Prediction of CAD model defeaturing impact on heat transfer FEA results using machine learning techniques

Florence DANGLADE 1, Philippe VERON 1, Jean-Philippe PERNOT 1, Lionel FINE 2

(1) : LSIS - UMR CNRS 7296
Arts et Métiers ParisTech
Aix-en-Provence, France
{Florence.Danglade, Philippe.Veron, Jean-Philippe.Pernot}@ensam.eu

(2) : EADS France IW
Suresnes, France
Lionel.Fine@eads.net

Abstract: Essential when adapting CAD model for finite element analysis, the defeaturing ensures the feasibility of the simulation and reduces the computation time. Processes for CAD model preparation and defeaturing tools exist but are not always clearly formalized. In this paper, we propose an approach that uses machine learning techniques to design an indicator that predicts the defeaturing impact on the quality of analysis results for heat transfer simulation. The expertise knowledge is embedded in examples of defeaturing process and analysis, which will be used to find an algorithm able to predict a performance indicator. This indicator provides help in decision making to identify features candidates to defeaturing.

Key words: Defeaturing, CAD model, Finite Element Analysis, Machine Learning, a priori estimation.

1.- Introduction

In the field of transfer from Computer Aided Design (CAD) to Finite Element Analysis (FEA), CAD model adaptation ensures the quality and the reliability of analysis results. Among all techniques [TB1] (deleting parts, defeaturing, simplification, merging, idealization) defeaturing is essential for CAD model adaptation. Defeaturing consists in removing irrelevant features (protrusion, pockets, holes, fillets/rounds, chamfers).

The choice of candidates for defeaturing and tools used depend on the target of the simulation (structural, dynamic/fluid/heat transfer analysis, assembly/disassembly procedure evaluations) as well as on the type of method adopted for solving it (Finite Differences, Finite Element Analysis and so on) and the type of the CAD model (B-REP, STEP, Mesh, ...). Today, the choice of candidates for defeaturing is often empirical and leads, by precaution, to have model for analysis more precise than necessary. This significantly increases the cost of meshing and solving. It is important to know the thresholds that drive engineers during their decision to propose rules.

Machine learning [M1] tools, like neural networks, support vector machine or decision tree, are able to imitate and accurately predict behaviour from carefully selected examples. Moreover, these techniques can take into account equations or known relations. Therefore, they can capitalize the knowledge from a set of process examples for CAD model adaptation and predict impact of defeaturing process for a new case.

We propose in this paper to validate one of the process steps of the preparation of CAD models for heat transfer analysis. It will be shown that the machine learning techniques can be used to estimate the defeaturing impact on analysis results by predicting a performance indicator of defeaturing. Section 2 presents a state of art related to impact of defeaturing and the use of machine learning tools on CAD. Then, in section 3 an algorithm for estimating the impact of defeaturing on FEA results is proposed. Some results are discussed in section 4.

2.- Related works

For convective heat transfer analysis, the numerical analysis is applied to a mesh of a fluid volume. Without simplification, a huge number of elements to mesh local detail is necessary. The volume meshed can be difficult to obtain (often the meshing is impossible), and the computing time is very high. Thus, defeaturing significantly reduces processing time.

The experts select the candidates for defeaturing based on the solving feasibility and on the result accuracy target of analysis. For that, experts take into account boundary conditions (adiabatic, constant temperature, heat flux surfaces). Then, they estimate if the impact on the results is small and the computation time faster.

In the field of finite element static analysis Tang [TG1] studied the defeaturing impact on analysis by using the change of a model’s strain energy. In recent years, a posteriori evaluation of defeaturing impact has aroused great interest [FC1]. However, only few attentions have been paid so far to the need for an a priori evaluation.

In the field of heat transfer analysis, Gopalakrishnan [GS1] propose a theory for estimating analysis errors in case of heat
transfer with a high accuracy of the error estimated. This method can be used for local applications or simple cases (with a reduced number of parts and features). In the case of a complex product, a very large number of details are deleted; the characteristics of the CAD model are strongly impacted, so it is difficult in these conditions to implement such a theoretical method.

Most of approaches about decision making for defeaturing propose [DP1] a feature by feature analysis. It is also difficult to take a global decision on the overall product when we have a large number of features. Therefore, we propose to predict impact of defeaturing on heat transfer analysis from global examples of defeaturing processes.

Machine Learning tools are widely used in design to address optimization problems [SL1], decision-making problems [L1], shapes recognition [JK1], item recognition or extraction for reuse, recognition from point cloud scans [GM1], feature recognition [SA1] and transfer CAD/CAM [DM2]. In this present paper, we propose a use of machine learning techniques for the prediction of the performance indicator of process defeaturing. A process defeaturing is defined by a list of features to delete, techniques used for features removal (feature removal for native CAD model, deleting and reconstruction of faces, deleting and reconstruction of meshes ...) and the sequencing of operations. The main objective of this paper is to evaluate defeaturing process in order to identify an optimal list of features to remove and to estimate the impact of defeaturing on Finite Element Analysis. Defeaturing impact evaluation is performed by estimating the quality of the analysis results and the costs of defeaturing, meshing and analysis operations.

3.- Algorithms for prediction of FEA result quality

Figure 1 represents the general algorithm of the proposed approach. The first step “pre-processing” consists in building a database of defeaturing process examples. For this, we extracted all data from initial CAD model and from prepared CAD model and simulation. Data are information like the format (e.g. CATIA native, STEP, IGES, tessellated model), the material, the dimensional quantities (e.g. size, surface area, volume, number of meshes elements, number of faces) and relations between features and boundary condition. Data about simulation are information on boundary conditions and analysis results. CAD preparation process should be described by a list and sequencing of simplification operations. Proposed by several experts, the initial CAD model data must be as exhaustive as possible (the database contains more than 250 attributes). Useful prepared CAD model data are selected in next step “learning”.

In the second step, machine learning techniques are used for carrying out learning models for the prediction of process performance indicators. For each learning model, determinant attributes are selected from the extracted examples. The final step consisted in selecting candidates for defeaturing by estimating the quality of analysis results and costs (duration of defeaturing, meshing and analysis). It was performed in 4 sub-steps, namely features classification, prediction of quality of analysis results, prediction of cost and decision making.

Details of various models are described in the following sections.

![Algorithm for evaluation of defeaturing process](image_url)

**Figure 1: algorithm for evaluation of defeaturing process**

3.1- Pre-processing

The database used for learning must be as exhaustive as possible. It contains a set of collected information:

- a global description of the initial CAD model (type of CAD model, boundary conditions, geometry, materials);
- a description of each feature (type of feature, relationship between features and with boundary conditions, geometry);
- a description of the prepared CAD model (type of CAD model, gains on volume/surface/mesh elements, geometry, mesh characteristics, list of deleted features);
- information about simulation (simulation goal, accuracy of results, targets of analysis) and industrial constraints.
- information about the process of preparation (list of deleted features, tools used for defaturing/meshing, cost of meshing, cost of defaturing).

This database must include a significant number of already known examples.
Then, for a new case, only initial CAD data (geometrical characteristics, material, type of data...), objectives (physical quantity to calculate, position of boundary conditions, values on boundary condition) and constraints are indicated.

3.2- Learning

We use Machine Learning tools for prediction of analysis results quality and for prediction of defaturing / meshing / analysis costs. In this paper, we describe the prediction of the analysis result quality (ARQ).
In this section, an algorithm to predict a performance indicator, illustrated in figure 2, is proposed.

![Figure 2: ARQ prediction](image)

The quality of analysis results is evaluated by predicting an indicator ARQ (analysis result quality). For learning, ARQ (1) is calculated from the value of analysis result for the initial cad data ($I_{AR}$) and for the defeatured CAD data ($D_{AR}$). For a new case, ARQ indicator will be predicted from determinant attributes (described in section 3.2.1) by machine learning techniques.

\[
ARQ = 1 - 6.67 \frac{\sum (D_{AR} - I_{AR})}{I_{RA}}
\]  

(1)

Where $R$ is the number of results observed. $ARQ=0$ if analysis is impossible, $ARQ=0.2$ if the error is greater than 15%, $ARQ=1$ if the error is negligible.

A significant number of examples are extracted, whose results of heat transfer analysis from initial CAD model and from defeatured CAD model are both known.

3.2.1- useful attributes selection

The major stage in using Machine Learning tools like Neural Network is to define input data. These data are represented by a vector of attributes which is defined by specific attributes for each learning goal. Attributes are selected from the database by means of classification methods of input variables according to the impact of each attribute on the output.

Regarding the prediction of the quality indicator, 12 relevant attributes has been identified. Selected attributes are described below:

Table 1 gives some examples of CAD model configurations (CAD models with different level of simplification) with their associated attribute values for each configuration. Attributes x1 to x7 describe set of features deleted

<table>
<thead>
<tr>
<th>x1</th>
<th>x2</th>
<th>x3</th>
<th>x4</th>
<th>x5</th>
<th>x6</th>
<th>x7</th>
<th>Configuration</th>
</tr>
</thead>
<tbody>
<tr>
<td>k1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>Initial CAD model before defaturing</td>
</tr>
<tr>
<td>k2</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>all holes are deleted, other features are retained</td>
</tr>
<tr>
<td>k3</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>Small holes are deleted, other features are retained</td>
</tr>
<tr>
<td>k4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>Small rounds are deleted, other features are retained</td>
</tr>
<tr>
<td>k5</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>Small holes, rounds and chamfer are deleted, other features are retained</td>
</tr>
<tr>
<td>k6</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>All features far from boundary conditions are deleted</td>
</tr>
</tbody>
</table>

Table 1. Examples of CAD model configurations and associated attribute values
f is a factor equal to 0 if the feature is retained or $k=1$ if the feature is deleted. Attributes $x_8$ and $x_9$ give indications on the weight of features deleted relative to initial CAD model.

- **attribute $x_8$** is the volume ratio (2) between the set of feature deleted and the product, where $VolF_i$ is the volume of each feature $F_i$ (for $i=1$ to $i_{max}=$number of features of the product), $I_{VolP}$ is the volume of the initial product:

$$VolRatio = \frac{\sum_{i=1}^{i_{max}} (VolF_i \times f)}{I_{VolP}} \tag{2}$$

- **attribute $x_9$** is a ratio (3) between the number of mesh element in the set of feature deleted and in the product ($I_{MeshP}$), where $MeshFi$ is the number of mesh element for each feature $F_i$:

$$MeshRatio = \frac{\sum_{i=1}^{i_{max}} (MeshFi \times f)}{I_{MeshP}} \tag{3}$$

- **attribute $x_{10}$** is the minimal Euclidean distance (4) between the feature deleted and the nearest boundary conditions:

$$D_{F_{BC}} = MIN(D_{F_{i_{BC}}} \times f) \tag{4}$$

Attributes $x_{11}$ and $x_{12}$ give gains obtained by defeating.

- **attribute $x_{11}$** is the gain (5) of volume between initial ($I_{VolP}$) and defeated product ($D_{VolP}$):

$$VolGain = \frac{D_{VolP} - I_{VolP}}{I_{VolP}} \tag{5}$$

- **attribute $x_{12}$** is the gain (6) of number of mesh element between initial ($I_{MeshP}$) and defeated product ($D_{MeshP}$):

$$MeshGain = \frac{D_{MeshP} - I_{MeshP}}{I_{MeshP}} \tag{6}$$

### 3.2.2- learning model selection

In this section, we identify the best Machine Learning tools for predicting the Analysis Result Quality (ARQ) from the 12 selected attributes described in section 3.2.1. Machine Learning tools are selected [DM1] for each learning model on the overall examples by cross-validation methods. The evaluation criteria of learning model are the percentage of correctly classified instances and the average quadratic error $AQE$ (7):

$$AQE = \frac{1}{N} \sum_{k=1}^{K} (y(x^k) - p(x^k))^2$$

Where $x^k$ is the vector of attributes for the example $k$, $y(x^k)$ is the known output value, $p(x^k, w)$ is the output value predicted by learning model tool, $w$ is the vector of weight between nodes.

In the context of our database and for our study (preparation with CATIA V5, heat transfer analysis with ANSYS [41], and learning with Weka [W1]), the learning model whose quadratic error is the smallest for a high percentage of correctly classified instances is a neural network (Multi Perceptron) as shown table 2.

Figure 3 illustrates the selected neural network with 3 layers using back-propagation method:

- input layer contains 12 nodes matched to the 12 selected attributes;
- output layer is the discrete value of the ARQ indicator ($0 \leq ARQ \leq 1$);
- hidden layer: contains $m$ nodes whose weight and number depend on learning examples.

Weights $w$ are found by means of Tan Sigmoid methods.

![Figure 3: Neural Network](image)

<table>
<thead>
<tr>
<th>Learning model</th>
<th>AQE</th>
<th>% correctly classified instances</th>
</tr>
</thead>
<tbody>
<tr>
<td>Naive Bayes</td>
<td>0.06</td>
<td>75</td>
</tr>
<tr>
<td>Multilayer Perceptron</td>
<td>0.03</td>
<td>78</td>
</tr>
<tr>
<td>RBF Network</td>
<td>0.05</td>
<td>63</td>
</tr>
<tr>
<td>Super Vector Machine</td>
<td>0.05</td>
<td>78</td>
</tr>
<tr>
<td>Decision tree</td>
<td>0.05</td>
<td>69</td>
</tr>
</tbody>
</table>

Table 2: Evaluation of learning models

Table 3 gives some examples of Analysis Result Quality predicted with the selected learning model.

### 3.3- Use on a new unknown case

Each feature $F_i$ belonging to the global product according to:

- its type (5 classes = protrusion, pocket, hole, fillet/round and chamfer);
- its size (2 classes = small and large);
- its Euclidian distance between boundary conditions (2 classes = near from boundary conditions and far from boundary conditions).
A new unknown case is a CAD model configuration whose one or several type of features is deleted. For a new case, several configurations should be studied. A first approach consists in evaluating a defeating process proposed by an expert. Indication thus obtained does not identify an optimal process. Another approach consists in studying different typical configurations with some features deleted. Table 1 shows examples of CAD models configurations. A great number of random configurations should be used. The 12 determinant attributes are extracted from the data of new case. For a new case, we know after priori only initial CAD model characteristics and simulation goal. Some selected attributed described in section 3.2 are unknown (volume and number of mesh elements). These values can be estimated from known rules or by using machine learning techniques (this step is not described in this paper).

The next stage consists in predict the indicator of quality ARQ described in section 3.2.2. Then, costs of defeating, meshing and analysis are estimated by predicting the duration of these operations. The crossed analysis of the defeating impact on simulation for all configurations makes it possible to take a decision for each set of feature. These set of features should be classified in 3 groups "the set of feature must be retained", "the set of feature should be deleted" and "the set of feature should be deleted on approval". If the decision is conditional, an individual analyze on each feature is necessary. From first hypotheses and criteria proposed by experts, which will be completed as and when the study, we propose to retain a feature if:

- the feature is parent of another feature to retain;
- the size of the feature is larger than a limit size;
- the distance between the feature and the boundary condition is smaller than a limit distance.

From results on all cases, a best global configuration should be compiled.

### Table 3: Examples of prediction

<table>
<thead>
<tr>
<th></th>
<th>Volume gain</th>
<th>Mesh element gain</th>
<th>Quality of analysis results (predicted)</th>
<th>Quality of analysis results (calculated)</th>
</tr>
</thead>
<tbody>
<tr>
<td>k1</td>
<td>0.00</td>
<td>0.0</td>
<td>1.0</td>
<td>1.00</td>
</tr>
<tr>
<td>k2</td>
<td>2.22</td>
<td>-18.8</td>
<td>0.8</td>
<td>0.86</td>
</tr>
<tr>
<td>k3</td>
<td>0.03</td>
<td>-6.2</td>
<td>1.0</td>
<td>0.97</td>
</tr>
<tr>
<td>k4</td>
<td>-0.04</td>
<td>-15.7</td>
<td>1.0</td>
<td>0.86</td>
</tr>
<tr>
<td>k5</td>
<td>0.45</td>
<td>-25.3</td>
<td>0.4</td>
<td>0.32</td>
</tr>
<tr>
<td>k6</td>
<td>0.46</td>
<td>-7.1</td>
<td>0.8</td>
<td>0.73</td>
</tr>
<tr>
<td>k7</td>
<td>-0.54</td>
<td>-21.0</td>
<td>1.0</td>
<td>0.86</td>
</tr>
<tr>
<td>k8</td>
<td>-0.52</td>
<td>17.2</td>
<td>1.0</td>
<td>0.95</td>
</tr>
</tbody>
</table>

A crossed analysis of the defeating impact on simulation for all configurations makes it possible to take a decision for each set of feature. These set of features should be classified in 3 groups "the set of feature must be retained", "the set of feature should be deleted" and "the set of feature should be deleted on approval". If the decision is conditional, an individual analyze on each feature is necessary. From first hypotheses and criteria proposed by experts, which will be completed as and when the study, we propose to retain a feature if:

- the feature is parent of another feature to retain;
- the size of the feature is larger than a limit size;
- the distance between the feature and the boundary condition is smaller than a limit distance.

From results on all cases, a best global configuration should be compiled.

#### 4.- Results

This paper focuses on convective heat transfer analysis with ANSYS CFD. Examples for learning were single parts and products whose defeating process was proposed by experts. Defeating examples were distributed in 2 sets: a training set (66%) and a test set (33%) statically equivalent.

Meshing was carried out by ICEM [11]. Meshes characteristics were the same for all configurations on a new case (triangular volume mesh, medium size, without adaptation).

Table 3 shows volume gain, mesh element gain, defeating cost and quality of analysis results for the new case illustrated in figure 3. Impact of defeating was done for 20 configurations. Examples of configurations k1 to k8 are described in table 1. Configuration k7 is the configuration proposed by an expert.

Machine learning technique was selected and performed by the Weka platform [W1]. For prediction of quality analysis results (ARQ), the quadratic error for Multi Layer Perceptron technique (3 nodes layers) is about 14% with a cross validation. This score should be increased by adding more examples. Table 3 gives values of ARQ predicted with neural network and calculated. These values show that the choice of the learning model is suitable for decision on the quality of the analysis. Comparative studies between configuration k7 proposed by expert and k8 predicted by learning model, shows that this last configuration is more accurate (features on inlet and outlet are retained for k8).

The set of features to delete after decision based on results from different configurations are rounds, chamfers far to boundary conditions and small holes far to boundary conditions. Configuration k8 in table 3 gives predicted values for the global configuration resulting to this decision.

In all cases tested, it turned out that the defeating made possible meshing and analysis for most complex cases and allowed to accelerate processing time for simple cases (from 5 to 35%).

The global configuration is shown in figure 3. 3a) represents the initial CAD model. In 3b) red features are features to retain and green features are features to deleted after global decision.

3c) represents defeatured CAD model. As indicated, the analysis duration is reduced for a little average error on temperature value. The mesh quality is significantly increased.
5.- Conclusion

The results of this study have shown that machine learning techniques can be used to learn how to defeature CAD model for heat transfer analysis. The proposed approach to predict the simulation quality can be easily applied to the prediction of cost and to the global decision making. Machine learning techniques can be a good mean to capitalize the knowledge embedded in empirical processes.

The proposed approach need to have information on post defeature (volume and number of mesh elements of CAD model). A first solution consists in defeaturing the new case. The next step is to predict the quality indicator of the analysis without pre defeaturing. This will require to predict unknown attributes or not to use them (which require to have a greater number of attributes and thus a very large number of examples for learning). Future work should take into account other CAD model preparation steps (deleting parts, defeaturing, simplification, merging, idealization). The global configuration proposed at the end of our workflow is not the optimal configuration. Further studies should therefore implement an optimization loop so that using the developed indicator, the best defeaturing configuration can be suggested to the designers. At the end, the proposed approach and tools should reduce significantly the number and duration of design iteration.

REFERENCES


[I1] http://www.ansys.com/Products/Other+Products/ANSYS+ICEM+CFD


