction commits, the global state records its effects, and forgets the association between the effect and the transaction; and, all data accesses performed by a transaction are performed serially with respect to other concurrently executing transactions. These four traits correspond to isolation, atomicity, durability, and consistency properties respectively in classical transaction models.

There have been several attempts which explore the integration of transactional semantics within a coordination language framework. Most notably, JavaSpaces [18] and TSpaces [27] combine Linda-like semantics augmented with transactional features into Java. A formal treatment of JavaSpaces, along with

several important extensions, is given by Busi and Zavattaro [11, 12]. The Java-Spaces implementation is based on a *pessimistic* concurrency control semantics: locks are acquired when an entry is accessed, preventing other transactions from witnessing that value until the owning transaction commits. An entry can be physically removed from the shared space only if the removing action is part of the same transaction as any uncommitted reads to that object. Writes to the space become visible to other transactions only when their initiator commits. Although relatively simple to describe, a pessimistic treatment of transactions for coordination languages has two notable disadvantages:

1. *Deadlock*: Since shared data are locked by reads, two transactions each of whom read distinct entries may deadlock depending upon the order in which the locks are acquired. Moreover, it is not possible to acquire locks prior to executing a transaction since determining whether an object will be read depends upon the transaction's dynamic control-flow. The ability to select entries based on pattern-matching further complicates the issue of deadlock detection and avoidance.
2. *Scalability*: Pessimistic concurrency control negatively impacts scalability of distributed implementations as read operations require the acquisition of global locks and thus increase synchronization between distributed nodes. Non-blocking operations, such as Linda's RDP and INP (which test for the presence of a tuple) further complicate the implementation of pessimistic concurrency control protocols as they require that every *potential* entry that may match the operation's template argument be locked for the duration of the transaction [11] to preserve desired serializability invariants.

In this paper, we explore an alternative treatment of transactions within a tuple space coordination framework that addresses these issues. We define a semantics for a transactional variant of Linda [19] based on the synchronous π-calculus [24]. This language can be viewed as the computational core of systems such as Java-Spaces (without some features such as leases and notification [18]) in which effects on the shared data space are controlled via a transactional mechanism based on an *optimistic* concurrency protocol. In this protocol, every transaction conceptually operates over its own local copy of the space and performs actions restricted to this copy. When a transaction is to commit, the state of the local copy is compared to the current state of the global space. If serializability invariants have been violated, the transaction aborts and its local copy is discarded. Otherwise, the changes can be propagated to the global space. The rate at which aborts occur in this scheme is closely correlated to the rate at which concurrent threads operate in incompatible ways over the same tuples.

Our work is distinguished from previous efforts in three main respects: (*a*) our transactional semantics is based on an *optimistic* concurrency model that eliminates issues of deadlock, and obviates scalability limitations which arise in more pessimistic lock-based transactional schemes; (*b*) our treatment of non-blocking operations that test for the presence or absence of tuples, allow us to reason

about the validity of these operations using transaction-local, rather than global, judgments; and (*c*) the semantics naturally allows transactions to be arbitrarily nested. Taken together, these features suggest that expressive, scalable, and deadlock-free transactional implementations for coordination languages are feasible.

2   Coordination Languages and Transactions

We are interested in defining robust, scalable concurrent programs that use a Linda-style coordination framework to mediate access to shared data. Classical coordination models such as Linda allow atomic access and modification of content-addressable data. However, explicit support for treating a locus of computation and the collection of data the computation accesses as a single atomic unit is not provided. An atomic transaction is a control abstraction that provides such functionality. The transactional semantics of concurrent actions in the context of a coordination model ensures serializability and atomicity properties of operations that read, remove, write, and test for the presence of tuples in a global data space.

2.1   JavaSpaces

Sun Microsystem's JavaSpaces is a well-known attempt at integrating a Linda-like coordination language enriched with transactional support into a general purpose programming language. Although expressive, the JavaSpaces programming model is somewhat unintuitive and difficult to implement. We consider its main features here and refer interested readers to the specification [18] and related research papers [8, 11, 12, 9, 26] for further clarification. In particular, we do not consider leasing and event notification which are treated in detail by Busi and Zavattaro in [11, 12].

In JavaSpaces a shared data space is an object which supports at least the following operations: `write()`, `read()`, `take()`, `readIfExists()` and `takeIf-Exists()`. The first operation deposits an entry in the shared space. The operations `read( template)` and `take( template)` scan the shared space for an entry matching `template`, reading or removing a matching tuple, respectively, if one exists and blocking otherwise. Although the match operation in JavaSpaces is user-defined, we restrict ourselves in this paper to traditional value-based structural pattern matching familiar to pure Linda systems. If the template is the distinguished `null` value, any entry in the data space may match. `readIfExists()` and `takeIfExists()` are non-blocking equivalents of `read()` and `take()` respectively (similar to Linda's `rdp` and `inp`). If no value matching the template is currently in the shared space, these operations return `null`. Busi and Zavattaro have discussed in detail the implications of the ability to test for the presence of a value [11, 12]. We revisit these operations in later sections.
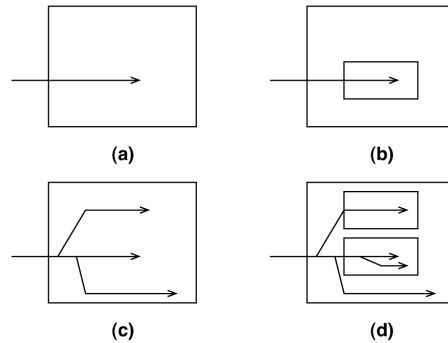
Fig. 1. Threads and transactions may be interleaved in various ways: (a) plain, (b) nested, (c) multi-threaded, (d) multi-threaded and nested.

2.2 Transactions

In this paper, we consider a multi-threaded nested transactional model. When transactions are *nested* [25], each top-level transaction is divided into a number of child transactions; each of which can also have nested transactions. Nested transactions commit from the bottom up, and child transactions must always commit before their parent. A transaction abort at one level does not necessarily affect a transaction in progress at a higher level. The updates of committed transactions at intermediate levels are visible only within the scope of their immediate predecessors. Support for nested transactions is an important feature of our semantics. A nested transaction defines a locus of computation whose effects are bounded by its parent. Thus, the effects of operations on shared data can be controlled by defining an appropriate transaction nesting structure, leading to improved modularity and scalability. Programs are more modular because the effects of a nested transaction are localized to its parent. Programs are scalable because effects are not propagated immediately as the nested transaction commits; instead, changes are aggregated and made globally visible only when the parent transaction commits.

In a multi-threaded transaction model, each transaction can have multiple concurrent threads operating on the same view of the shared state. In a nested transaction model, multiple threads can be executing in parent and child transaction concurrently. Fig. 1 summarizes the different interactions that arise between transactions and threads.

A transactional facility can be modeled by a simple API consisting of the operation `start()` and `commit()`. We assume an implicit transactional context so that `start()` initiates a new transaction. If the current thread is already running within a transaction, the new transaction is nested in the current one, otherwise the new transaction becomes a top-level transaction. The `commit()` operation
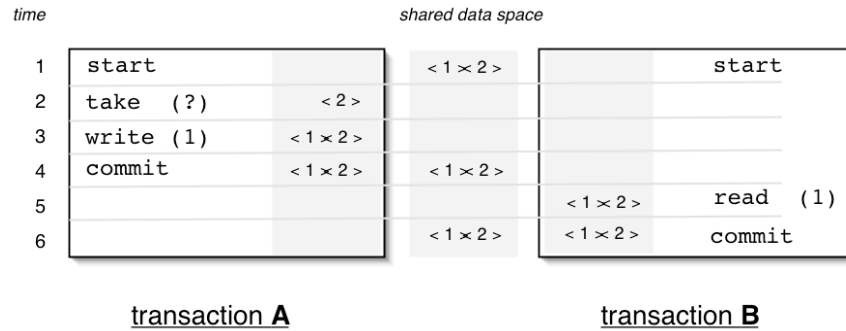
time	transaction **A**		shared data space		transaction **B**
1	start		< 1 × 2 >		start
2	take (?)	< 2 >			
3	write (1)	< 1 × 2 >			
4	commit	< 1 × 2 >	< 1 × 2 >		
5				< 1 × 2 >	read (1)
6			< 1 × 2 >	< 1 × 2 >	commit

Fig. 2. Serialized execution of transaction `A` and `B`. The shared data space is only updated upon commit. Each transaction has its own view of the data space. The shared space starts out as { $\langle 1 \rangle$, $\langle 2 \rangle$ }. At time 2, transaction `A` performs a `take(?)` which removes entry $\langle 1 \rangle$ from the space. Thus from `A`'s point of view the shared space contains a single entry ($\langle 2 \rangle$). A commit reconciles a transaction's local view with the state of the shared space.

attempts to finalize the changes made by the current transaction, propagating its results to the parent if one exists.

2.3   Motivating examples

To illustrate some salient issues related to the incorporation of transactional semantics into a coordination language, consider the actions defined in Fig. 2. Transactions `A` and `B` perform read and write actions on a global data space which initially has two singleton tuples $\langle 1 \rangle$ and $\langle 2 \rangle$. Transaction `A` takes $\langle 1 \rangle$ and then writes it back. Once `A` is done, transaction `B` performs a non-destructive read of $\langle 1 \rangle$ and commits. This is a valid interleaving as all actions of `A` precede the actions of `B`.

Figs. 3 and 4 illustrate possible interleavings if the actions performed by `A` and `B` are enclosed within transactions under different transactional semantics. In Fig. 3 transaction `B` reads $\langle 1 \rangle$ after it was taken by `A`. In a *pessimistic* model such as the one used by JavaSpaces [11], this interleaving is not valid as the `take()` performed by `A` locks $\langle 1 \rangle$ and prevents `B` from performing a read. In an *optimistic* semantics, transaction `B` can proceed and, in fact, both transactions can commit successfully.

Fig. 4 illustrates a case in which, under an *optimistic* semantics, transaction `B` has to abort. In this example, `A` reads $\langle 1 \rangle$ but writes back another copy of $\langle 2 \rangle$, while `B` still reads $\langle 1 \rangle$. Of course when `B` gets to the point of committing, it finds that the global space does not have tuple $\langle 1 \rangle$ and thus is forced to abort.

Fig. 5 illustrates a more serious problem with JavaSpaces. We use `inp` as a non-blocking take (*i.e.* `takeIfExists()`). Assume the global data space is initially empty. Transaction A checks if tuple $\langle 2 \rangle$ is present, and will take if it is, while transaction B attempts to take $\langle 1 \rangle$. Both operations fail as the space contains
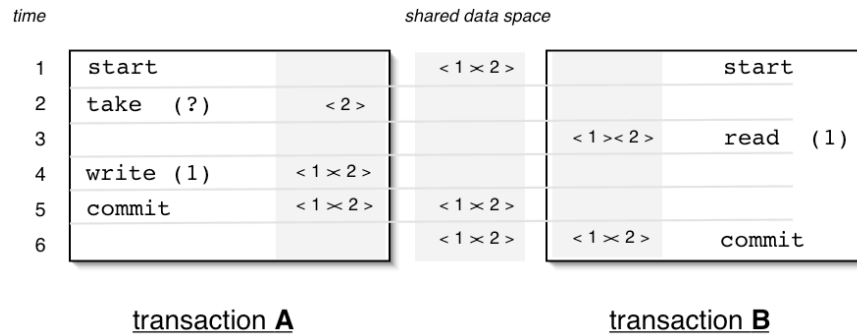
time	transaction A		shared data space		transaction B
1	start		<1×2>		start
2	take (?)	<2>			
3				<1><2>	read (1)
4	write (1)	<1×2>			
5	commit	<1×2>	<1×2>		
6			<1×2>	<1×2>	commit

Fig. 3. Valid optimistic interleaving. In a pessimistic semantics transaction B would have to block until A commits, *i.e.* it would behave as illustrated in Fig. 2.
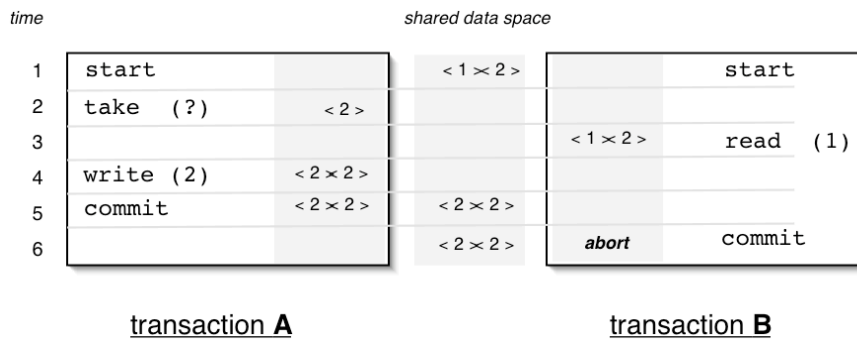
time	transaction A		shared data space		transaction B
1	start		<1×2>		start
2	take (?)	<2>			
3				<1×2>	read (1)
4	write (2)	<2×2>			
5	commit	<2×2>	<2×2>		
6			<2×2>	*abort*	commit

Fig. 4. Valid optimistic interleaving. Transaction B fails to commit due to a conflict in the global data space (entry $\langle 1 \rangle$ is absent).

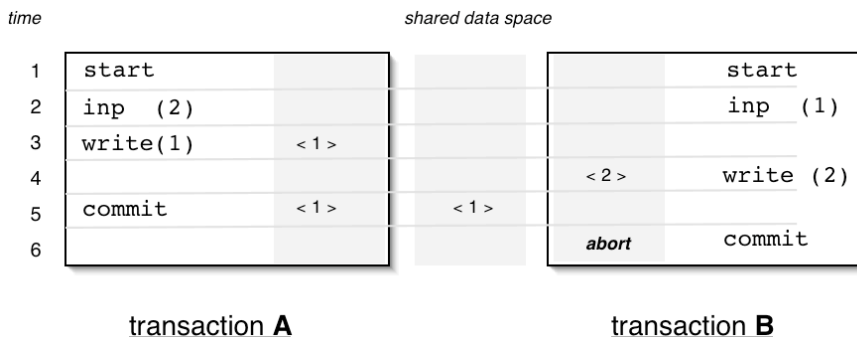| time	transaction A		shared data space		transaction B
1	start				start
2	inp (2)				inp (1)
3	write(1)	<1>			
4				<2>	write (2)
5	commit	<1>	<1>		
6				*abort*	commit

Fig. 5. Valid optimistic interleaving. Transaction B abort due to a conflict in the state of the shared space ($\langle 1 \rangle$ is present). A pessimistic semantics would deadlock.

no matching value. A proceeds to write $\langle 1 \rangle$ and B to write $\langle 2 \rangle$, after which both transactions commit. In an *optimistic* semantics transaction B would abort because at time 6 the contents of the global space is $\{\langle 1 \rangle \}$ while transaction B expect $\langle 1 \rangle$ *not* to be in the space (since it has checked this at time 2). In Sun's JavaSpace *implementation*, both transactions are allowed to commit successfully. This behavior is incorrect as noted by Busi and Zavattaro [11]. In a correct *pessimistic semantics* (*i.e.* [11]) this problem can be redressed by prohibiting other transactions from depositing any matching writes to the tuple-space once A performs its test. However, in this example, such a semantics would lock any transaction from depositing a single element tuple, and would thus prevent B from committing since it subsequently performs a `write` on $\langle 2 \rangle$.

Thus, the example of Fig. 5 would deadlock in a pessimistic semantics because each transaction is trying to acquire a lock held by the other transaction. In general, a correct implementation of pessimistic concurrency control may entail locking large portions of the shared data space. Consider for instance the action `readIfExists( null)` which, in JavaSpaces, matches *any* entry in the data space. If the data space is empty when the action is evaluated, the current transaction will lock the entire space and no other transaction will be able to write to the space until this one commits if serializability invariants are to be preserved.

Transaction protocols for coordination languages are further complicated when nested transactions are considered. The semantics of nested transactions permit child transactions to see the effects of their parents, even before the parent commits. Using pessimistic concurrency control, the rules governing when tuple operations become visible, and when read and write locks are released must be strengthened. Fig. 6 describes a valid execution sequence in which transaction A defines a nested transaction C. Tuples written by C are unavailable to B until A commits, although they are available to A once C commits. Tuples written by A are available to C even before A commits. This figure reveals that fine-grained pessimistic concurrency protocols require sophisticated lock management schemes to allow nested transactions to inherit locks on tuples owned by their ancestors. In constrast, an optimistic concurrency protocol maintains multiple local logs and directly incorporates notions of visibility into the log structure. When a nested transaction attempts to read a tuple, writes performed by its parent that match the tuple pattern are visible to the nested transaction. To ensure that writes performed in a nested transaction are propagated correctly, an optimistic concurrency mechanism need only guarantee that nesting of logs reflect transaction nesting. No special lock management protocol is required.

3   A transactional coordination calculus

The transactional Linda calculus is based on the synchronous π calculus [1, 4, 22] which provides a small and elegant computational core. The main departure
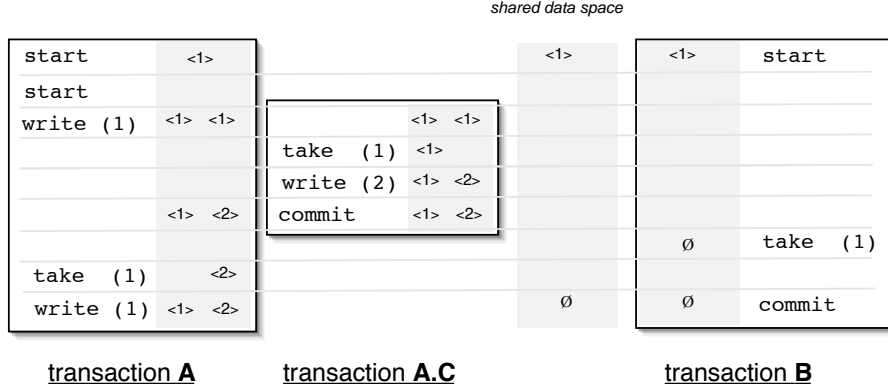
shared data space

start	<1>			<1>	<1>	start
start						
write (1)	<1> <1>		<1> <1>			
		take (1)	<1>			
		write (2)	<1> <2>			
	<1> <2>	commit	<1> <2>			
					∅	take (1)
take (1)	<2>					
write (1)	<1> <2>			∅	∅	commit

transaction **A**   transaction **A.C**   transaction **B**

Fig. 6. Actions performed by a nested transaction are not visible to the parent until the inner transaction commits; the effects of a nested transaction are not visible to other transactions until its parent commits.

from the π calculus is that rather than communicating by means of channel-based message passing, processes exchange tuples through a common shared space. Embeddings of Linda in process calculi have been explored in previous work [6, 17]. We shall focus only on essential features from the point of view of the transactional behavior of the calculus.

3.1 Syntax

The syntax of the calculus is summarized in Fig. 7. We take an infinite set of names ranged over by meta-variable x. Tuples are sequences of formal and actual entries ranged over by meta-variable v. The empty tuple is denoted by $\langle\rangle$. The expression x.v denotes the tuple resulting from concatenating an actual name x with tuple v, while $x^?$.v denotes the concatenation of a formal variable x with v. Meta-variable t ranges over transaction identifiers, and ℓ over sequences of transaction identifiers.

The syntactic category of processes, ranged over by P and Q, includes the empty process **0** and parallel composition of processes P | Q. Synchronous output, written v.P, deposits the tuple denoted by v in the shared data space and proceeds to execute P. The input operation (v).P evaluates the template v against the shared space; if a matching value is found, the tuple is taken from the shared space and the process continues as P with the formals of v replaced by actuals from the tuple; the operation blocks if no matching value can be found. Guarded replication, written (v)!P, performs an input of v and, if a match is found, proceeds as P with (in parallel) a copy of the original process (v)!P. The test operation, (v)?P ; Q is a non-blocking test for presence which tries to find a tuple matching the template v, and if one is found, proceeds to evaluate P with the tuple; if no matching tuple is found in the space Q is evaluated. The restriction operator

Syntax:

$$P ::= 0 \mid P|Q \mid (\nu\,x)\,P \mid \textbf{trans}\ P \mid t[P] \mid \textbf{commit} \mid (v).P \mid v.P \mid (v)?P\,;\,Q \mid (v)!P$$

$$v ::= \langle\rangle \mid x.v \mid x^?.v \qquad \tau ::= v \mid {}^{v}\!/_{v} \mid \neg v \mid \textbf{com} \mid \textbf{srt} \qquad \ell ::= \langle\rangle \mid \ell.t$$

$$\mathcal{E} ::= \langle\rangle \mid \mathcal{E}.\ell:\rho \qquad\qquad \rho ::= \langle\rangle \mid \rho.v \mid \rho.\neg v \mid \rho.v^?$$

$$(v).P \xrightarrow{\ {}^{v'}\!/_{v}\ }_{\langle\rangle} {}^{v'}\!/_{v}\,P \qquad \text{(L-INP)}$$

$$(v)?P\,;\,Q \xrightarrow{\ v\ }_{\langle\rangle} {}^{v'}\!/_{v}\,P \qquad \text{(L-POS)}$$

$$(v)?P\,;\,Q \xrightarrow{\ \neg v\ }_{\langle\rangle} Q \qquad \text{(L-NEG)}$$

$$(v)!P \xrightarrow{\ {}^{v'}\!/_{v}\ }_{\langle\rangle} {}^{v'}\!/_{v}\,P \mid (v)!P \qquad \text{(L-LOOP)}$$

$$v.P \xrightarrow{\ v\ }_{\langle\rangle} P \qquad \text{(L-OUTP)}$$

$$\textbf{trans}\ P \xrightarrow{\ \textbf{srt}\ }_{t} t[P] \qquad \text{(L-START)}$$

$$t[\,\textbf{commit} \mid P\,] \xrightarrow{\ \textbf{com}\ }_{t} 0 \qquad \text{(L-COMM)}$$

$$\frac{P \xrightarrow{\ \tau\ }_{\ell} P'}{t[P] \xrightarrow{\ \tau\ }_{t.\ell} t[P']} \qquad \text{(L-TRANS)}$$

$$\frac{P \xrightarrow{\ {}^{v'}\!/_{v}\ }_{\ell} P' \quad v',\mathcal{E}' = \texttt{take}(\mathcal{E},\ell) \quad v \le v'}{P\ \mathcal{E} \xRightarrow{\ \ell\ } P'\ \mathcal{E}'} \qquad \text{(G-INP)}$$

$$\frac{P \xrightarrow{\ v\ }_{\ell} P' \quad \mathcal{E}' = \texttt{put}(\mathcal{E},\ell,v)}{P\ \mathcal{E} \xRightarrow{\ \ell\ } P'\ \mathcal{E}'} \qquad \text{(G-OUTP)}$$

$$\frac{P \xrightarrow{\ \neg v\ }_{\ell} P' \quad \mathcal{E}' = \texttt{neg}(\mathcal{E},\ell,v)}{P\ \mathcal{E} \xRightarrow{\ \ell\ } P'\ \mathcal{E}'} \qquad \text{(G-NEG)}$$

$$\frac{P \xrightarrow{\ \textbf{srt}\ }_{\ell} P' \quad \mathcal{E}' = \texttt{start}(\mathcal{E},\ell)}{P\ \mathcal{E} \xRightarrow{\ \ell\ } P'\ \mathcal{E}'} \qquad \text{(G-START)}$$

$$\frac{P \xrightarrow{\ \textbf{com}\ }_{\ell} P' \quad \mathcal{E}' = \texttt{commit}(\mathcal{E},\ell)}{P\ \mathcal{E} \xRightarrow{\ \ell\ } P'\ \mathcal{E}'} \qquad \text{(G-COMM)}$$

$$\frac{P\ \mathcal{E} \equiv P'\ \mathcal{E} \quad P'\ \mathcal{E} \xRightarrow{\ \ell\ } Q'\ \mathcal{E}' \quad Q' \equiv Q}{P\ \mathcal{E} \xRightarrow{\ \ell\ } Q\ \mathcal{E}'} \qquad \text{(G-CONG)}$$

Evaluation Contexts:

$$E[\bullet] \quad \mid \quad E[\bullet] \mid P$$

$$\frac{P \xrightarrow{\ \tau\ }_{\ell} P'}{t[E[P]] \xrightarrow{\ \tau\ }_{t.\ell} t[E[P']]}$$

$$\frac{P \xrightarrow{\ \tau\ }_{\ell} P' \quad x \notin \mathit{fn}(\tau)}{(\nu\,x)\,P \xrightarrow{\ \tau\ }_{\ell} (\nu\,x)\,P'}$$

Structural Congruence:

$$P \mid Q \equiv Q \mid P \qquad (\nu\,x)\,(\nu\,y)\,P \equiv (\nu\,y)\,(\nu\,x)\,P$$

$$P \mid (\nu\,x)\,Q \equiv (\nu\,x)\,P \mid Q \qquad \text{if } x \notin \mathit{fn}(P)$$

$$t[\,(\nu\,x)\,P\,] \equiv (\nu\,x)\,t[P]$$

$$(\nu\,x)\,P\ \mathcal{E} \equiv P'\ \mathcal{E} \qquad \text{if } x \notin \mathit{fn}(\mathcal{E}) \wedge P \equiv P'$$

Fig. 7. Syntax and Semantics

Free Names:

$$fn(\mathbf{0}) = \{\} \qquad fn(\mathtt{P} \mid \mathtt{Q}) = fn(\mathtt{P}) \cup fn(\mathtt{Q}) \qquad fn((\nu\,\mathtt{x})\,\mathtt{P}) = fn(\mathtt{P}) - x \qquad fn(\mathtt{t}[\mathtt{P}]) = fn(\mathtt{P})$$

$$fn(\mathbf{commit}) = \{\} \qquad fn((\mathtt{v}).\mathtt{P}) = fn(\mathtt{P}) \cup fn(\mathtt{v}) - bn(\mathtt{v}) \qquad fn(\mathtt{v}.\mathtt{P}) = fn(\mathtt{P}) \cup fn(\mathtt{v}) - bn(\mathtt{v})$$

$$fn(\mathbf{trans}\;\mathtt{P}) = fn(\mathtt{P}) \qquad fn((\mathtt{v})?\,\mathtt{P}\,;) = fn(\mathtt{P}) \cup fn(\mathtt{v}) - bn(\mathtt{v}) \qquad fn((\mathtt{v})!\,\mathtt{P}) = fn(\mathtt{P}) \cup fn(\mathtt{v}) - bn(\mathtt{v})$$

$$fn(\mathtt{x}.\mathtt{v}) = \{x\} \cup fn(\mathtt{v}) \qquad fn(\mathtt{x}^?.\mathtt{v}) = fn(\mathtt{v}) \qquad fn(\langle\rangle) = \{\}$$

$$bn(\mathtt{x}.\mathtt{v}) = bn(\mathtt{v}) \qquad bn(\mathtt{x}^?.\mathtt{v}) = \{x\} \cup bn(\mathtt{v}) \qquad bn(\langle\rangle) = \{\}$$

Matching:

$$\langle\rangle \leq \langle\rangle \qquad \mathtt{x}.\mathtt{v} \leq \mathtt{x}.\mathtt{v}' \;\; \text{if } \mathtt{v} \leq \mathtt{v}' \qquad \mathtt{x}^?.\mathtt{v} \leq \mathtt{x}.\mathtt{v}' \;\; \text{if } \mathtt{v} \leq \mathtt{v}'$$

Fig. 8. Free Names and Matching Rules.

$(\nu\,\mathtt{x})\,\mathtt{P}$ generates a fresh name $\mathtt{x}$; the calculus is lexically scoped, so $\mathtt{x}$ is visible only in process P. Finally **trans** $\mathtt{P}$ and **commit** are used to, respectively, start and terminate a new transaction. The transaction **trans** $\mathtt{P}$ will execute the process $\mathtt{P}$ until a **commit** is evaluated, at which point all changes effected by the transaction will become visible to all other transactions. If a conflict is detected the transaction will not be able to commit and the evaluation of the thread will remain stuck. Note also that all other threads within the same transaction and nested transaction are terminated upon a commit. The calculus does not provide an explicit **abort** operation; we model aborted transactions as stuck terms on **commit** transitions.

3.2 Semantics

The semantics of the calculus is shown in Figs. 7, 8, and 9. The semantics is stratified so that the local reduction relation $\mathtt{P} \xrightarrow{\tau}_\ell \mathtt{Q}$ defines that process $\mathtt{P}$ can reduce to $\mathtt{Q}$ in one step. The transition is labeled by an action τ and a

$$\tfrac{x.\mathtt{v}}{x.\mathtt{v}'}\,\mathtt{P} = \tfrac{\mathtt{v}}{\mathtt{v}'}\,\mathtt{P} \qquad \tfrac{x.\mathtt{v}}{y^?.\mathtt{v}'}\,\mathtt{P} = \tfrac{\mathtt{v}}{\mathtt{v}'}\,\tfrac{x}{y}\,\mathtt{P} \qquad \tfrac{\langle\rangle}{\langle\rangle}\,\mathtt{P} = \mathtt{P}$$

$$\tfrac{x}{y}\,\mathbf{0} = \mathbf{0} \qquad \tfrac{x}{y}\,\mathtt{P} \mid \mathtt{Q} = \tfrac{x}{y}\,\mathtt{P} \mid \tfrac{x}{y}\,\mathtt{Q} \qquad \tfrac{x}{y}\,(\nu\,\mathtt{x})\,\mathtt{P} = (\nu\,\mathtt{x})\,\mathtt{P}$$

$$\tfrac{x}{y}\,(\nu\,\mathtt{x}')\,\mathtt{P} = (\nu\,\mathtt{x}')\,\tfrac{x}{y}\,\mathtt{P} \qquad \tfrac{x}{y}\,\mathtt{t}[\mathtt{P}] = \mathtt{t}[\tfrac{x}{y}\,\mathtt{P}] \qquad \tfrac{x}{y}\,\mathbf{commit} = \mathbf{commit}$$

$$\tfrac{x}{y}\,(\mathtt{v}).\mathtt{P} = (\tfrac{x}{y}\,\mathtt{v}).\tfrac{x}{y}\,\mathtt{P} \;\; \text{if } x \notin fn(\mathtt{v}) \qquad \tfrac{x}{y}\,(\mathtt{v}).\mathtt{P} = (\tfrac{x}{y}\,\mathtt{v}).\mathtt{P} \;\; \text{if } x \in fn(\mathtt{v})$$

$$\tfrac{x}{y}\,(\mathtt{v})!\,\mathtt{P} = (\tfrac{x}{y}\,\mathtt{v})!\,\tfrac{x}{y}\,\mathtt{P} \;\; \text{if } x \notin fn(\mathtt{v}) \qquad \tfrac{x}{y}\,(\mathtt{v})!\,\mathtt{P} = (\tfrac{x}{y}\,\mathtt{v})!\,\mathtt{P} \;\; \text{if } x \in fn(\mathtt{v})$$

$$\tfrac{x}{y}\,(\mathtt{v})?\,\mathtt{P}\,;\,\mathtt{Q} = (\tfrac{x}{y}\,\mathtt{v})?\,\tfrac{x}{y}\,\mathtt{P}\,;\,\tfrac{x}{y}\,\mathtt{Q} \;\; \text{if } x \notin fn(\mathtt{v}) \qquad \tfrac{x}{y}\,(\mathtt{v})?\,\mathtt{P}\,;\,\mathtt{Q} = (\tfrac{x}{y}\,\mathtt{v})?\,\mathtt{P}\,; \;\; \text{if } x \in fn(\mathtt{v})$$

$$\tfrac{x}{y}\,\mathtt{v}.\mathtt{P} = (\tfrac{x}{y}\,\mathtt{v}).\tfrac{x}{y}\,\mathtt{P} \;\; \text{if } x \notin fn(\mathtt{v}) \qquad \tfrac{x}{y}\,\mathtt{x}.\mathtt{v} = y.\tfrac{x}{y}\,\mathtt{v} \qquad \tfrac{x}{y}\,\mathtt{x}^?.\mathtt{v} = \mathtt{x}^?.\mathtt{v}$$

$$\tfrac{x}{y}\,\mathtt{z}^?.\mathtt{v} = \mathtt{z}.\tfrac{x}{y}\,\mathtt{v} \qquad \tfrac{x}{y}\,\langle\rangle = \langle\rangle$$

Fig. 9. Substitutions.

transaction ℓ. The global relation $P \, \mathcal{E} \xrightarrow{\ell} Q \, \mathcal{E}'$ defines the behavior of a program P in an environment $\mathcal{E}$. We assume congruence of environments under element reordering.

In the local reduction relation the tuple space remains implicit. Each reduction step is labeled by one of the following actions: $v, {}^{v'}\!/_{v}, \neg v, \textbf{srt}$, and $\textbf{com}$, indicating a write of v, a take of v' matching v, a test for absence of v, the start of a transaction and a commit, respectively. Furthermore transitions are labeled by the issuing transaction name, ℓ, composed of a sequence of transaction identifiers; the different identifiers correspond to the levels of nesting, thus $\mathtt{t.t'.t'''}$ denotes an action performed by transaction $\mathtt{t'''}$ nested in transactions $\mathtt{t'}$ and $\mathtt{t}$.

Starting a transaction with $\textbf{trans}\ P$ creates a process P running within some transaction $\mathtt{t}$, and is denoted $\mathtt{t}[P]$. We assume that each new transaction identifier $\mathtt{t}$ is chosen to be distinct. Transactions can be nested, for instance the expression $\mathtt{t}[\mathtt{t'}[P] \mid Q]$ denotes two parallel processes P and Q such that Q is running within a top level transaction $\mathtt{t}$ and a process P in a transaction $\mathtt{t'}$ nested within $\mathtt{t}$. For example, the expression $\textbf{trans}\ (\textbf{trans}\ v.\mathbf{0})$ denotes a process that will spawn two transactions, the second nested within the first. Then the expression will reduce in one step to the inactive process:

$$ \mathtt{t}[\ \mathtt{t'}[v.\mathbf{0}]\ ] \quad \xrightarrow{v}_{\mathtt{t.t'}} \quad \mathtt{t}[\ \mathtt{t'}[\mathbf{0}]\ ] $$

Notice that in this configuration, the tuple v was output from transaction $\mathtt{t.t'}$ and since there is no commit for that transaction that value will never become available to other processes.

The global reduction relation $\xrightarrow{\ell}$ manages the shared data space and defines the semantics of the transactional facility. The semantics does not fix a specific implementation or a particular transactional model as it is parameterized by an environment $\mathcal{E}$ which contains the tuple space and five operations over environments that implement the actual transactional model. Although we provide definitions for these operations that capture the essential traits of an optimistic concurrency protocol, other specifications that define alternative implementations of the protocol, or which express different transactional semantics (such as pessimistic concurrency) can be developed without modifying the global reduction relation.

The rule (G-INP) defines the behavior of a destructive read over the tuple space. Given an environment and a transaction, $\texttt{take}$ returns a value available to that transaction; if that value matches the requested template, the transition proceeds with an updated environment. (G-OUTP) is similar in that it relies on the $\texttt{put}$ operation to record that a transaction ℓ has output tuple v. (G-NEG) applies in case there is no tuple matching template v available to transaction ℓ. When this occurs, the $\texttt{neg}$ operation records this fact in the environment. (G-START) sets up the environment for a new transaction. (G-COM) attempts to commit a transaction. Finally, (G-CONG) allows reduction up to structural congruence

$$\frac{\begin{array}{c} \mathtt{v} \in \mathtt{visible}(\mathcal{E},\ell) \\ \mathcal{E}' = \mathcal{E}\,.\,(\ell:\rho\,.\,\mathtt{v}^?) \end{array}}{\mathtt{v},\mathcal{E}' = \mathtt{take}(\mathcal{E},\ell)} \qquad \frac{\begin{array}{c} \mathtt{match}(\mathtt{visible}(\mathcal{E},\ell),\mathtt{v}) = \{\} \\ \mathcal{E}' = \mathcal{E}\,.\,(\ell:\rho\,.\,\neg\mathtt{v}) \end{array}}{\mathcal{E}' = \mathtt{neg}(\mathcal{E},\ell,\mathtt{v})}$$

$$\frac{\mathcal{E}' = \mathcal{E}\,.\,(\ell:\rho\,.\,\mathtt{v})}{\mathcal{E}' = \mathtt{put}(\mathcal{E},\ell,\mathtt{v})} \qquad \frac{\mathcal{E}' = \mathtt{reflect}(\mathcal{E}'',\rho,\ell)}{\mathcal{E}' = \mathtt{commit}(\mathcal{E}''\,.\,(\ell\,.\,\mathtt{t}:\rho),\ell)}$$

$$\mathcal{E}\,.\,(\ell:\langle\rangle) = \mathtt{start}(\mathcal{E},\ell)$$

Fig. 10. Transactional semantics.

of processes. Top level ν-binders can be erased to allow names to flow into the environment.

3.3 Optimistic transactional facility

The implementation of an optimistic transactional model is shown in Figs. 10 and 11. In this scheme, the shared tuple space is represented by an environment $\mathcal{E} = \ell_1 : \rho_1 \ldots \ell_n : \rho_n$ which contains per-transaction logs $\rho_1 \ldots \rho_n$. Each of these logs is a sequence comprised of events, $\mathtt{v}, \mathtt{v}^?, \neg\mathtt{v}$; these events denote an output of $\mathtt{v}$, a removal of a tuple $\mathtt{v}^?$ and an absence of a match for template $\neg\mathtt{v}$, respectively.

$$\mathcal{E} = \mathtt{reflect}(\mathcal{E},\langle\rangle,\ell) \qquad \frac{\mathcal{E}' = \mathtt{reflect}(\mathcal{E}\,.\,(\ell:\rho'\,.\,\mathtt{v}),\rho,\ell)}{\mathcal{E}' = \mathtt{reflect}(\mathcal{E}\,.\,(\ell:\rho'),\mathtt{v}\,.\,\rho,\ell)}$$

$$\frac{\begin{array}{c} \mathtt{v} \in \mathtt{visible}(\mathcal{E},\ell) \\ \mathcal{E}' = \mathtt{reflect}(\mathcal{E}\,.\,(\ell:\rho'\,.\,\mathtt{v}^?),\rho,\ell) \end{array}}{\mathcal{E}' = \mathtt{reflect}(\mathcal{E}\,.\,(\ell:\rho'),\mathtt{v}^?\,.\,\rho,\ell)} \qquad \frac{\begin{array}{c} \mathtt{match}(\mathtt{visible}(\mathcal{E},\ell),\mathtt{v}) = \{\} \\ \mathcal{E}' = \mathtt{reflect}(\mathcal{E}\,.\,(\ell:\rho'\,.\,\neg\mathtt{v}),\rho,\ell) \end{array}}{\mathcal{E}' = \mathtt{reflect}(\mathcal{E}\,.\,(\ell:\rho'),\neg\mathtt{v}\,.\,\rho,\ell)}$$

$$\mathtt{visible}(\mathcal{E},\ell) = \mathtt{find}(\mathtt{merge}(\mathcal{E},\ell)) \qquad \mathtt{match}(\mathcal{V},\mathtt{v}) = \{\,\mathtt{v}' \mid \mathtt{v}' \in \mathcal{V} \wedge \mathtt{v} \le \mathtt{v}'\,\}$$

$$\begin{array}{l} \mathtt{find}(\rho\,.\,\mathtt{v}) = \mathtt{find}(\rho \cup \mathtt{v}) \\ \mathtt{find}(\rho\,.\,\mathtt{v}^?) = \mathtt{find}(\rho - \mathtt{v}) \\ \mathtt{find}(\rho\,.\,\neg\mathtt{v}) = \mathtt{find}(\rho) \end{array} \qquad \begin{array}{l} \mathtt{merge}(\mathcal{E},\langle\rangle) = \langle\rangle \\ \mathtt{merge}(\mathcal{E}\,.\,(\ell\,.\,\mathtt{t}:\rho),\ell\,.\,\mathtt{t}) = \mathtt{merge}(\mathcal{E},\ell)\,.\,\rho \end{array}$$

Fig. 11. Auxiliary functions.

The operations on the environment are $\mathtt{take}, \mathtt{neg}, \mathtt{put}, \mathtt{commit}$ and $\mathtt{start}$. Whenever a new transaction is created $\mathtt{start}(\mathcal{E},\ell)$ is used to extend the environment with an empty log for transaction ℓ. The operation $\mathtt{put}(\mathcal{E},\ell,\mathtt{v})$ extends the log of transaction ℓ with a tuple $\mathtt{v}$. The operation $\mathtt{take}(\mathcal{E},\ell)$ returns an arbitrary tuple visible by transaction ℓ and records in the log that the tuple has been removed. The operation $\mathtt{neg}(\mathcal{E},\ell,\mathtt{v})$ records the fact no tuple matching template $\mathtt{v}$ is visi-

ble to the transaction ℓ. Finally, operation $\texttt{commit}(\mathcal{E}, \ell)$ attempts to commit the change performed by transaction ℓ to the log of the parent transaction. Commit operates by replaying the changes performed by transaction ℓ. The commit fails if the environment for ℓ's parent is not in the expected state.

3.4 Soundness

Proving the correctness of the semantics requires that we show our treatment of transactions preserves desired isolation and atomicity properties. To do so, we first define a notion of a *well-defined* state. Intuitively, a state is well-defined if the contents of the logs associated with each transaction are such that no conflicts would arise. One way to manifest this idea is to compare the state of a transaction's parent with its child at the point the child attempts to commit. If the parent reflects commits from other transactions that violate invariants recorded in the log of the committing child, the child state is not considered well-defined. We formalize this intuition thus:

Definition 1. *$\mathcal{E}$ is well-defined if for any transaction ℓ such that $\mathcal{E} \equiv \mathcal{E}' . (\ell : \rho)$ and $\ell = \ell' .\texttt{t}$, the function $\texttt{reflect}(\mathcal{E}', \ell', \rho)$ is defined.*

Thus, the log of a child transaction is well-defined with respect to the parent if actions observed by the child are consistent with the actions seen by the parent. Satisfiability of this condition implies, that after a child commits, $\texttt{write}$ and $\texttt{take}$ operations performed by the child must be visible in the parent, and tuples that were observed to be absent by the child must still be so in the parent. Observe that the $\texttt{reflect}$ operation propagates tuple events from a child transaction to a parent transaction provided these events do not violate natural visibility and serializability invariants.

Definition 2. *A trace $tr(R) = \texttt{P}_0 \, \mathcal{E}_0 \overset{\ell_{0}}{\Longrightarrow} \ldots \overset{\ell_{n}}{\Longrightarrow} \texttt{P}_n \, \mathcal{E}_n$ is serial iff $\forall i, j, k$ such that $0 \leq i \leq j \leq k \leq n$, and $\ell_i = \ell_k$, $l_i \lhd l_j$ (read "l_i is a prefix of l_j).*

A serial trace is one in which for all pairs of reduction steps with the same transaction label ℓ, all actions that occur between these two steps are taken on behalf of that transaction or its children. Given a notion of a serial trace, we can define a soundness theorem that captures our desired notion of serializability:

Theorem 1. *Let R be a sequence of reductions $\texttt{P}_0 \, \mathcal{E}_0 \overset{\ell_{0}}{\Longrightarrow} \ldots \overset{\ell_{n}}{\Longrightarrow} \texttt{P}_{n_1} \, \mathcal{E}_{n+1}$. If $\mathcal{E}_{n+1}$ is well-defined, then there exists a sequence $R' = \texttt{P}_0 \, \mathcal{E}_0 \overset{\ell'_{0}}{\Longrightarrow} \ldots \overset{\ell'_{n}}{\Longrightarrow} \texttt{P}_{n_1} \, \mathcal{E}_{n+1}$ such that R' is serial.*

The proof of this theorem appeals to notions of permutability on global actions. Informally, two actions α_1 and α_2 executed in transactions ℓ_1 and ℓ_2 are

permutable if they have no data or control-dependency with each other. For example, a `take` operation performed by a transaction ℓ has a data dependency with any `write` operation performed by any of ℓ's parents that matches the read tuple. Simlarly, a `write` action logged in any parent transaction of ℓ has a data dependency with any $\neg v$ action recorded in ℓ. This means that any valid serializable permutation cannot move a `write` action in a parent above a child action that successfully tested for the absence of the tuple being written.

4 Related Work and Conclusions

The Linda coordination model [19, 14] uses generative communication on anonymous structured data to facilitate interactions among concurrent programs. Its simplicity and generality make it a fertile vehicle in which to explore and formalize various concurrency abstractions [9, 17, 7, 13]. However, Linda does not directly support atomic operations on aggregate shared data. To alleviate this drawback, systems such as JavaSpaces [18] or TSpaces [27] allow operations on tuple-spaces to be encapsulated within transactions.

There is a large body of work that explores the formal specification of various flavors of transactions [23, 16, 20]. Most closely related to our work is that of Bussi, Gorrieri and Zavattaro [8] and Busi and Zavattaro [11] who formalize the semantics of JavaSpaces and discuss the semantics of important extensions such as leasing [10]. However, their work is presented in the context of a pessimistic concurrency control model. Our contribution is a formal characterization of transactions for Linda based on optimistic concurrency that provides scalable and deadlock-free execution.

Other related efforts include the work of Black *et. al.* [3] and Choithia and Duggan [15]. The former presents a theory of transactions that specify atomicity, isolation and durability properties in the form of an equivalence relation on processes. Choithia and Duggan present the pik-calculus and pike-calculus, extension of the pi-calculus that supports various abstractions for distributed transactions and optimistic concurrency. Their work is related to other efforts [11, 5] that encode transaction-style semantics into the pi-calculus and its variants. Our work is distinguished from these efforts in that it provides a simple operational characterization and proof of correctness of transactions that can be used to explore different trade-offs when designing a transaction facility for incorporation into a language. Haines *et.al.* [21] describe a composable transaction facility in ML that supports persistence, undoability, locking and threads. Their abstractions are very modular and first-class, although their implementation does not rely on optimistic concurrency mechanisms to handle commits. Berger and Honda [2] examine extensions to the pi-calculus to handle various forms of distributed computation include aspects of transactional processing such as two-phase commit protocols for handling commit actions in the presence of node failures.

While we have argued that Linda implementations that use optimistic concurrency protocols to subsume transactional behavior are likely to be more robust and exhibit better scalability than pessimistic variants, this argument depends on the way tuples are manipulated by transactions. If many transactions compete for a few tuples by attempting to update them in ways that cannot be expressed under a serializable schedule, aborts are likely to be high. Such applications are likely to exhibit better performance using a pessimistic protocol. Similarly, scalabilty limitations of pessimistic protocols that arise because of locking overheads incurred to support test-for-presence or absence operations are irrelevant if applications do not employ these actions. In an optimistic concurrency protocol, atomic propagation of updates from child to parent transactions when a commit action occurs incurs a cost proportional to the amount of work performed within the child transaction; we intend to explore implementation techniques to quantify these different overheads in scalable distributed environments.

References

1. Roberto M. Amadio, Ilaria Castellani, and Davide Sangiorgi. On Bisimulations for the Asynchronous π-Calculus. In Ugo Montanari and Vladimiro Sassone, editors, *CONCUR '96*, volume 1119 of *LNCS*, pages 147–162. Springer-Verlag, Berlin, 1996.
2. Martin Berger and Kohei Honda. The Two-Phase Commitment Protocol in an Extended pi-Calculus. In Luca Aceto and Bjorn Victor, editors, *Electronic Notes in Theoretical Computer Science*, volume 39. Elsevier, 2003.
3. Andrew Black, Vincent Cremet, Rachid Guerraoui, and Martin Odersky. An Equational Theory for Transactions. Technical Report CSE 03-007, Department of Computer Science, OGI School of Science and Engineering, 2003.
4. Gérard Boudol. Asynchrony and the π-calculus (Note). Rapport de Recherche 1702, INRIA Sofia-Antipolis, May 1992.
5. R. Bruni, C. Laneve, and U. Montanari. Orchestrating Transactions in the Join Calculus. In 13th *International Conference on Concurrency Theory*, 2002.
6. Nadia Busi, Roberto Gorrieri, and Gianluigi Zavattaro. A process algebraic view of Linda coordination primitives. *Theoretical Computer Science*, 192(2):167–199, February 1998.
7. Nadia Busi, Roberto Gorrieri, and Gianluigi Zavattaro. A process algebraic view of linda coordination primitives. *Theoretical Computer Science*, 192(2):167–199, 1998.
8. Nadia Busi, Roberto Gorrieri, and Gianluigi Zavattaro. On the Semantics of JavaSpaces. In *Formal Methods for Open Object-Based Distributed Systems IV*, volume 177. Kluwer, 2000.
9. Nadia Busi, Roberto Gorrieri, and Gianluigi Zavattaro. Process calculi for coordination: From Linda to JavaSpaces. *Lecture Notes in Computer Science*, 1816, 2000.
10. Nadia Busi, Roberto Gorrieri, and Gianluigi Zavattaro. Temporary Data in Shared Dataspace Coordination Languages. In *FOSSACS'01*, pages 121–136. Springer-Verlag, 2001.
11. Nadia Busi and Gianluigi Zavattaro. On the serializability of transactions in JavaSpaces. In *Proc. of International Workshop on Concurrency and Coordination*

(CONCOORD'01). Electronic Notes in Theoretical Computer Science 54, Elsevier, 2001.

12. Nadia Busi and Gianluigi Zavattaro. On the serializability of transactions in shared dataspaces with temporary data. In *Proc. of ACM Symposium on Applied Computing (SAC'02)*, pages 359–366. ACM Press, 2002.

13. N. Carriero, D. Gelernter, and L. Zuck. Bauhaus Linda. In P. Ciancarini, O. Nierstrasz, and A. Yonezawa, editors, *Object-Based Models and Languages for Concurrent Systems*, volume 924 of *LNCS*, pages 66–76. Springer-Verlag, Berlin, 1995.

14. Nicholas Carriero and David Gelernter. Linda in Context. *Communications of the ACM*, 32(4):444–458, 1989.

15. Tom Chothia and Dominic Duggan. Abstractions for Fault-Tolerant Computing. Technical Report 2003-3, Department of Computer Science, Stevens Institute of Technology, 2003.

16. Panos Chrysanthis and Krithi Ramamritham. Synthesis of Extended Transaction Models Using ACTA. *ACM Transactions on Database Systems*, 19(3):450–491, 1994.

17. R. DeNicola and R. Pugliese. A Process Algebra based on Linda. In P. Ciancarini and C. Hankin, editors, *Proc. 1st Int. Conf. on Coordination Models and Languages*, volume 1061 of *Lecture Notes in Computer Science*, pages 160–178. Springer-Verlag, Berlin, 1996.

18. Eric Freeman, Susanne Hupfer, and Ken Arnold. *JavaSpaces principles, patterns, and practice*. Addison-Wesley, Reading, MA, USA, 1999.

19. David Gelernter. Generative Communication in Linda. *ACM Transactions on Programming Languages and Systems*, 7(1):80–112, January 1985.

20. Jim Gray and Andreas Reuter. *Transaction Processing*. Morgan-Kaufmann, 1993.

21. Nicolas Haines, Darrell Kindred, Gregory Morrisett, Scott Nettles, and Jeannette Wing. Composing First-Class Transactions. *ACM Transactions on Programming Languages and Systems*, 16(6):1719–1736, 1994.

22. Kohei Honda and Mario Tokoro. On Asynchronous Communication Semantics. In M. Tokoro, O. Nierstrasz, and P. Wegner, editors, *Object-Based Concurrent Computing. LNCS 612*, pages 21–51, 1992.

23. Nancy Lynch, Michael Merritt, William Weihl, and Alan Fekete. *Atomic Transactions*. Morgan-Kaufmann, 1994.

24. Robin Milner, Joachim Parrow, and David Walker. A calculus of mobile processes, Parts I and II. *Journal of Information and Computation*, 100, September 1992.

25. J. Eliot B. Moss. *Nested Transactions: An Approach to Reliable Distributed Computing*. MIT Press, Cambridge, Massachusetts, 1985.

26. Jaco van de Pol and Miguel Valero Espada. Formal specification of JavaSpaces™ Architecture using μCRL. *Lecture Notes in Computer Science*, 2315, 2002.

27. P. Wyckoff, S. McLaughry, T. Lehman, and D. Ford. T Spaces. *IBM Systems Journal*, 37(3):454–474, 1998.