Computerised Modelling for Developmental Biology

Bertens, L.M.F.

CHAPTER 7. INTEGRATION OF MODELLING TECHNIQUES

In the previous chapters we have addressed several case studies of modelling techniques, applied to processes in developmental biology. The development of the heart in the turtle *Emys orbicularis* was studied using 3D reconstructions in chapter 2, followed by a detailed study of the outflow tract development, using 3D reconstructions based on high resolution images in chapter 3. Subsequently, the use of ontologies in describing vertebrate anatomy, development and physiology was examined in chapter 4. Finally, gradient formation was modelled using the modelling framework of Petri nets. Both a conceptual, qualitative approach was presented, in chapter 5, and a more concrete, quantitative approach, incorporating aspects of differential equation models, in chapter 6.

The used techniques each have their own particular merits, but also their own particular limitations. The visual reconstructions gave us insight into the three-dimensional structure of the developing turtle heart, but did not tell us much about its specific physiology. The Petri net model of gradient provided us with a way to simulate gradient formation, but did not inform us about the visual appearance of the resulting developmental patterns. In order to extend the functionality of these modelling techniques and try and overcome some of these limitations the goal of the research has been twofold. In the case studies, in particular in chapters 2 and 3, developmental processes or structures have been studied within a biological context. In this way they contributed to our understanding of developmental biology. But at the same time, for each of the cases an integration of modelling techniques was strived after, in order to extend the methodology used in models of developmental biology. With the increase in data production in the life sciences, due to *e.g.* high throughput techniques, novel methods are needed for analysis. Here integration plays an important role, both of data and of methods. The rising interest in these types of integration is illustrated by the publication of a journal specifically dedicated to this topic, the *Journal of Integrative Bioinformatics*. This journal covers data integration, method integration and combinations of both, in particular for the fields of molecular and systems biology. In general, data and method integration are currently focused on the field of systems biology (Hoehndorf *et al.*, 2011; Takai-Igarashi, 2005; Zhang and Verbeek, 2010), dealing mainly with pathway modelling. However, the field of higher level developmental biology (*i.e.* on tissue and organ level) also stands to gain from integration techniques. Method integration is in this case more relevant than data integration, since data increase is less prominent in higher level developmental biology. It is for this reason that we have investigated method integration for this field by means of the presented case studies.

For each of the separate studies the biological context and relevance has been addressed in the corresponding chapters. This concluding chapter will focus on the ways in which integration of modelling techniques has been used or proposed to increase the (accessibility of) knowledge contained within the model system. In chapter 1 four categories of models were presented (verbal, visual, algorithmic and equation based) and the current chapter is based on this same distinction. For each category, integration methods have been investigated, albeit not yet fully implemented or presented for all cases in the previous chapters. Here I present a short overview, looking at these combinations one by one. Additionally, in the case studies we have looked at combinations of models of the same type, for different purposes. These are discussed in section 7.5. At the end of the chapter a table is provided, in which an overview is given of all combinations described here.

7.1 VERBAL AND VISUAL: COMBINING ONTOLOGIES WITH DIAGRAMS AND 3D MODELS

In chapter 4 the functionality of ontologies for developmental biology has been discussed. We have looked at new ways to model knowledge of biological structures using ontologies. One way of doing this was to change the focus of the organization of the captured knowledge, from one particular species to the group of vertebrates in general. Another proposed way of extending the functionality was by integrating other ways of presenting the data. This has been done by including both visual and algorithmic functions, of which the latter will be discussed in the next paragraph.

Two ways of visualising the knowledge can be used for the vertebrate heart ontology. First of all the information captured in the object and data properties (linking classes and instances to one another or to pieces of data respectively) can be visualised in diagrams, as mentioned in 1.3.2. In these **visualisations of ontology information in diagrams**, classes and instances are represented as nodes and the connections between these as edges; these diagrams can be categorized as class diagrams. As can be seen from the blood flow diagram presented as an example in Fig. 7.1, this way of visualising the data helps the user get an overview of information contained in the ontology. When only taking into account the verbal information in the ontology, it is hard to get the full picture of for instance blood flow. Diagrams such as these can be produced for all properties modelled in the ontology. They are particularly useful for our dynamic ontology, in which different physiological situations are modelled; a diagram can provide insight into a particular situation of interest.

A second way in which visual models and ontologies can be combined is by **visualising results of ontology queries in 3D reconstructions**, as discussed in 4.4.4, *cf.*

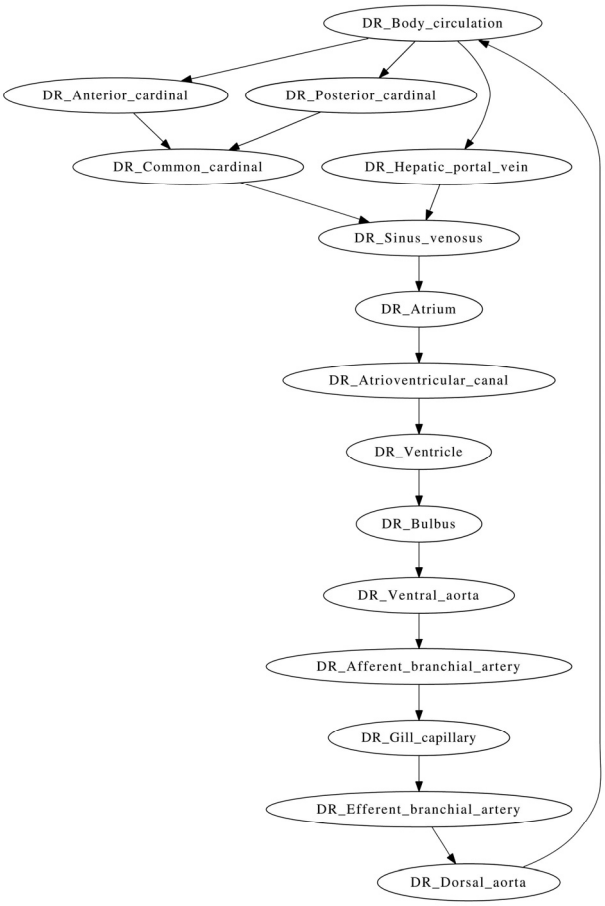

Figure 7.1. Diagrammatic representation of the blood flow information for *Danio rerio*, contained in the ontology system, *cf.* chapter 4.

Fig. 4.4. Combining these methods has benefits both for the 3D reconstructions and for the ontology. By linking the 3D reconstructions to an ontology one is forced to use a standard vocabulary when annotating the structures, in order to identify these with classes and instances in the ontology. This standard vocabulary makes it easier to compare and study different reconstructions and helps prevent miscommunication. The overview of turtle heart development (provided on the website http://bio-imaging.liacs.nl/galleries/), as presented in chapter 2, illustrates this. At the same time, combining these methods also enriches the ontology. A query of the ontology results in a list of names of relevant structures and/or instances and visualising these in a 3D reconstruction provides the user with more information. Other examples of combining ontologies with 3D reconstructions exist (*e.g.* Köhn *et al.*, 2004) and in chapter 4 our particular method and its benefits have been explained.

7.2 VERBAL AND ALGORITHMIC: COMBINING ONTOLOGIES WITH PETRI NETS

The vertebrate heart ontology has been developed with future extensions in mind, in particular with a complementary Petri net model. The information on context-dependent heart physiology, modelled using object and data properties in the ontology, lends itself for modelling in a process oriented method like Petri nets. Our aim is therefore to **connect the ontology to a Petri net simulating the blood flow**, in which the Petri net takes information on blood pressures and shunting from the ontology.

Currently we are working on new tools for the simulation and analysis of Petri nets, for instance using mathematical software. Several tools exist for the implementation and simulation of Petri nets, *e.g.* Snoopy, CPNTools and Pipe2, each restricted by certain limitations. Developing a more generic tool, which allows a range of input file types for the Petri nets and gives the user control over arc functions and firing rules, we hope to overcome these limitations; one of the strengths of Petri nets is the ease with which its theory can be extended to include *e.g.* new arc types or firing rules. This makes the modelling framework easily amenable to particular needs of a research field like developmental biology. In response to the expansion of Petri net theory (*e.g.* with Hybrid Petri nets), existing tools are constantly being adapted and new tools are developed. A tool which allows the user direct access to the theoretical basis will create greater freedom in modelling complex biological processes.

In this way we hope to be able to extract variable values from the ontology and use them in the weight functions of the corresponding Petri net. Consequently, we would combine the semantic properties of ontologies with the ability to model (concurrent) processes of Petri nets.

Combinations of Petri nets and ontologies have been used and reported sporadically for other fields (*e.g.* Takai-Igarashi, 2005; Recker and Indulska, 2007), but there the focus has been different; ontologies were used in these cases to establish a standard 'language' for Petri net components within a particular research area (*e.g.* signalling pathways; Takai-Igarashi, 2005). This systematization of Petri net semantics enables researchers to communicate and share Petri net models more easily. To date no combinations of Petri nets and ontologies have been presented in literature for the field of developmental biology and in particular for physiological processes. Furthermore no methods have been described in which simulation of a Petri net model directly relies on information provided by the ontology in a dynamic way, such as proposed for the vertebrate heart ontology. Developing such a method will help extend the functionality of both ontologies and Petri nets for the field of developmental biology.

7.3 ALGORITHMIC AND VISUAL: COMBINING PETRI NETS WITH 3D MODELS

In chapters 5 and 6 the problem of modelling gradient formation is addressed and a solution is presented with Petri nets. This model serves as a proof of concept for the use of Petri nets for higher level developmental biology. Originally, the biological process selected as a case study was the formation of the anterior-posterior axis (AP-axis) in the frog *Xenopus laevis*. This process comprises several smaller processes, one of which is gradient formation. The model of gradient formation has been constructed with the underlying future aim of combining it with other modelled elements of the AP-axis formation, in order to obtain an all-encompassing Petri net model of the process (this is addressed in more detail in 7.5).

An important aspect of the process of AP-axis formation is the expression of a series of *Hox* genes, which play a pivotal role in the development of all vertebrates. Changes in the order of the *Hox* gene expression in this process can lead to severe developmental aberrations (Wolpert, 2002; McNulty *et al.*, 2005; Stern *et al.*, 2006; Iimura and Pourquié, 2007). The expression patterns of the *Hox* genes as seen in embryonic development can be visualised by staining the gene products using the technique of (fluorescent) *in situ hybridization*. By performing experiments which interfere with normal AP-axis formation (for instance by knocking out *Hox* genes, *cf.* McNulty, 2005; or by removing the organiser mesoderm, *cf.* Jansen *et al.*, 2007), anomalous patterns can be obtained that provide us with new insights in the process. Ideally, a Petri net model of the entire process would be able to reproduce and possibly predict experimental results. When executing the Petri net, it would be helpful if users were provided with visual information about the process in addition to the abstract information on token movement. Therefore we would like to **connect states of the Petri net model to 3D reconstructions with corresponding *Hox* gene expression patterns**.

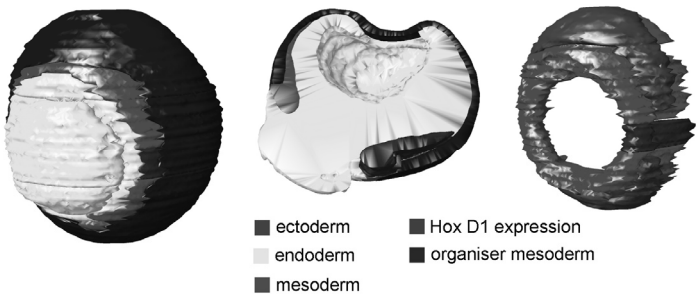

Figure 7.2. 3D reconstruction of a stage 11 *Xenopus laevis* embryo, showing the three germ layers and the Spemann organiser, as well as the expression pattern of *Hox* gene D1.

To this end we have, in collaboration with the Institute of Biology Leiden (IBL), obtained several sectioned *Xenopus laevis* embryos of different developmental stages in which different *Hox* genes have been visualised by *in situ hybridization*. From section images of these embryos, 3D reconstructions have been created (using the methods described in chapter 2), of which an example is shown in Fig. 7.2. Since in situ hybridization is only possible for two (or sometimes three) genes in one experiment, we cannot visualise expression patterns for more than two genes at a time in one embryo. Therefore work is being done on developing a template model, onto which multiple 3D reconstructions of *Hox* gene expression patterns can be projected simultaneously. The end result of this will allow us to produce visual information on many different situations of gene expression, corresponding to particular states of the Petri net (and thus to particular states in the biological process). In future we hope to connect these visualisations to the Petri net model itself, *i.e.* we would like for the model to present visual output whenever a state is reached for which this information is available.

7.4 EQUATION BASED AND ALGORITHMIC: COMBINING DIFFERENTIAL EQUATIONS WITH PETRI NETS

In chapter 6 we have combined the Petri net model of gradient formation with differential equation models of the same process. As described in 1.4, algorithmic and equation based models differ significantly and both have particular merits. Equations allow precise quantitative analysis of the continuous behaviour (both spatially and temporally) of a process, while an algorithmic approach is better suited to study particular and possibly concurrent events in the process. By **incorporating parameters from differential equations in a Petri net model**, we hope to attain exact analysis and insight into complex processes consisting of different types of events.

While differential equation models describe a process in a continuous manner but can be discretised in time and space, Petri nets start off by describing changes in the system states in discrete space and time steps. Petri nets allow one to model concurrent events and hierarchical coherence of processes, while a hierarchical approach for differential equations, using multiscale modelling, is feasible but quite involved. Thus the two modelling approaches each hold particular advantages for modelling in developmental biology and combining aspects opens up new possibilities in the simulation and analysis of complex developmental processes. The current Petri net model of gradient formation has been designed to accommodate the parameters of existing differential equation models and has been validated using data obtained from these equations (described in detail in chapter 6).

In future we hope to be able to directly implement differential equations in the Petri net, by enabling the weight functions of arcs to include more complex formulas. The current work on simulating Petri nets using mathematical tools (*cf.* section 7.2) is specifically focused on this. This will result in a hybrid model, which will allow more complex computation, but will still operate in discrete temporal and spatial steps.

7.5 INTEGRATING MODELS OF THE SAME MODELLING APPROACH

In addition to integrating different types of models, each of the case studies has also featured combinations of models of the same modelling approach. The vertebrate heart ontology system, described in chapter 4, is deliberately referred to as an ontology *system*, instead of merely an ontology, since this model is a **system of three combined ontologies**. We constructed an ontology for anatomical information and one for developmental information and used the NCBI database ontology for taxonomic information. By separating these different types of information the anatomy and development ontologies could be used on their own, increasing their functionality (as explained in 4.3). Furthermore, using the NCBI database provided us with a standardized way of addressing and including species information which increased the inter-operability of the system. While all the information currently present in the system could have been modelled in one all encompassing ontology, dividing information over separate ontologies provides the system with more freedom for future expansion and integration.

Secondly, 3D reconstructions of cardiogenesis in the turtle *Emys orbicularis*, presented in chapters 2 and 3, were combined to form a developmental series, illustrating essential stages in this process in a chronological order. In addition to these models of cardiogenesis, several other series of models have been produced during the research in its entirety, for instance the *Xenopus laevis* models, mentioned in section 7.3, and series of zebrafish gene expression models of different developmental stages. These models are all presented in an **online atlas of 3D reconstructions** (Potikanond, 2012) which can be browsed and queried, allowing users access to all information in the models as a whole, as opposed to letting them study the models one at a time. Incidentally, this approach is also related to modelling with ontologies, since the database organizes the models and their annotations in a structural way that can be examined in context and queried.

Another way in which 3D reconstructions could be combined is by **linking low resolution 3D reconstructions with high resolution 3D reconstructions** (*i.e.* models constructed from low and high resolution images, respectively). In chapters 2 and 3 we have presented models of the entire turtle heart in low resolution and detailed models

of the outflow tract in high resolution. Apart from the biological insights gained from these models, a second purpose of this approach was to integrate these models in a hierarchical manner. The high resolution models provide a detailed view of one particular structure, the outflow tract, but at the cost of losing the overview of the entire heart and the position of the outflow tract in it; on the other hand, models of the entire heart cannot provide the level of detail necessary to answer particular questions about the development of the outflow tract. The in-house software, 3D acq, used to capture the section images, stores the stage coordinates of the microscope for each image in a database (*cf.* chapter 3), making it possible to relate high and low resolution images of the same section on a pixel level. This in turn can be used to connect structures in an all encompassing low resolution model to high resolution models of these structures. Ideally the visualisation tool would include a function allowing the user to select certain biological structures and zoom in on these.

A final modelling approach for which integration has been investigated concerns Petri net models. As mentioned in 7.3, gradient formation is an important part of AP-axis formation in the species *Xenopus laevis*. In future we aim to construct additional models for the other events in the AP-axis formation and combine these models and to **combine Petri net models of separate processes in an all encompassing model of the entire process**. The other subprocesses of AP-axis formation include *Hox* gene expression, cell layer movement and vertical signalling. We would like to design separate Petri net models for each of these subprocesses, based on the same assumptions and modelling decisions as the current model.

The models used to construct an all encompassing model would be of the same biological level (the same level of magnification as it were). However, as with the 3D reconstructions, a second way of integrating Petri nets is possible, by using models on different levels, creating as it were a possibility to zoom in (*cf.* combining low and high resolution 3D reconstructions). In the case of Petri nets this would amount to **a hierarchical Petri net**, which allows one to refine particular events in the net in more detailed subnets. In analogy with (microscopic) imaging techniques, it would correspond to changing resolutions through different imaging modalities. This can be achieved by identifying a transition or place with a complete net on a lower level of refinement; when for instance such a transition fires, the net inside is activated and generates output to the transition.

The model of gradient formation, presented in chapter 6, is amenable to this approach. Here the process of morphogen dissipation through the tissue has been treated as ‘effective diffusion’, corresponding to the differential equation model presented by Kicheva *et al.* (2007); no distinction is made between different spreading mechanisms, such as active and passive diffusion or endocytosis. For the current model this generalization suffices, but in case where more detailed information is available this

could be added to the net by means of a subnet specifying all transitions t'_i and t''_i. Similarly, the process of endocytosis, which accounts for the degradation of morphogens, could be specified by connecting subnets to all transitions p'_i. Due to the modular nature of the system only one subnet design is required for each of these detailed mechanisms; general nets for endocytosis and diffusion will serve as building blocks and can be added to each of the cells (or pairs of neighbouring cells).

	VERBAL ONTOLOGY	VISUAL 3D MODEL	ALGORITHMIC PETRI NET	EQUATION BASED DIFFERENTIAL EQUATION
VERBAL ONTOLOGY	system of combined ontologies of the vertebrate heart	- diagrammatic visualisation of information - visualisation of ontology query results	- process model of blood flow	
VISUAL 3D MODEL		- online atlas of 3D reconstructions - linking low and high resolution models	- *Hox* gene expression reconstructions linked to PN	
ALGORITHMIC PETRI NET			- all encompassing model of AP-formation - hierarchical net with specification of diffusion and degradation mechanisms	- Petri net of gradient formation incorporating parameters from differential equations

Table 7.1. Overview of ways in which the modelling techniques, presented in this dissertation, were integrated. Elements in black have been realised, elements in gray are planned and have been taken into account in the modelling decisions of the current models.

This concludes the discussion of integrative modelling methods, used in this research, and the presentation of the research as a whole. In order to investigate and further the use of modelling techniques for developmental biology, four modelling approaches were selected, each on the basis of particular merits for modelling phenomena in developmental biology and each representing a different category from chapter 1. This has led to a broad scope of the study. The 3D reconstructions underlined the significance of spatial information and high resolution imaging in understanding complex developmental anatomy. Secondly, developmental, anatomical and physiological information was structured formally in the vertebrate heart ontology, in order to investigate the uses of formal organization of knowledge for this particular field. Thirdly, the Petri net framework was chosen for its ability to model the dynamics of complex processes, including modular and hierarchical elements, which are particularly relevant to developmental biology. And finally, by combining this with aspects from differential equations commonly used to model developmental processes such as gradient formation, the prospects for simulation and analysis were widened. In each of these projects, the integration of modelling methods has been taken into account as well, as described in this chapter. The combinations of the techniques, used in this research, are summarized in Table 7.1, which includes both projects that have been realised and projects that are planned for future work, as discussed above.

The modelling approaches, discussed in this dissertation, have thus provided us with novel ways to modelling biological phenomena, as well as new insights into developmental biology and have laid the groundwork for integrating diverse modelling methods for developmental biology.