Componential Explanation in Philosophy, Cognitive Science and Computer Science

Tibor Bosse (tbosse@cs.vu.nl), Catholijn M. Jonker[1] (C.Jonker@nici.ru.nl), Jan Treur[2] (treur@cs.vu.nl)
Vrije Universiteit Amsterdam, Department of Artificial Intelligence
De Boelelaan 1081a, 1081 HV Amsterdam, The Netherlands

Abstract

This paper shows how Componential Explanation as discussed within Cognitive Science and Philosophy of Science relates to Compositional Verification in Computer Science. It is shown how formal techniques and methods developed for Compositional Verification provide a formal basis and automated support for Componential Explanation. The role of formalised interlevel relations is shown to be crucial for formalisation of the analysis on which a componential explanation rests. A case study is used to illustrate the thoroughness of the approach.

Introduction

The notion of componential explanation plays a role in different disciplines such as Philosophy, Biology, Cognitive Science, Computer Science and AI. Roughly spoken, *componential explanation* describes how properties of a system that is organised according to a number of components, can be explained from properties of the components and their interactions. For componential explanation, Clark (1997) draws the analogy with modelling and analysis methods within AI, referring to, among others, Newell and Simon (1972) and Dennett (1978).[3] He also claims that componential explanation has a role to play in less classical AI areas such as connectionist approaches: in advanced connectionist work, complex tasks require highly structured multi-layer networks.[4] Clark (1997) gives suggestions, but does not address in more detail how to formalise componential explanation. This is the subject of the current paper. To this end methods developed originally in Computer Science are considered.

The area within Computer Science in which properties of component-based systems are analysed in terms of properties of their components is called *compositional verification*; e.g., Roever et al. (1998, 2001), Jonker and Treur (2002). Formalisation and automation are important in the contributions to this area. The considered (software and hardware) systems are assumed to be hierarchically structured according to a number of aggregation levels. A central role is played by *interlevel relations* between properties at different levels of aggregation. For example, for a system S with property G that consists of two components A and B that have properties DP1 and DP2, respectively, the implication DP1 & DP2 & T $\Rightarrow$ G is an example of an interlevel relation expressing that S has property G in virtue of connectivity T and properties DP1 and DP2 of components A and B. Here the connectivity property T denotes a property that describes the connection or interaction between the components: transfers between the components. Compositional verification analyses properties of systems based on such interlevel relations.

In this paper it is explored how the notion of compositional verification developed within Computer Science relates to the notion of componential explanation as developed within Philosophy and Cognitive Science (cf. Cummins, 1975, 1983; Clark, 1997; Davies, 2001), and how it can be used to obtain a formalisation of componential explanation in a more technical sense, opening doors to existing or new software tools to support the explanation process. First the notion of componential explanation is briefly described. Next compositional verification is summarised, and it is shown by a case study on the circulatory system, how the notions relate to each other.

Componential Explanation in Philosophy

Hempel (1959) and Nagel (1961) focus on functional explanations why certain items I (such as the heart) are present within an organised system S (e.g., a human being). They base the explanation on an attempted form of deduction, concluding that the item I is necessary in the context of the overall system S (for a certain function F). In this line of reasoning the existence of functional equivalents is problematic: why would another item I' with the same functional contribution F not be possible instead? The dilemma is that:

- either functional equivalents exist, then the necessity of the existence of an item cannot be claimed deductively,
- or the necessity of the existence of an item can be claimed deductively, but functional equivalents are not allowed.

Hempel (1959) takes the first horn of this dilemma, Nagel (1961) the second one. Hempel's explanation does not provide a deductive argument. Nagel's is deductive, but requires a premise excluding the existence of functional equivalents, which is problematic (since there are no laws to derive it).

Cummins (1975) avoided this dilemma by a change of perspective. Instead of attempting to obtain a deduction concluding the existence of a certain item I, his deductive analysis A aims at concluding the systemic capacity C of the

[1] Currently at: Radboud Universiteit Nijmegen, Nijmegen Institute for Cognition and Information, The Netherlands.

[2] Part of this work was performed as part of a position at Utrecht University, Department of Philosophy, The Netherlands

[3] 'Modular programming methods in classical AI lent themselves quite nicely to a componential form of explanation. In attempting to understand the success of such a program, it is often fruitful to isolate the various subroutines, modules, etc. and to display their role in dividing the target problem into a manageable series of subproblems.'
Clark, (1997, pp. 104-105)

[4] 'In such cases it is possible to advance our understanding of how the system succeeds by asking after the roles of these gross components (layers and subnets).' Clark, (1997, p. 105)

overall system S, on the basis of properties of the components of S. Within this analysis A the item I contributes function F. This function F is needed in A in the sense that, if it would be left out of A, capacity C cannot be deductively concluded anymore. Davies (2001, Chapter 2, pp. 25-27), discusses Cummins' account on componential explanation, also called *systemic functional analysis*; see also Clark (1997, Ch. 6). The idea is as follows. For a system S, one of its capacities C can be analysed: by virtue of what does S exercise C? For example, the capacity C of an animal to stay alive can be analysed in terms of different components within the animal and the jobs they perform: e.g., circulation, digestion, respiration.

According to Davies' analysis, first the subsystems performing such jobs are identified, and the relevant capacities specified. For example, within Biology the circulatory system contributes to C by a capacity C' to transport oxygen and nutrients to the places within the animal where they are used. A next level of functional analysis focusses on a capacity of one of these subsystems, for example the capacity C' of the circulatory system. Considering the next level, the analytical approach also needs to be performed for this subsystem, i.e., identification of the main components and the jobs they perform. Example capacities for this system are assimilation of oxygen and nutrients in the blood, propulsion of blood, and absorption of oxygen and nutrients. The heart is one of the contributing components for these capacities; in the context of capacity C' it can be attributed the (systemic) function F of pumping blood. After presenting a brief overview of Cummins' account, Davies (2001, Chapter 4) presents his own account on componential explanation. A main addition is that the phenomena analysed are *hierarchically organised*:

Let A denote the analysis of system S into its components, and C the systemic capacity analysed. The item I within S has systemic capacity function F if and only if:

- (i*) I is capable of doing F
- (ii*) A appropriately and adequately accounts for S's capacity to C in terms of the organised structural or interactive capacities of components at some lower level of organisation
- (iii*) I is among the lower-level components cited in A that structurally or interactively contribute to the exercise of C
- (iv*) A accounts for S's capacity to C, in part, by appealing to the capacity of I to F
- (v*) A specifies the physical mechanisms in S that instantiate the systemic capacities itemised

Here (i*), (iv*), and (v*) are items of Cummins' account, and (ii*) and (iii*) are adding hierarchical organisation. Clark (1997)[5] considers componential explanation ('from parts to wholes', pp. 103-105) as a major explanatory strategy, to be used in conjunction with other types of explanation (for example, based on reciprocal input thought-action cycles, pp. 105-106), to explain interaction with the environment.

Compositional Verification

In this paper the formalization of Jonker and Treur (2002) of compositional verification for Computer Science and Artificial Intelligence, summarized in this section, is used as starting point for the formalization of componential explanation. Within software engineering, the purpose of verification is to prove that, under a certain set of assumptions, a system will adhere to a certain set of properties, for example the design requirements. In this approach, verification is accomplished by a formal analysis of relations between properties and assumptions that respects the levels of aggregation already present in the compositional structure of the system.

A component-based system can be viewed at different levels of aggregation. Viewed from the top level, denoted by L_0, the complete system is one component S. At the next lower level of aggregation, level L_1, the system component S is a composition of components, and connections between these sub-components. Each component is again composed of its sub-components, and so on, until the lowest level of aggregation is reached, in which components are no longer composed of other components: primitive components.

The primitive components can be verified using dedicated verification methods, such as described in, e.g., (Leemans, Treur and Willems, 2002). Verification of a composed component is done using properties of the sub-components it embeds, and environmental properties of the component (i.e., assumptions on its embedding in the rest of the system). Given a set of environmental properties, the proof that a certain component adheres to a set of properties depends on the properties of its sub-components, and properties of the interactions between those sub-components. The compositional verification method can be formulated in more detail as follows:

A. Verifying one Aggregation Level Against the Other

- Determine which properties are of interest (for the higher level).
- Determine which assumptions (for the lower level) and which environment properties guarantee the higher-level properties.
- Prove the higher-level properties on the basis of these assumptions for the lower level and environmental properties.

B. The Overall Verification Process

- Determine the properties that are desired for the whole system.
- Apply procedure **A** iteratively until primitive components are reached.
- Verify the primitive components using techniques specialised for the type of component.

The results of verification are a hierarchy of properties at the different aggregation levels, and the logical relations between the properties of different aggregation levels, see Figure 1. In the picture, $P^j_{t,m}$ is the set of properties or assumptions of a component labeled j belonging to aggregation level L_t. This set is used in the proof for a component labeled m that is part of aggregation level L_{t-1}. Let $P_{t,m} = \cup_j P^j_{t,m}$. Then, the hierarchy is constructed such that

[5] (1) An account of the gross behaviors of the well-functioning organism in the environment - an account that may invoke collective variables whose componential roots span brain, body, and world.

(2) An account that identifies the various components whose collective properties are targeted by the explanations proper to (1). Two important subtasks here are to identify relevant neural components and to account for how these components interact.

(3) An account of the varying information-processing roles played by the components (both internal and external) identified in (2) – an account that may well assign specific computational roles and representational capacities to distinct neural subsystems. Clark (1997, p. 126).

$P_{t,m} \Rightarrow P^{m}_{t-1,r}$, for some r, being the label of the parent component of m on level L_{t-2}.

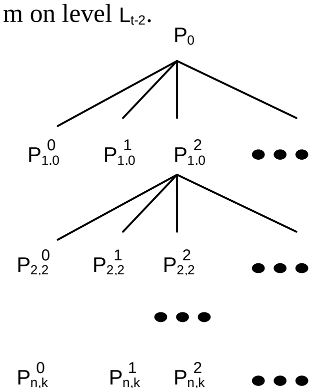

Figure 1: Hierarchy of properties for compositional verification.

Case Study: Circulatory System

In this section, a case study in the domain of the circulatory system in mammals is used to illustrate how the philosophical idea of componential explanation can be worked out using the methods in compositional verification within Computer Science. This case study is often used as an example in philosophical literature. The analysis of the system's capacities in the case study is described in terms of *dynamic properties*: temporal statements that relate different states of a system (at different time points) to each other. Such dynamic properties are identified at different aggregation levels. Next, *interlevel relations* are established, relating dynamic properties at different levels to each other. The properties have been formalised using the Temporal Trace Language TTL introduced in Jonker and Treur (2002) (see also Bosse et al., 2006); for reasons of readability most of them are presented here in semiformal form. It is shown how this analysis can be used to obtain a componential explanation according to Cummins' and Davies' perspective.

The circulatory system (see Figure 2) takes care of a number of capacities, such as providing nutrients and oxygen to the body and taking waste (e.g., CO_2) out of the body; e.g., Noordergraaf (1978), Rideout (1991). The main property to focus on in this example is that the system provides oxygen for all parts of the body. The organisation of the circulatory system S is analysed as consisting of the following active components that (by showing their specific behaviours) all play their roles within the overall process: heart, capillaries in lungs and other organs, arteries (pulmonary artery channels, from the heart to the capillaries in the lungs; aorta channels, from heart to the capillaries in the body), veins (pulmonary veins, from the capillaries in the lungs to the heart; inferior and superior vena cava, from the capillaries in the body to the heart). These active components work together due to a structure as depicted in Figure 2.

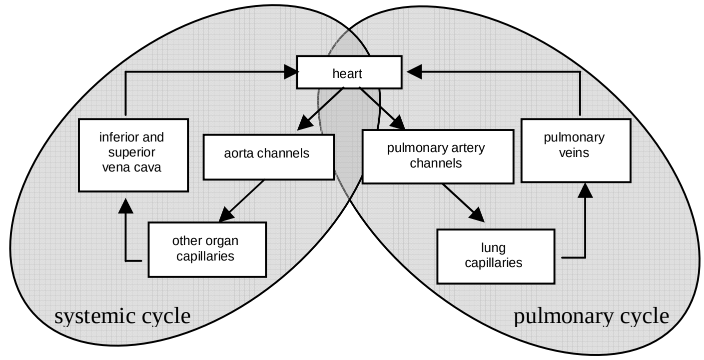

Figure 2: Schema for the circulatory system.

In the next sections, a number of dynamic properties relevant for the analysis of the system's capacities are presented. In particular, the following properties are shown (all related to oxygen supply):

- environmental assumptions
- dynamic properties specifying component capacities
- dynamic properties for interaction between components (transfers)

At the top level, the system can be seen as one component. At lower levels, properties of sub-components can be identified, as well as properties of transfers between these sub-components. The lowest level comprises properties of primitive components and transfers between them.

Environmental Assumptions

The following environmental assumptions are considered:

EA1(i) Heart Stimulus Occurrence (with maximal interval i)
For any point in time t there exists a time point with $t < t' \leq t + i$ such that at t' a heart stimulus occurs. Formalisation in TTL:
$\forall t\ \exists t'\ [\ t < t' \leq t + i\ \&\ \mathrm{state}(\gamma, t') \models \mathrm{stimulus\_occurs}\ ]$

EA2(w_{init}) Heart Initialisation
There exists a time point t with $0 \leq t \leq w_{init}$ such that at t the heart generates a fluid volume V with any ingredients I

EA3 Oxygen Availability
At any point in time t oxygen is present in the air within the lungs.

EA4 Carbonacid Availability
At any point in time t carbonacid is present within the organs.

Here V is an amount of fluid, and I is a specification of ingredients (a list of them).

Component Capacities

Below, the properties of components at the different aggregation levels are discussed.

Component Properties at Aggregation Level 0

At the top level (level 0), it is expressed that oxygen is successfully provided to the organs:

CP0(d) Oxygen Delivery Successfulness (with maximal interval d)
For any point t there exists a time point t' with $t < t' \leq t + d$ such that at t' oxygen is delivered to the organs. Formalisation in TTL:
$\forall t\ \exists t'\ [\ t < t' \leq t + d\ \&\ \mathrm{state}(\gamma, t') \models \mathrm{oxygen\_delivered}]$

Component Properties at Aggregation Level 1

At one aggregation level lower (level 1), the following properties are expressed, for the systemic cycle component and for the pulmonary cycle component. For reasons of presentation the remaining properties are presented only in semiformal form

CP1a(u,v,u',v')   Systemic Cycle Successfulness
At any point in time t,
If at t the systemic cycle component receives a fluid volume V with ingredients I (including oxygen)
and carbonacid is present within the organs
then there exist time points $t' \leq t''$ with $t + u \leq t' \leq t + v$ and $t + u' \leq t'' \leq t + v'$ such that at t' oxygen is delivered to the organs
and carbonacid is taken from the organs
and at t'' the systemic cycle component generates a fluid volume V with ingredients I - oxygen + carbonacid

Formalisation in TTL:

$\forall t\ \forall V\ \forall I$ state(γ, t) |= systemic_cycle_component_receives_fluid(V,I)
& state(γ, t) |= carbonacid_present $\Rightarrow$
$\exists t',t''\ I'$ [$t+u \leq t' \leq t+v$ & $t+u' \leq t'' \leq t+v'$
& state(γ, t') |= oxygen_delivered
& state(γ, t') |= carbonacid_taken
& state(γ, t'') |= systemic_cycle_component_generates_fluid(V,I')
& I' = I - oxygen + carbonacid]

CP1b(u,v,u',v')   Pulmonary Cycle Successfulness
At any point in time t,
if at t the pulmonary cycle component receives a fluid volume V with ingredients I (including carbonacid)
and oxygen is present in the air within the lungs
then there exist time points $t' \leq t''$ with $t + u \leq t' \leq t + v$ and $t + u' \leq t'' \leq t + v'$ such that at t' carbonacid is delivered to the air within the lungs
and oxygen is taken from the air within the lungs
and at t'' the pulmonary cycle component generates a fluid volume V with ingredients I - carbonacid + oxygen

Here V is an amount of fluid and I is a specification of ingredients, as before, and I - A + B specifies the ingredients of I except A and augmented by B.

Component Properties at Aggregation Level 2

At the lowest aggregation level (level 2), properties of the primitive components are expressed. Notice that, considered within the systemic cycle, the heart receives input from outside this cycle, and generates output within this cycle. A similar comment can be made for the heart within the pulmonary cycle.

Component Properties within the Systemic Cycle

CP2a(e, f) Heart Effectiveness in Systemic Cycle
At any point in time t0
if at some $t \leq t0$ the heart (within the systemic cycle) receives from outside the systemic cycle a fluid volume V with ingredients I
and at t0 a heart stimulus occurs
then there exists a time point t1 with $t0 + e \leq t1 \leq t0 + f$ such that at t1 the heart (in the systemic cycle) generates within the systemic cycle a fluid volume V with ingredients I

CP2b(e, f) Aorta Channels Effectiveness
At any point in time t
if the aorta channels receive a fluid volume V with ingredients I
then there exists a time point t' with $t + e \leq t' \leq t + f$ such that at t' they generate a fluid volume V with ingredients I

CP2c(e, f) Organ Capillaries Effectiveness
At any point in time t
if the organ capillaries receive a fluid volume V with ingredients I (including oxygen)
and carbonacid is present within the organs
then there exists a time point t' with $t + e \leq t' \leq t + f$ such that at t' oxygen is delivered to the organs
and carbonacid is taken from the organs
and the organ capillaries generate a fluid volume V with ingredients I - oxygen + carbonacid

CP2d(e, f) Vena Cava Effectiveness
At any point in time t
if the vena cava receive a fluid volume V with ingredients I
then there exists a time point t' with $t + e \leq t' \leq t + f$ such that at t' they generate a fluid volume V with ingredients I

Component Properties within the Pulmonary Cycle

CP2e(e, f) Heart Effectiveness in Pulmonary Cycle
At any point in time t0
if at some $t \leq t0$ the heart (in the pulmonary cycle) receives from outside the pulmonary cycle a fluid volume V with ingredients I
and at t0 a heart stimulus occurs
then there exists a time point t1 with $t0 + e \leq t1 \leq t0 + f$ such that at t1 the heart (in the pulmonary cycle) generates within the pulmonary cycle a fluid volume V with ingredients I

CP2f(e, f) Pulmonary Artery Channels Effectiveness
At any point in time t
if the pulmonary channels receive a fluid volume V with ingredients I
then there exists a time point t' with $t + e \leq t' \leq t + f$ such that at t' they generate a fluid volume V with ingredients I

CP2g(e, f) Lung Capillaries Effectiveness
At any point in time t
if the lung capillaries receive a fluid volume V with ingredients I (including carbonacid)
and oxygen is present in the air within the lungs
then there exists a time point t' with $t + e \leq t' \leq t + f$ such that at t' carbonacid is delivered to the air within the lungs
and oxygen is taken from the air within the lungs
and the lung capillaries generate a fluid volume V with ingredients I - carbonacid + oxygen

CP2h(e, f) Pulmonary Veins Effectiveness
At any point in time t
if the pulmonary veins receive a fluid volume V with ingredients I
then there exists a time point t' with $t + e \leq t' \leq t + f$ such that at t' they generate a fluid volume V with ingredients I

Interaction Properties

Interaction or transfer properties express that the different components are connected in an appropriate manner to enable proper interaction. Such connections are from one component's output to another component's input, or (in the special case of TP1a and TP1b) from one component's input to another component's input. In a general form, delays can be taken into account for the transfers. Note that in this case the output of one component often is the input of the connected component. Therefore, the input state property is then taken identical to the previous output state property and thus, the delays for transfers are assumed to be 0. As a result all g's and h's in the specifications given below can be taken 0.

Interaction Properties at Aggregation Level 1

At level 1, the following transfer properties address the interaction between the pulmonary cycle component and the systemic cycle component.

TP1a(g, h) Systemic Cycle connects to Pulmonary Cycle
At any point in time t
if at some $t \leq t0$ the heart within the systemic cycle receives from within the systemic cycle a fluid volume V with ingredients I
then there exists a time point t' with $t + g \leq t' \leq t + h$ such that at t' the heart within the pulmonary cycle component receives from outside the pulmonary cycle a fluid volume V with ingredients I

TP1b(g, h) Pulmonary Cycle connects to Systemic Cycle
At any point in time t
if at some $t \leq t0$ the heart within the pulmonary cycle receives within the pulmonary cycle a fluid volume V with ingredients I
then there exists a time point t' with $t + g \leq t' \leq t + h$ such that at t' the heart within the systemic cycle component receives from outside the systemic cycle a fluid volume V with ingredients I

Interaction Properties at Aggregation Level 2

At level 2, the following transfer properties were identified. These properties correspond to the arrows in Figure 2.

Interaction Properties within the Systemic Cycle

TP2a(g, h) Heart connects to Aorta Channels
At any point in time t
if the heart generates a fluid volume V with ingredients I
then there exists a time point t' with $t + g \leq t' \leq t + h$ such that at t' the aorta channels receive a fluid volume V with ingredients I

Property TP2a would not be fulfilled, for example, if the heart opening were not connected to the aorta channels, so that the generated fluid volume would stream away without reaching the aorta channels.

TP2b(g, h) Aorta Channels connect to Organ Capillaries
At any point in time t
if the aorta channels generate a fluid volume V with ingredients I
then there exists a time point t' with $t + g \leq t' \leq t + h$ such that at t' the organ capillaries receive a fluid volume V with ingredients I

TP2c(g, h) Organ Capillaries connect to Vena Cava
At any point in time t
if the organ capillaries generate a fluid volume V with ingredients I
then there exists a time point t' with $t + g \leq t' \leq t + h$ such that at t' the inferior and superior vena cava receive a fluid volume V with ingredients I

TP2d(g, h) Vena Cava connect to Heart
At any point in time t
if the inferior and superior vena cava generate a fluid volume V with ingredients I
then there exists a time point t' with $t + g \leq t' \leq t + h$ such that at t' the heart receives a fluid volume V with ingredients I

Interaction Properties within the Pulmonary Cycle

TP2e(g, h) Heart connects to Pulmonary Artery Channels
if the heart generates a fluid volume V with ingredients I
then there exists a time point t' with $t + g \leq t' \leq t + h$ such that at t' the artery channels receive a fluid volume V with ingredients I

TP2f(g, h) Pulmonary Artery Channels connect to Lung Capillaries
At any point in time t
if the artery channels generate a fluid volume V with ingredients I
then there exists a time point t' with $t + g \leq t' \leq t + h$ such that at t' the lung capillaries receive a fluid volume V with ingredients I

TP2g(g, h) Lung Capillaries connect to Pulmonary Veins
At any point in time t
if the lung capillaries generate a fluid volume V with ingredients I
then there exists a time point t' with $t + g \leq t' \leq t + h$ such that at t' the pulmonary veins receive a fluid volume V with ingredients I

TP2h(g, h) Pulmonary Veins connect to Heart
At any point in time t
if the pulmonary veins generate a fluid volume V with ingredients I
then there exists a time point t' with $t + g \leq t' \leq t + h$ such that at t' the heart receives a fluid volume V with ingredients I

Interlevel Relations for the Case Study

The idea of specifying dynamic properties at different aggregation levels is that the dynamics of the whole componential system can be (logically) related to the dynamics of lower levels. At the highest level, the following interlevel relation (between level 0 and level 1) holds:

EA1 & EA2 & EA3 & EA4 & CP1a & CP1b & TP1a & TP1b $\Rightarrow$ CP0

Thus, global property CP0 is implied by the lower level properties. Or, in other words, in all situations in which properties EA1 through TP1b hold, property CP0 also holds. In a similar manner, the following interlevel relations can be established between properties at level 1 and 2:

CP2a & CP2b & CP2c & CP2d & TP2a & TP2b & TP2c & TP2d $\Rightarrow$ CP1a
CP2e & CP2f & CP2g & CP2h & TP2e & TP2f & TP2g & TP2h $\Rightarrow$ CP1b

An overview of all interlevel relations that are related to global property CP0 is depicted graphically in Figure 3 (comparable to Figure 1). These interlevel relations have been automatically checked using the model checker SMV (http://www.cs.cmu.edu/~modelcheck/smv.html; see also McMillan, 1993). This analysis also proved that none of the antecedents can be left out; in particular, if the heart's effectiveness fails, then CP0 cannot be concluded.

Componential Explanation for the Case Study

In the previous subsections a componential analysis A for the circulatory system S has been formalised by compositional verification methods from Computer Science. But in how far does this indeed address componential explanation according to Cummins (1975, 1983) and Davies (2001)? As an example, consider the aorta channels as item I. The function F for this item is given by the property Aorta Channels Successfulness, CP2b(e, f): if it receives a blood stream at one point, it will generate a comparable blood stream at another point. The system's capacity C is Oxygen Delivery Successful-

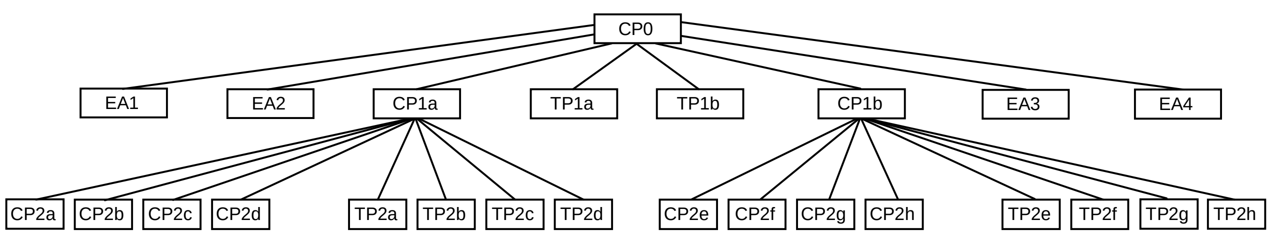

Figure 3: Interlevel Relations for Global Property CP0.

ness CP0(d). Then the function Successfulness of the Aorta Channels within S is described by the following instantiated pattern according to Davies:

The item Aorta Channels within S has systemic capacity function Aorta Channels Successfulness if and only if:

(i*) The Aorta Channels satisfy Aorta Channels Successfulness

(ii*) The analysis appropriately and adequately accounts for S's capacity Oxygen Delivery Successfulness in terms of the organised structural or interactive capacities of components at some lower level of organisation

(iii*) The Aorta Channels are among the lower-level components cited in the analysis that structurally or interactively contribute to the exercise of Oxygen Delivery Successfulness

(iv*) The analysis accounts for S's capacity Oxygen Delivery Successfulness, in part, by appealing to the capacity of the Aorta Channels to satisfy Aorta Channels Successfulness

(v*) The analysis specifies the physical mechanisms in S that instantiate the systemic capacities itemised

Indeed, (i*) to (iv*) are satisfied by the analysis above. However, to satisfy (v*), some specification of the physical mechanisms of the Aorta Channels has to be added, for example by referring to, e.g., Noordergraaf (1978).

Discussion

The article contributes to componential explanation in the area of Philosophy and Cognitive Science by introducing a formal framework of compositional verification as developed within Computer Science. In particular, one of the formal approaches to compositional verification has been applied to a case study to provide a formal analysis, which can serve as the basis for a componential explanation that corresponds to the work of Davies (2001) and Cummins (1975, 1983). In addition, the article contributes to the area of Computer Science and Artificial Intelligence by making clear the conditions on componential explanation to bear on computer software, and provides an additional foundation for the ideas of Clark (1997), Dennett (1978), Newell and Simon (1972). The case study also shows the level of detail necessary to complete a formal analysis of only one aspect of the circulatory system that itself contributes to the capacity of an organism to live. The rigorousness of a formal approach to componential explanation therefore also begs for the development and use of dedicated software support. In the mean time, the formalization opens the doors to the use of existing tools that support verification in Computer Science, such as the model checker SMV.

The case study of the circulatory system has shown to be an appropriate example for the application of compositional verification. It may be expected that the approach is also applicable to other compositional systems (in particular in cognitive domains). For example, many authors (e.g., Fodor, 1983) claim that the human mind can also be structured by components. In future work, it will be explored to what extent the presented approach can be used to explain functions of the mind.

Furthermore, componential explanation can also contribute to the analysis of organisation models. Central issues in organisation modelling are (Lomi and Larsen, 2001):

- how to identify properties of the whole, given properties of parts
- how to identify properties of parts, given desired or required properties of the whole

These issues are similar to the challenges discussed in this paper. The circulatory system has been modelled from an organisation modelling perspective in Bosse et al. (2004).

References

Bosse, T., Jonker, C.M., Meij, L. van der, Sharpanskykh, A., & Treur, J. (2006). A Temporal Trace Language for the Formal Analysis of Dynamic Properties. Technical Report, Vrije Universiteit Amsterdam. http://www.few.vu.nl/~treur/TTL.pdf

Bosse, T., Jonker, C.M. & Treur, J. (2004). Organisation Modelling for the Dynamics of Complex Biological Processes. In: Lindemann, G., Moldt, D., and Paolucci, M. (eds.), Proc. of the International Workshop on Regulated Agent-Based Social Systems: Theories and Applications, RASTA'02. Lecture Notes in Artificial Intelligence, vol. 2934. Springer Verlag, pp. 92-112.

Clark, A. (1997). Being There: Putting Brain, Body and World Together Again. MIT Press, Cambridge, Mass.

Cummins, R. (1975). Functional Analysis. The Journal of Philosophy, vol. 72, pp. 741-760

Cummins, R. (1983). The Nature of Psychological Explanation, MIT Press, Cambridge, Mass.

Davies, P.S. (2001). Norms of Nature: Naturalism and the Nature of Functions. MIT Press, Cambridge, Mass.

Dennett, D. (1978). Brainstorms. MIT Press, Cambridge, Mass.

Fodor, J.A. (1983). *The Modularity of Mind,* Bradford Books, MIT Press: Cambridge, Massachusetts.

Hempel, C.G. (1959). The Logic of Functional Analysis. In: Gross, L. (ed.), Symposium on Sociological Theory; New York: Harper and Row, p. 271-287.

Jonker, C.M. & Treur, J. (2002). Compositional Verification of Multi-Agent Systems: a Formal Analysis of Pro-activeness and Reactiveness. In: (Roever et al., 1998), pp. 350-380. Extended version in: International Journal of Cooperative Information Systems, vol. 11, 2002, pp. 51-92.

Leemans, N.E.M., Treur, J. & Willems, M. (2002). A Semantical Perspective on Verification of Knowledge. Data and Knowledge Engineering, vol. 40, pp. 33-70.

Lomi, A., and Larsen, E.R. (2001). Dynamics of Organizations: Computational Modeling and Organization Theories, AAAI Press, Menlo Park.

McMillan, K.L. (1993). Symbolic Model Checking: An Approach to the State Explosion Problem. PhD thesis, School of Computer Science, Carnegie Mellon University, Pittsburgh, 1992. Published by Kluwer Academic Publishers, 1993.

Nagel, E. (1961). The Structure of Science. London: Routledge & Kegan Paul.

Newell, A. & Simon, H. (1972). Human Problem Solving. Prentice Hall.

Noordergraaf, A. (1978). Circulatory System Dynamics. Academic Press, New York.

Rideout, V.C. (1991). Mathematical and Computer Modelling of Physiological Systems. Prentice Hall, Englewood Cliffs.

Roever, W.-P. de, Langmaack, H. & Pnueli, A. (eds.) (1998). Proceedings of the International Workshop on Compositionality, COMPOS'97. Lecture Notes in Computer Science, vol. 1536, Springer Verlag.

Roever, W.-P. de, Boer, F. de, Hanneman, U., Hooman, J., Lakhnech, Y., Poel, M. & Zwiers, J. (2001). Concurrency verification: introduction to compositional and noncompositional methods. Cambridge University Press.