A Shape Calculus for Biological Processes *

E. Bartocci, F. Corradini, M. R. Di Berardini, E. Merelli, and L. Tesei

Dipartimento di Matematica e Informatica, University of Camerino
Via Madonna delle Carceri 9, 62032 Camerino (MC) - Italy
Email: {ezio.bartocci, flavio.corradini, mariarita.diberardini, emanuela.merelli, luca.tesei}@unicam.it

Abstract. We introduce the main concepts of a bio-inspired calculus for describing 3D shapes moving in a space. Each shape contains a timed CCS-like process specifying an internal behaviour based on shape-dependent channels and splitting actions. Interactions among such entities, called 3D processes, are collision-driven, i.e. collisions are detected and resolved yielding bounces or joining of 3D compatible processes. We discuss how the features of the calculus can be used to model biological processes, for instance biochemical reactions.

1 Introduction and Motivation

In the near future, systems biology will profoundly affect healthcare and medical science. The aim is to design and test “in-silico” drugs giving rise to individualized medicines that will take into account physiology and genetic profiles [5]. The advantages of performing in-silico experiments by simulating the model, instead of arranging expensive in-vivo or in-vitro experiments, are evident. But of course the models should be as faithful as possible to the real system. Since physical concepts like space occupancy, intra-cellular movement, contacts (collisions) and shapes transformation determines biomolecular interactions and therefore cell life, there is the need to provide physical characteristics (shape, weight, size, position) to entities, to collocate them in the continuum space, to allow them to autonomously move and to perceive their spatial neighbour/colliding entities, reacting accordingly to biochemical laws.

Many process algebras have been proposed in systems biology for modelling biological systems [9, 10, 2, 3], accomplishing different kinds of abstractions. The common assumption in these calculi is that the systems are always well-stirred, which means the positions become randomly uniform over a contained volume. This distribution is often generated by several simulation methods [6] based on the theory of stochastic chemical kinetics. When systems are not well-stirred, the ideal way to simulate the time evolution of a chemical system would be to use molecular dynamics, in which the exact positions and velocities of all the molecules in the systems are tracked. In these cases the concept of space plays a

* The authors were partly supported by the FIRB/LITBIO project of the Italian Research Fund.
fundamental role and only recently it has also been taken into account in process
calculi for systems biology. BioAmbients [10] considers space as a set of communicat-
ing compartments, while in Spatial CLS [1] and SpacePI [7] the entities
involved are modeled as spheres situated in space. Time is another fundamen-
tal notion to describe motion. SpacePI [7] proposes an extension of π calculus
equipped with time and space. In this algebra processes can communicate if they
are sufficiently close. No shapes are considered. However, in biochemistry, the
shape of an enzyme plays a very important role in biochemical interactions. The
behaviour or the function of an enzyme is mostly determined by its 3D structure.

The paradigm of our calculus goes towards this direction defining 3D pro-
cesses as entities composed of a 3D shape and a dynamic behaviour. Processes are
situated in a space, move accordingly to specific motion laws (as, for instance, ac-
celeration in a gravitational field or Brownian motion), collide and possibly bind
with others processes becoming compound 3D processes. The binding depends
on channels \((a, X)\), derived from classical CCS channels, where \(a\) is the channel
name and \(X\) is a certain region on the surface of the 3D process in which the
channel is “active”. The binding corresponds to communication on these chan-
nels. It occurs if and only if the surface of collision of two 3D processes belongs
to active channels of both processes and the names of the channels are co-names
à la CCS. Compound processes can split weakly, by non-deterministically releasing
a previously established bond, or “react”, by splitting urgently in as many
pieces as the products of the reaction are. If communication (i.e. binding) does
not occur, the collision of two 3D processes is considered elastic, i.e. the shapes
bounce and proceed independently. In the next sections we provide the main
ideas of our calculus.

2 The calculus

We introduce the syntax and sketch the semantics of a calculus suitable to de-
scribe 3D processes with shapes moving in a space. We give some intuitions
on how the calculus can be used to model cellular processes, e.g. biochemical
reactions.

2.1 3D Shapes

We consider a global coordinate system in a three dimensional space represented
as \(\mathbb{R}^3\). We describe a basic three dimensional shape in a certain instant of time
by a tuple \((V, m, p, v)\) where \(V \subseteq \mathbb{R}^3\) are the points of the basic shape, \(m\) is
the mass of the shape, \(p\) is the centre of mass of the shape and \(v\) is the vector
representing the current velocity of the shape. \(V\) must be a measurable set with
non-zero volume and must be a closed, limited and convex set of points. We
mostly use simple basic shapes such as polyhedra, spheres, cones and cylinders
as we can represent them by suitable and efficient data structures (see [4]) and
they are handled by the most popular algorithms for motion simulation and
collision detection.
Two (basic or compound) shapes $S_1$ and $S_2$ can be combined into a compound shape represented by $S_1 \langle X \rangle S_2$ provided that they have a non-empty, zero-volume, connected boundary fraction $X$ in common, i.e. they physically touch, but they do not interpenetrate. Moreover, it is required that they have the same velocity at any instant in time. 3D shapes can move in space according to the given velocity. The motion of a shape is a trajectory determined by a function $\text{move}$ that updates the shape velocity over time. A well-formedness notion can be defined to guarantee that the structural constraints of shapes and their space consistency are continuously respected.

Collision detection [4] is a well-studied problem in the field of computer games. Very efficient techniques exist to determine - starting from a consistent state, i.e. the shapes are not interpenetrating - the first time of contact $t$ of a given set of moving shapes, i.e. the maximum time that can pass without a collision, together with the identification of those shapes that will be colliding at time $t$. The interface of a generic collision detection system can be formalized and used as a function in the definition of the semantics of our calculus.

2.2 3D Processes

In the biological metaphor, the shape of entities - for instance proteins - is deeply related to their behaviour. Usually, interactions take place because a collision between two (or more) entities occurs and the involved entities have specific compatible ligands on the surface of collision. In this case, the entities bind and - after internal activities - react, i.e. transform and split becoming different entities with possibly other behaviours and ligands. Otherwise, they bounce and continue to move freely in the space.

To represent this scenario, we associate a behaviour to a shape in order to form a 3D process. A (basic) 3D process $S[B]$ is composed of a timed CCS-like $B$ together with a shape $S$. Unlike from CCS [8] process, channels are of the form $(a,X)$. The channel name $a$ can be thought as a “type of binder” and the set of 3D points $X$ is the surface of $S$ in which the channel is active, i.e. the binding of type $a$ is possible there. 3D processes are represented by the terms:

$$ P ::= S[B] \mid P \langle (a,X) \rangle P $$

Communication can only occur when two 3D processes $P_1$ and $P_2$ collide and have, on the surface of collision, active corresponding channels, i.e. $P_1 \xrightarrow{(a,X)} P_1'$, $P_2 \xrightarrow{(a,X)} P_2'$ and $X = X_1 \cap X_2 \neq \emptyset$. The communication produces a $\tau$ action (urgent in a Timed CCS-like setting) and the resulting process is the compound 3D process $P_1' \langle (a,X) \rangle P_2'$. The compound process has, as behaviour, the parallel composition of the behaviours of $P_1$ and $P_2$, and - as associated shape - $S_1 \langle X \rangle S_2$ where $S_1$ is the shape of $P_1$ and $S_2$ is the shape of $P_2$. The compound process becomes a new process that can freely move in space and possibly bind with other processes if other communications are possible. The full syntax for the behaviours is the following:

$$ B ::= \text{nil} \mid (a,X).B \mid \omega(a,X).B \mid \tau.B \mid \rho(L).B \mid \epsilon(d).B \mid B + B \mid K $$

where $L$ is a finite set of channels, $K$ is a process variable and $d \in \mathbb{R}_{\geq 0}$ is
a time delay. Note that inside a compound process interactions of type $\omega$ and $\rho$ are possible. A communication between $\omega(a,X)$ and $\omega(\overline{a},Y)$ produces an $\omega$ (non-urgent action) and causes a compound processes to physically split from a previously established bond $(a,Z)$, $Z = X \cap Y$, and continue as two new independent 3D processes. This kind of communication is needed to represent a non-deterministic non-urgent split of processes that is typical, for instance, in biochemical reactions. Finally, a communication of type $\rho(L)$ is used to model a reaction, i.e. an urgent (produces $\tau$) split of all the existing bonds of $L$ where the pieces represent the products of the reaction and are new 3D independent processes.

The final syntactic level of our calculus is that of networks of 3D processes $N ::= P \mid N \parallel N$ that are collections of 3D processes moving into space, colliding, binding and splitting. A proper operational semantics can be defined to precisely describe the evolution of a network of 3D processes taking into account collision detection, weak splitting ($\omega$) and reactions ($\rho$). The resulting observables are timed traces of the network configurations.

3 Conclusions and Future Work

We are working in defining the semantics in a detailed way, in studying interesting properties of the calculus, in comparing its expressiveness with other process algebra for systems biology, and in applying it to non-trivial case studies.

References