A CBR APPROACH FOR SUPPORTING ECODESIGN WITH SYSML

Bougain, Sébastien Joël; Gerhard, Detlef
TU Wien, Austria

Abstract
Taking environmental and sustainability issues adequately into account within the design process is mandatory but leads to additional complexity since even more engineering domains have to be involved in the process and since lifecycle information, like use of resources, has to be integrated and managed. Model-Based Systems-Engineering (MBSE) is an approach for managing the product’s complexity by using interconnected models. Yet, means for integrating information on environmental impacts are missing. Moreover, the knowledge integrated in the model, which reflects decision making processes, is not reused properly as a design experience for future products. This paper presents previous research about integrating environmental impacts and ecodesign in SysML to introduce a new Case-Based Reasoning (CBR) approach for adequately reusing previous MBSE experiences and results, such as environmental impacts. To do so, two databases are developed: one for the SysML models and one for storing own environmental impacts. Moreover, a specific retrieval process for selecting the relevant experiences is proposed. The paper ends with a discussion and a brief outlook about the research.

Keywords: Ecodesign, Systems Engineering (SE), Product Lifecycle Management (PLM), Case-Based Reasoning, SysML

Contact:
Sébastien Joël Bougain
TU Wien
Institute for Engineering Design and Logistics Engineering
Austria
sebastien.bougain@tuwien.ac.at

Please cite this paper as:
1 INTRODUCTION

In the recent years, products have gotten increasingly complex, integrating more and more functions. Sophisticated products require experts from various engineering domains like mechanic, electronic and informatics among others. Model-Based Systems Engineering (MBSE) is an approach to manage the complexity of the development of such products based on models instead of documents. According to the International Council on Systems Engineering (INCOSE), “Systems Engineering is an interdisciplinary approach and means to enable the realization of successful systems. It focuses on defining customer needs and required functionality early in the development cycle, documenting requirements, then proceeding with design synthesis and system validation while considering the complete problem” (INCOSE, 2016). The Systems Modelling Language SysML (OMG, 2016) is its main modelling language. One driver for complexity is the adequate consideration of environmental threats respectively the integration of sustainability aspects and ecodesign methods in the development process of modern products. The intricacy of relations between actions of mankind and consequences for the nature and hence mankind, is not yet fully discovered. Yet some main parameters remain. Green House Gases Potential (GHGP) is for instance identified as the core cause of climate change (GEO, 2005). Moreover, the Cumulative Energy Demand (CED) is a relevant value for considering the benefits or harmful effects of a change of design concerning the energy consumption saved in the use phase. That is why they are being integrated in SysML but still not completely. For instance, environmental impacts are integrated in SysML models in order to help design engineers to foresee consequences of engineering decisions (Eigner, 2014).

But what happens after that? What if an engineer leaves a design team? How can engineering knowledge be preserved for future projects? New designers likely make the same mistakes as someone before without having read all the reports made in the design team before. As George Santayana (1905) wrote: "Those who cannot remember the past are condemned to repeat it". That is why this paper presents a Case-based Reasoning (CBR) approach to help design engineers utilizing information that has previously been generated with a focus on corresponding environmental impacts that were linked with a design solution. The aim of such a tool is to better inform design engineers about certain design decisions. Knowledge inside a design team would be better transmitted over the years. Moreover, this approach allows a personalization of an eco-database being closer to the actual impacts made by the products instead of using general ones.

First, related work about CBR approaches and Lifecycle Assessment (LCA) is presented. In the third chapter the integration of environmental impact in SysML from a previous research is presented. In the fourth chapter of this paper the CBR approach for SysML model is introduced in details followed by a discussion and a conclusion about further work and outlook of the research in the last chapter.

2 RELATED WORK ABOUT CBR DATABASES

2.1 Overview of CBR

CBR is a method to follow the quote from the Spanish philosopher mentioned above by supporting design in reminding the designer of the previously designed products that can help solving his new problem instead of repeating mistakes from the past. The basic process of CBR consists of the following steps: New Problem, Recalling previous cases, Adapt Cases, Obtain a new solution, Feed the database for later recall (see Figure 1). The first and most important aspect of a CBR approach is the Case Memory or the database. This includes a representation of a set of previously designed products or systems and provide the basis for using the CBR. The 3 following points have to be addressed in order to define a Case Memory:

• Cases’ content
• Representation of the contents of each case
• The organization of the set of cases in Case Memory.
In working with CBR and SysML, the case's content and its representation and organization are similar because everything has already been well organized and structured. This would not be the case with basic information like emails, CAD parts, etc. That is why the CBR approach is well suited for an application with SysML.

Recalling a case from the database is a pattern matching problem that depends on the formulation of the new issue. This is decomposed as: index, retrieve and select. Indexing is used for comparing the new problem with the previously designed systems. The new problem is indexed to be understood by the database by using the same kind of information. The retrieval process searches the database for one or more previous systems that have similarities with the problem. This could be detailed with different threshold determining if a case is good enough or not. Then the best option is selected and shown to the design engineer.

Sometimes when the best case is selected, it is enough information for the design engineer. For example, a case retrieving a value concerning whether or not a function has to be implemented does not need further work. The simple answer yes or no is enough. But sometimes the information needed is a qualitative value as for environmental impacts. In this case, the adaptation of the selected case is necessary to fit the new problem. For instance, the design of a 10 cm plastic bin has specific environmental impacts that are different to those of a 15 times bigger bin. A basic similarity can be executed for retrieving the needed information.

### 2.2 Forecasting the design of eco-products by integrating TRIZ evolution patterns with CBR and simple LCA

Yang and Chen used TRIZ evolution patterns with CBR and simple LCA for developing eco-products. Here the CBR is used for accelerating the implementation of a specific solution previously chosen with TRIZ. 75 products from various fields, without further information, were collected to construct the database. For indexing the 75 products, 3 types of index were used, first the TRIZ evolution pattern, towards what aims the evolution, then information regarding the product's function and then the limitation of the product's function. The similarity between cases is calculated in Equation (1):

\[
\text{Totalsimilarity} = \sum_{i=1}^{n} w_i \cdot \text{sim}_i(D_i; C_i)
\]

(1)

\(n\) refers to the total number of characteristics in an index system, and \(w_i\) refers to the weighting value of character \(i\). \(\text{sim}_i(D_i; C_i)\) represents the individual similarity between \(D_i\) and \(C_i\) (Yang and Chen, 2012). The weights in the formula are between 0 and 1.

After the calculation, a simple LCA (Graedel et al, 1995) is conducted on the new solution for evaluating and comparing it with the previous design that was already assessed with the same method. No adaptation of simple LCA method from previously designed products is made to assess new ones.

### 2.3 Research on Estimation of Energy Consumption in Machining Process Based on CBR

Gong and Ma developed a CBR method for estimating the energy consumption in machining process (Gong and Ma, 2011). The indexing of products from the database is based first on products attributes
like name, serial number, type, technical parameters, etc. Different formulas analyse the similarity between the different attributes of two products. The texts are compared by determining how many different letters they have. The numeric attributes are compared separately and averaged. The overall similarity between products is calculated via a similar formula as for Yang and Chen. Thanks to this, a set of similar products could be retrieved. In this case, an adaptation was necessary to fit the energy consumption for the new product. Gong and Ma chose a linear equation for adapting the energy consumption of the previous products for the new one.

2.4 A Framework for stepwise Lifecycle Assessment during Product Design with Case-Based Reasoning

This approach from Jeong et al (2010) uses a Function Behaviour Structure (FBS) and Environmental Effect. In the early phase a qualitative evaluation can be used but later in the detailed phase environmental effect can be evaluated more precisely. The FBSe process is used for indexing and organizing the memory case as follows:

- Function: The purpose of the design (product) (e.g. the purpose of a cellular phone is to make or receive wireless calls; the purpose of a beam is to carry a load over a distance; the purpose of a cup is to hold liquid)
- Behaviour: Required functional module/parts and manufacturing process (e.g. cellular phone shows graphical information by LCD panel (functional module); cup holds liquid by providing an enclosed volume (functional part); enclosed volume of cup is made by steel forging (manufacturing process))
- Structure: Description of the physical characteristics of the object (e.g. geometry size, material, color)
- Environmental Effect: Description of the environmental effect of a product (e.g. global warming, ozone layer depletion, acidification) (Jeong et al, 2010)

The Case Memory is organised with a LCA value for each level. The LCA of the whole product is saved at the top level. The case retrieval is also made as previous researches mentioned above with the weighting formula according to the importance of each matching attribute. The weighting scale is between 1, less important attribute, and 5, very important attribute. After having selected the best match in the memory case it is adapted depending on the attributes. If the new product is bigger than the one selected, an adaptation is necessary. Jeong et al (2010) stated that LCA results usually depend on the product material, the mass, the manufacturing process, and the scenario of delivery, use, and disposal stage. And they suggested that from those parameters the material and mass have the greatest effect. By taking the mass as the most important, after the case has been selected, it is adapted according to a proportional equation for the new problem.

Yet, concerning mechatronic products, the use phase is usually the most intensive phase over the product's lifecycle. Moreover, the behaviour used in this research doesn't contain the full possibility of SysML. Indeed, it only concerns functional modules and manufacturing processes instead of taking the interaction between the product and the user into account. How is the user using the cell phone? Is he calling 30% of the time and surfing on the internet for the rest of the time? The behaviour changes the environmental impact of the product during its use phase which is commonly the most intensive phase of mechatronic products.

3 INTEGRATION OF ENVIRONMENTAL IMPACTS WITH SYSML

The following method for mechatronic products extends from the requirements definition to the end of product development and follows the V-model, explained in the VDI guideline 2206 (VDI, 2004). The presentation of the method for integrating Cumulative Energy Demand (CED) and Green House Gas Potential (GHGP) in SysML models is divided in three sections. In the first section the requirement phase including the use of the Ecodesign Pilot (EP) software is explained. Then, different modelling approaches of CED and GHGP for the different lifecycle phases, i.e. Extraction, Manufacturing, Use and End-of-Life (ISO, 2016), are detailed and finally the iteration process of the method is defined. To illustrate the method, engineering design of MendelMax 3D Printer (RepRap, 2016) is used as an example. The following SysML models are made with MagicDraw (Nomagic, 2016) and the database illustration with Microsoft Access (Microsoft, 2016).
3.1 Requirement phase

The requirement phase is a very important phase for engineering design. Little differences and deviations in transforming user or customer requirements to functional features usually leads to major changes and costs in the later phases, like manufacturing. According to the EN NF X-50-100 (AFNOR, 2011) one requirement diagram is needed for one lifecycle phase of the product. Each phase is then detailed within this diagram, defining what the product is doing and with what it is in contact. In general, the use phase of mechatronic products like the 3D Printer is the most intensive phase concerning CED and GHGP because it is the phase where the product is consuming energy or fuel. That is why the use phase is considered the most intensive phase by default in the SysML diagram. To prepare the designer with the future challenges of ecodesign for the product, the Ecodesign Pilot (EP) by Wimmer et al (2002) is integrated in the SysML requirements models. EP contains more than a hundred strategies for developing eco-intelligent products. Each strategy belongs to one or more of the 5 lifecycle phases' goals that optimize the product, i.e. Raw material intensive phase, Manufacturing intensive phase, Transportation intensive phase, Use intensive phase and Disposal intensive phase. Each strategy is detailed with measures that are described with the following attributes:

- Designation: title of the measure
- Priority (P)= R*F (from 0 to 40)
- Relevance (R): how relevant is the measure for the product (0,5 or 10)
- Fulfilment (F): how fulfilled it is yet (0,1,2,3 or 4)
- Idea for realization:
- Costs:
- Feasibility:
- Action:

These attributes and strategies are all integrated by default in the SysML model. Moreover, a value "importance" for each diagram phase is added. This value defines the most intensive phase in the SysML model. Importance 2 defines the most intensive phase, importance 1 defines the second most intensive phase which would optimize a bit less the product, and finally importance 0 corresponds to a phase with little to no improvement for the environmental aspects of the product. An example of a strategy for a use intensive product with 2 measures is illustrated in Figure 2. With such strategies, available soon in the design, the best strategies for optimizing the product can be adopted. As described in the section 3.3, in this chapter, the importance of the lifecycle phases is dynamically changed along with the design of the product.

---

Figure 2. one of the strategies for a use intensive product with its two measures.
3.2 Design

As part of systems engineering, the different lifecycle phases of the product are detailed with the two environmental impacts, CED and GHGP, as follows.

3.2.1 Modelling of the Extraction and End-of-Life phases

For modelling the environmental impacts of the Extraction and End-of-Life phases, the GHGP and CED of the material as well as its mass are needed. To be able to calculate it, as soon as, the design engineer determines the material used in the product in the model, the program is looking for the values of GHGP and CED in general eco-database, like Granta Design (Granta Design, 2016) and when the design engineers defines the mass of the material, the value is calculated and implemented directly in the model to be shown. Moreover, a specific property is added for each part, the maintenance factor. If a part is going to be replaced once or several times over the product's lifecycle, the CED and GHGP of the part have to be multiplied by the number of times it is replaced. For assemblies like the 3D Printer's casing, containing structural beams and holding parts, no value is added for the extraction phase but concerning the End-of-Life phase, a CED and GHGP could be added if a disassembly process is required for recycling or disposal. When the design is detailed in a CAD Software, the material and mass of the parts and the product available in the CAD Software or the PDM System are directly linked with the SysML model and updated for a more detailed analysis of the environmental impact of the product.

3.2.2 Modelling of the Manufacturing phase

For modelling the Manufacturing phase, GHGP and CED of the manufacturing process as well as the process parameters are needed. CED and GHGP values can be obtained from eco-databases. Other information has to be determined by design engineers at the beginning of design and implemented in the model. Yet, when the design is ready for manufacturing and available in an Enterprise Resource Planning System (ERPS), the data are directly linked to the SysML model and updated with the metadata (time of update, etc…). GHGP and CED values still come from a general eco-database but other values can be obtained from the manufacturing process. As some ERPS calculate the energy consumption of the machines, this data can directly be used instead of the value from the general eco-databases to calculate the environmental impacts.

3.2.3 Modelling of the Use phase

The modelling of environmental impacts of the Use phase is made via behaviour models of SysML. To do so, each Use Case is defined with how much it occurs over the product's lifecycle. For the 3D Printer, the assumption that 5% of the time the calibration of the nozzle occurs has been made. In the SysML model the energy consumption of the different parts needed for calculating the CED of this use case is available in the parametric diagram. With this data, the energy consumption of the product in the Use Case "calibrating the nozzle” can be calculated and implemented for each part. Modelling of the Extraction, Manufacturing, Use and End-of-Life phases is illustrated in Figure 3 with the positioning motor of the MendelMax 3D Printer.

![Figure 3. Modelling of Extraction, Manufacturing, Use and End-of-Life phases in SysML](image)

3.3 Iteration

When a product is designed, the values of CED and GHGP for each lifecycle phase are stored and compared with one another. This redefines which lifecycle phases is the most intensive and which is the second most intensive etc. The importance values of the EP strategies can be adapted accordingly in the
model to guide the design engineer further with his work of optimizing the product for the environment. The goal of the next chapter is to develop the CBR approach for optimizing the product development.

4 CONTINUOUS OPTIMIZATION THROUGH A CASE-BASED REASONING APPROACH

Knowing how much energy consumption is caused by a particular function or part is a valuable input for the design engineer. Knowing what a decision in the early phases of design could cause in the later phases is also a valuable input. A database and a CBR approach are made for this purpose and to allow the personalization of the environmental impacts database. This chapter is divided in the following sections. First the Case Memory where all the cases are stored is explained. Then the case retrieval when a new problem appears is detailed and finally the adaptation of cases and the reporting are detailed.

4.1 Case Memory

The Case Memory is divided in two separates parts. The first part explained in this section concerns the database used for the SysML models and the second part details the database used for the personalization of the environmental impact.

4.1.1 The SysML models database

Each case is stored in the Case Memory for a later recall. This first database is still under development as the exhaustiveness of SysML is not completely stored in it yet. At this time, some information that SysML can contain are still lost during the save in the database. But the important data necessary for our CBR approach like model elements and attributes, e.g. stereotypes and environmental impacts, are taken into account.

One table represents the products that contain the different models and elements. This table corresponds to the top level or project level and consists of a unique ID, the name of the product, its description and the date when it was developed.

Another table is created for the SysML elements. Each element is defined with an ID, uniquely identifying the element in the database, the product to which it belongs, the type, e.g. Requirement, Structure, Use Case, …, a title, describing the element and a description going more in details. The title uses a specific format for easing the indexing and recognition, like hashtags for twitter. Each title has to be formulated differently according to their type. A Requirement is formulated with first a verb, then an object and finally a level. For instance, with the example of a 3D Printer: "Print (verb) CAD part (Object) in less than 5 hours (level)”. A Use case is defined with a verb and an object. The title of a Structure block is formulated with an object. The title of an activity is made with a verb and an object and the title of a state is made with a text. Those texts are defined as text properties.

The table "relations" contains a unique ID, the source of the relation, the target and the type.

Finally, the table "additional attributes" describing the elements, e.g. stereotypes and environmental impacts among others. Attributes are a unique ID, the ID of the linked element, the type of the attribute (stereotypes, value property) and its value. This is shown in the Figure 4 below.

![Figure 4. Case Memory, from a SysML model to a database.](image-url)
4.1.2 The Environmental impact database

During design, first the design engineer determines the material, mass, manufacturing process and its parameters by writing them directly in the SysML model. When the design is in its advanced phase, those data are replaced with the actual materials, masses, and manufacturing processes. Concerning the values from general eco-databases they can be replaced with actual measurement or communication about the manufacturing process or extraction. This second database holds the values concerning the GHGP and CED allowing the personalization of the database and hence a better optimization of environmental assessment. Two tables for the manufacturing process and the materials are created. Moreover, a table for the GHGP and the CED are created to follow the changes over time of the different materials and processes. To do so each entry in the database for the GHGP or CED is defined with a unique identification number, a date when the change occurred, the object concerned, whether a material or a manufacturing process, the new value and the source where it comes from. To obtain the new value from one source, for instance when creating a block with a material and a weight, a request to the database is necessary and is obtained with the following Equation (2):

\[
\text{NewValue} = \frac{1}{\text{NbValues}} \sum_{i=1}^{n} \text{Value}_i
\]

Where \( \text{Value}_i \) are the different values from the same source for the same object and \( \text{NbValues} \) is the total number of value from a source for an object. An illustration of the database is in Figure 5. For the material and manufacturing process a parameter is defined that describes the parameter needed to obtain the GHGP or CED for a part. For instance, for a material, the mass is likely used, or the volume, linked with the mass with the density. This helps avoiding a 3D Printer with a bigger print room to use a bigger motor, although it doesn't need one.

![Figure 5. The environmental impacts database](image)

4.2 New problem and case retrieval

When the designer is working on a new product he first formulates the requirements. When the requirements are formulated according to a similar pattern as the indexing of the Case Memory, specific solutions can be found. Looking through the database an identical requirement can be found. In that case the process jumps directly to reporting. Whereas if by looking through the database no identical block is to be found, then a similarity approach for calculating the Block similarity between the new block \( B \) and one block in the database \( B_i \) is executed with Equation (3):

\[
\text{BlocSimilarity}_{B/B_i} = \frac{1}{\text{NbTextproperty}_{B_i}} \sum_{j=1}^{n} f(\text{Textproperty}_{j From B}, \text{Textproperty}_{j From B_i})
\]
Where \( N_{textprop} \) is the number of text property in the block and \( f(\text{Textproperty}_f, \text{Textproperty}_j) \), designs the finding of the longest common sequences, also known as LCS problem, divided by the maximal text length, as follows in Equation (4):

\[
f(a, b) = \frac{N_{LCS(a,b)}}{\max(|a|,|b|)}
\]  

(4)

Where: \( N_{LCS(a,b)} \) is the number of character in the longest common sequence and \( \max(|a|,|b|) \) is the number of character of the longest text between a and b. With this formula, the requirement "Use ABS as a printing material" and "Use PLA as printing material" are 75% identical due to the use of a different material and the absence of a "a" before the word "printing". The more the writing of requirement is formatted the better the recognition will be. Moreover, the use of acronym can increase the percentage of the recognition although the two words have nothing in common. Here ABS and PLA have 33% similarity but by using the full names Acrylnitril-Butadien-Styrol and Polylactide, they only have 24% similarity.

The parameters are for the different object defining the block's title. For instance, the 3 parameters of a requirement's title are the verb, the object and the level. The total similarity between two blocks is defined with the sum of the product between the parameter similarity and the weighting factor. The one that has the highest similarity is chosen

### 4.3 Adaptation

In the case of requirement, no adaptation is required. But for the structure parts it is important to adapt the environmental impacts values of previous cases to new ones. As Gong and Ma, a linear regression for adapting the CED and GHGP for linear parameter like mass and time, has been chosen as shown in Equation (5):

\[
\text{EnvironmentalImpact}_{\text{Object}} = \frac{\text{Parameter}_{\text{Object}}}{\text{Parameter}_{\text{Database}}} \times \text{EnvironmentalImpact}_{\text{Database}}
\]

(5)

Where Environmental Impact Object is the Environmental impact of an object, Parameter object is the physical parameter that defines the Environmental impact for this object (for instance the mass for a material in the extraction phase or the power in the use phase).

### 4.4 New solution / Reporting

After the best matching case has been selected and adapted, the blocks coming from that case and derived from the one being created are shown. For instance, the requirement "Use ABS as a printing material" (Verb + object + level) is being written in a new model. Two years before a product has been designed with the same requirement. The program is then showing the derived blocks that were linked with the requirement. The farther a block is from the new block, the lighter it is represented. The will to "use ABS as a printing material" caused the specification "Heating until 250°C" which was satisfied with the chosen extruder with specific environmental impacts, see Figure 6.

---

**Figure 6.** Derived specification, part and note with the CBR approach for SysML
By confirming it or not the design engineer can accept the proposition made in the SysML model. If a block proposed has been accepted, then the others blocks derived from the one accepted are shown in an opaque fashion and proposed as well until the design engineers refuse it or there is no block to show anymore. By also showing the notes the design engineer could be aware of the problem met that time for this requirements or functional solution and adapt his solution accordingly. As shown in the example in Figure 6, a note about the real-life expectancy of the extruder was added to warn that a second extruder was actually needed during the old product's life, hence doubling some environmental impacts because of the maintenance factor introduced.

5 DISCUSSION AND CONCLUSION

A CBR approach is very practical for use with SysML. SysML is indeed a formatted language that can be used as an indexing system for the Case Memory. The proposed method can support design engineers in ecodesign or general product development tasks by showing the different decisions that have been made before for similar cases as well as recommendations, mistakes warnings, and caused environmental impacts. Although the formatted requirement assists the recognition of similar requirements, it will possibly lead to a loss of information that is vital for such an early phase. A good balance between formatted requirement and recognition of similarity or a change of weighting factor before the function f, which is comparing parameters, has to be found and analysed deeper. The proposed CBR approach is an efficient way to support the engineer in developing his product by helping him with the previous choices that were made for similar requirements, functional unit, behaviour models or parts. Moreover, the environmental impact database allows the personalization of an eco-database getting closer and closer to a precise LCA and hence a better ecodesign of a product for a specific company. This approach as well as the integration of environmental impacts in SysML still have to be integrated as a support program. A first idea is to develop a plugin for MagicDraw with the help of the calculation plugin ParaMagic (Inter CAx, 2016).

REFERENCES

AFNOR, (2011), NF X50-100:2011-11-01: Value management - Functional analysis, basic characteristics - Functional analysis: Need (or external) functional analysis and technology/product (or internal) functional analysis - Requirements for deliverables and implementation approach.

Eigner et al., (2014), System Lifecycle Management: Initial Approach for a sustainable Product Development Process Based on Methods of Model Based Systems Engineering. IFIP.


Santayana G., (1905), The Life of Reason, chapter XII.


