*Meeting of the Working Group on 'Modeling of Machining Operations, STC Cutting, CIRP, CIRP General Assembly, Part II, Tianjin, China, 28-30, August, 1997*

# Prediction Augmentation Through Reinforcing Interactions Amongst Modeling, Sensing and Learning

**Patri K. Venuvinod**
**City University of Hong Kong**

## 1. Introduction

STC Cutting of CIRP had set up a Working Group on 'Modeling of Machining Operations' in August 1995. The scope of the Group was restricted to operations with defined cutting edges [1]. In January 1996 (at it's Paris meeting), the Group reached the conclusion that, despite extensive developments that have taken place over the last 50 years, there is a serious dearth of models for machining operations which are accurate and general enough to be used in industry. In particular several members, including the present author, suggested that the Group may wish to decide "whether we need to give a nudge or push in a specific direction to suit our engineering purposes as we perceive them in the short as well as the long term" [2]. There was substantial support to the suggestion at the meeting. However, so far, the nature of the specific 'nudge(s)' needed have remained undefined.

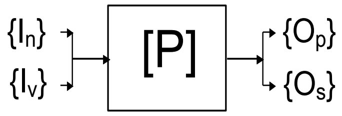
This paper describes and argues for some specific 'nudges' which the author believes would be particularly useful at this point in the history of machining research. Several of the ideas contained in this paper are not new. Bits and pieces of the ideas can be found in recent machining literature. However, these bits are yet to be integrated into a unified and collectively agreed strategy.

This paper does not intend to be a definitive paper. It merely aims to stimulate discussion in a specific direction. Hence, an intuitive approach will be used in developing the ideas.

The following sections follow up on the propositions included in [2] and are mainly inspired by the author's observations during the development of a literature database for the Working Group [3].

## 2. Process Inputs and Outputs

A process [P] responds to a given set of inputs by producing a set of outputs.




$\{I_n\}$ = Nominal Input
$\{I_v\}$ = Unanticipated Input Variations
P = The Process
$\{O_p\}$ = Perfromance Measure (Output)
$\{O_s\}$ = Output capabele of being sensed

Fig. 1

Process inputs can be classified into two types: The nominal inputs, $\{I_n\}$, and the unexpected input variations, $\{I_v\}$.

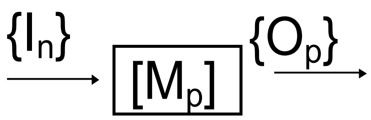
The nominal input array, $\{I_n\}$, consists of nominal values of the inputs deliberately set by the process designer or operator. The typical contents of $\{I_n\}$ are nominal tool and work material properties; tool geometry variables: e.g. rake angle, clearance angle, and cutting edge obliquity; and cutting conditions: e.g. cutting speed, feed, depth of cut, and dry or wet.

In practice, many unexpected variations, disturbances or perturbances occur in the inputs conditions. For instance, the process planner might be expecting a tempered structure in the steel being machined. However, it is quite possible that a particular batch of workpieces might unexpectedly have a martensitic structure. Such variations can not often be anticipated and lead to unexpected variations in the outputs. $\{I_v\}$ refers to this array of unexpected input variations.

Amongst the many outputs from a process, we are particularly interested in two subsets. The first is a set of performance measures (e.g. cutting forces, cutting temperatures, tool life) which we need for the purpose of process design, tool design, process control, etc. A partial list of performance measure outputs, $\{O_p\}$, is given in [2]. On other occasions, we need to measure (or sense) another subset, called the sensing subset $\{O_s\}$, for monitoring and control purposes. Acoustic emission, force components, power, acceleration, vibrations, noise, tool-work thermocouple temperature, and flank wear are amongst the outputs that have been sensed for monitoring purposes. In principle, there could be elements common to $\{O_p\}$ and $\{O_s\}$. However, in practice, it has not been possible to sense many of the elements of $\{O_p\}$.

## 3. The Current Status of Modeling of Machining Operations

A process [P] produces outputs $\{O_p\}$ as well as $\{O_s\}$ as a result of a series of phenomena natural to the process under the given conditions. A modeller speculates on these phenomena in terms of a network of cause-effect relationships and attempts to capture the more important relationships in a model $[M_p]$ which aims at quantitatively predicting $\{O_p\}$ for a given input array $\{I_n\}$:




$M_p$ = The process model aimed at predicting $\{O_p\}$ from given nominal inputs $\{I_n\}$

Fig. 2

The more detailed is the capture of the relationships, finer is the 'degree of brush' of the model. The greater is the range of cutting situations in which the model is effective, greater is the 'generality' of the model. The greater is the correlation of the predicted outputs with actual output magnitudes, greater is the 'accuracy' of the model.

The major aim of the Working Group is to promote a quicker development of models which can quantitatively predict the array of performance measures $\{O_p\}$.

The author was recently surprised read the following sentences in a news brief included in the 'Manufacturing Engineering' journal published by SME: "A metal cutting FE modeling software called Mach2D to be released by Third Wave to be realized in 1997. Capable of modeling forces, temperatures, material removal rate, chip growth, chip breaking, chatter, and vibration." [4].

There are substantial similarities between the capabilities the Working Group hopes to impart to machining models and the capabilities of Mach2D as described above.

But why was the author surprised to read the news brief? The answers lies in the following quotes from [3]:

- "[T]he progress of machining science has not been dramatic despite its history of over 50 years and the efforts put in by a large number of scientists. Notwithstanding the availability of powerful computers which can be used to analyze more complex operations, significant interest still persists in the relatively simpler case of single edge orthogonal cutting. It might be that many theoretical concepts can be tested more easily for orthogonal cutting. However, as long as the modellers do not progress well beyond single edge orthogonal cutting, they would have little impact on industrial practice."
- "[V]ery little work has been done in studying (leave alone modeling) machining with non-plane rake faces although the vast majority of modern tools have complex rake faces. This is another reason why modeling has not moved out of the laboratory (or the computer room in these days) and on to the shop floor."
- "[A]nalytical modeling continues to dominate the modeling scene. However, there is growing interest in Finite Element Modeling in recent years. That is good news. The bad news is that, except in the rare cases when the models are supported by custom-built machining databases, these models can still only predict qualitative trends. Their performance with regard to quantitative prediction continues to be questionable — another reason for the lack of impact on industry."
- "There is indeed a wealth of knowledge concerning machining. It is generally agreed that this knowledge is quite useful for process designers. The models so far created have been quite successful in terms of qualitative predictions. However, it is a different story when it comes to quantitative prediction. No wonder then that there have been very few automated industrial machining systems where a modeling package is a regular and critical component of the system control software."

Mach2D, whatever its virtues, is unlikely to be more than a computer-aided tool for exploring machining process trends. It would be surprising if it were to be able to make quantitative predictions at the accuracy levels acceptable to industry for a wide range of machining operations. And, it is only a 2D model whereas most machining operations (and the phenomena associated with them) are 3D in nature.

As noted in [3], the majority of current machining process models ($[M_p]$} are either analytical or computational (mainly FEM based).

Analytical models have only been partly successful. They have created a deep understanding of machining processes which is often valuable to the process designer. Some models have demonstrated the ability to quantitatively predict some of the performance measures — especially the mean magnitudes of cutting force components, cutting power and temperatures. However, the following problems persist:

- Only very few practical operations have been modeled to the required generality.
- Many performance measures cannot still be quantitatively predicted.
- They are reliable only when type II chips are ensured. If a built-up edge exists, the chips are discontinuous or serrated, and so on, success is questionable. And, such chip forms occur often and cannot be anticipated.
- They invariably require massive machining databases, MDB (consisting of data on the work material flow stress magnitudes, shear angles, chip flow angles, mechanical and thermal properties of the tool material, etc.), to guide the selection of the magnitudes of an array of model coefficients {C}. The creation of machining databases requires extensive off-line experimentation, $X_{off}$, and is therefore expensive.

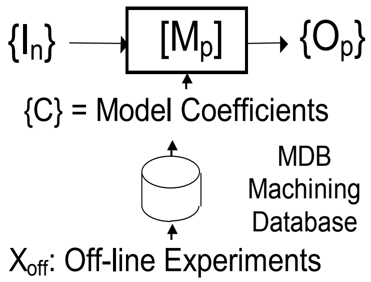



Fig. 3

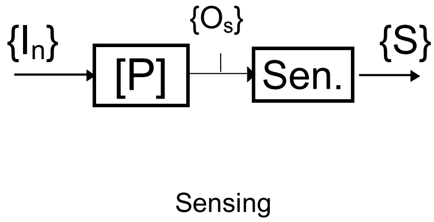



Fig. 4

With regard to the MDB, two approaches are in vogue.

The first attempts to obtain the properties of work materials from fundamental (or, more appropriately, non-machining) tests. The problem is that these tests need to be carried out at conditions simulating the large strains, strain rates, and temperatures prevailing in machining. Only limited success has been achieved in this regard so far. Further, the tests that are available are very expensive to conduct.

The second approach (pursued by Armarego, Venuvinod, etc.) utilizes machining (in particular, single edge orthogonal/oblique cutting) itself as the material test [2]. Methods are then found to extrapolate these data to more complex cutting situations.

The current status with regard to computational modeling is similar to the above except that there are additional uncertainties and problems with regard to the criteria that are to be used in relation to chip-work material separation, transition from transient to steady state cutting, etc.

Finally all modeling approaches have to face the reality that chip formation in machining is, in general, not uniquely defined. Minute perturbations in the input conditions (i.e. the presence of $\{I_v\}$), can lead to substantial changes in the process state. Consequently, model coefficients derived from one set of off-line experiments cannot be transported confidently to other instances of the same process.

It is clear from the above discussion that we need to explore other ways to obtain the magnitudes of the model coefficients ({C}).

**4. Sensing**

Sensing involves the measurement and monitoring of the desired subset $\{O_s\}$ of the process output and processing the signals collected through a signal processing algorithm or system to yield an array of sensed features {S}:

The following was noted in [3]:"[T]here is significant interest in sensing and monitoring. A more detailed study of the [literature] database has shown that much of this interest has appeared since the mid eighties and that the sensing techniques have generally been augmented by Artificial Neural Nets and such AI-based techniques. On occasion, there has been a model-based processing of sensory data." Further, "[A]coustic emission has emerged as the predominant sensing technique in on-line process monitoring."

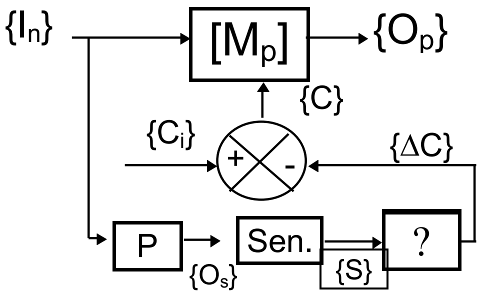
**4. Augmenting the Prediction Ability of Modeling Through Sensing**

We have noted that a major problem with the traditional approach to modeling is the need to calibrate the model coefficients through a machining database MDB. The traditional approach to compiling a machining database has been through extensive off-line experimentation. As we have noted, a reliance on such static databases leads to the inability of the model to perform effectively when there are minor perturbations in input conditions. These perturbations (inherent variability) cannot generally be anticipated and, owing to the inherently non-unique nature of the chip formation process, leads to large prediction errors.

In contrast, modern sensing technologies enable us to perform on-line or real-time measurement of outputs. Sensing and sensor fusion technologies are getting better every day. Is it time now to abandon the concept static databases in favor of dynamic databases, i.e. in favor of data obtained in real-time through sensing? Such real-time data are likely to be superior to static databases since they have been collected while the same perturbations exist in the input conditions.

However, there is one problem to address. Model $[M_p]$ has been created to predict $\{O_p\}$. But the sensors measure $\{O_s\}$ and may or may not be able to measure $\{O_p\}$ (in most cases, they do not). How can we solve this problem?

The answer probably lies in the realization that both $\{O_p\}$ and $\{O_s\}$ are outputs from the very same process [P] as the latter manifests in real time. Hence, $\{O_s\}$ should contain much information concerning the way [P] behaves for the nominal input $\{I_n\}$. All we need to do is find a way of utilizing this insight towards determining the corrections $\{\Delta C\}$ that need to be made to the model coefficients $\{C_i\}$ set initially through static machining databases (static):




Calibrating A Predictive Model by Using Sensed Output
Fig. 5

The author has attempted a variation of the above approach with some success in compensating for workpiece dimensional errors in turning. More about this will be presented in section 10.

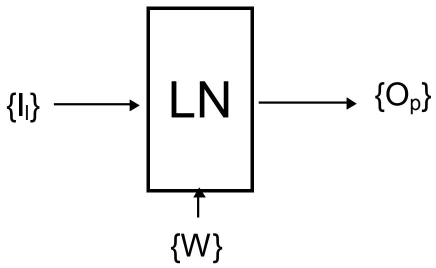
**5. Learning**

Learning is usually based on pattern recognition and clustering techniques such as Artificial Neural Nets (ANN). More recently, there have been applications of Chaos and Fractal theories and Wavelet theories towards learning.

An ANN can be supervised or un-supervised. Amongst the various types of supervised ANN, the Back Propagation Net (BPN) has become particularly popular:

A supervised ANN operates in two modes: Training Mode, and Test or Utilization Mode.

In the training mode a set of data/signals ($\{I_l\}$ are input to the learning network. In addition, the network's connection weights [W] are adjusted to an initial set. Off-line or on-line experiments are conducted with the given process using a training set of $\{I_l\}$ arrays. The sensed outputs $\{O_s\}$ are used as training data to yield a set of corrections $[\Delta W]$ for [W]. The cycles are repeated until the network outputs signals which agree with the experimental outputs within an acceptable error margin. [W ]is then frozen so that the network becomes the trained network. The network can then be used to 'predict' the outputs for new instances of $\{I_l\}$.




$\{I_l\}$: Inputs to the learning system
[W] : Matrix of Connection Weights
LN : Learning Network
Prediction by Using a Learning System
Fig. 6

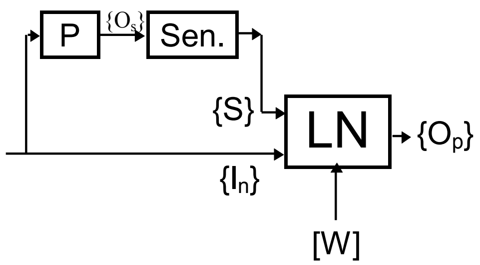
**6. Augmenting Learning Through Reinforcing Interactions with Modeling**

The performance of learning systems has been variable. Intuitively speaking, a learning system should be able to predict effectively if it is able to capture the essence of the process. It follows then that if the process is more complex (i.e. it has more uncertainty, more non-uniqueness, has a larger network of cause-effect-relationships, etc.) then one would require a more complex learning system.

What factors could influence prediction effectiveness of a learning system. To the author's knowledge no rigorous answer is available today to this question. Hence we will resort to the intuitive proposition that the learning effectiveness of a learning systems (such as BPN) can be increased by:

1. adopting a more complex network architecture (this is likely to increase the cost and processing time, hence we will leave it out of the present discussion);
2. increasing the number of training cycles (this also will increase the time required for training and, as encountered often in practice, there might be saturation effects; hence we will leave this alternative alone.); and
3. increasing the size of the input array $\{I_l\}$.

Let us focus on the last approach. How can we increase the size of the input array $\{I_l\}$? We have already used up all the known inputs. Hence we need to look elsewhere. How about $\{O_s\}$? The present ethos of modeling assumes that it can down without a knowledge of $\{O_s\}$ in the prediction of $\{O_p\}$. And, as noted earlier, this has only been partially successful. However, we have already noted that $\{O_s\}$ does contain some insight into [P]. If so, why not include $\{O_s\}$ as one of the inputs to the learning system?:

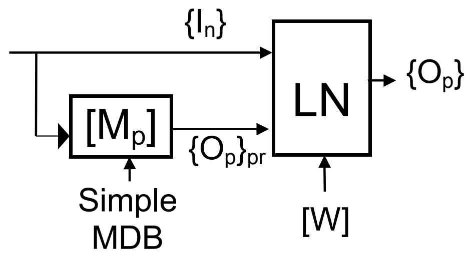



Augmenting a Learning System Through $\{O_s\}$
Fig. 7

The above strategy might look similar to many recent works where a combination of sensing and learning has been used. However, in these works, the output to be sensed was both $\{O_s\}$ and $\{O_p\}$, i.e. one or more of the performance measures were sensed. This has led to a flurry of actions to discover or invent methods for directly sensing each performance measure. The problem with this approach has been an unnecessary increase in the number of sensors and complexity of the system. In contrast, in Fig. 7, there need be no overlap between $\{O_s\}$ and $\{O_p\}$. One may use only those sensors which are convenient and are known to be able to reasonably capture the essence of the process and learn to predict any performance measure.

**7. Augmenting Learning Through Modeling**

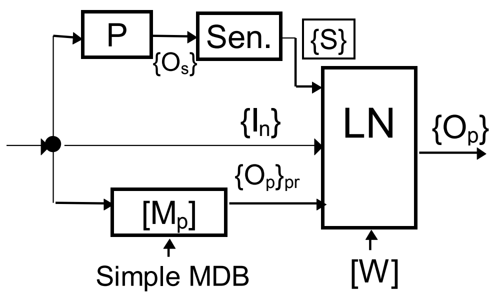
Another approach (which is probably more in line with the agenda of this Working Group) is to increase the size of the input array of the learning system by supplementing with the model predictions:




Augmenting Learning with Modeling
Fig. 8

The predictions, $\{O_p\}_{pr}$, from the model contain much insight into the behavior of the process by virtue of the accumulated knowledge on machining the model incorporates and the empirical information contained in its MDB. However, unlike in conventional modeling, the model $[M_p]$ is not expected to perform accurate quantitative prediction. It is left to do what it does best, i.e. anticipate the qualitative trends of $\{O_p\}$. Augmented with this additional insight, hopefully, the learning system will perform faster and better. The MDB supporting the model, in contrast to the current practice, need not be elaborate and highly accurate. The learning system will eventually learn to correct the errors arising from the MDB.

**8. Throwing Everything Together**




Augmenting Learning Through

Modeling as well as Sensing
Fig. 9

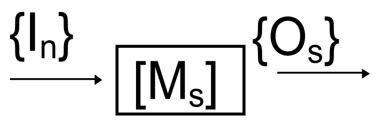
**9. The Problem**

Figures 5, 7, 8, and 9 have outlined 4 approaches towards obtaining synergy amongst modeling, sensing and learning. All the approaches have aimed at predicting $\{O_p\}$. With well researched implementation, they are likely to be effective when (I) the sensing system measures the desired $\{O_p\}$, and/or (ii) the learning system is trained to learn $\{O_p\}$. However, what should one do when, as is often the case, one does not have sensors to measure $\{O_p\}$, or when no reliable off-line data exists to lean $\{O_p\}$?

To illustrate the above problem consider the case of sensing using acoustic emission (AE) signals which has become particularly popular in recent years. Many ANN based learning systems have been developed in association with AE. But, note that AE is not a performance measure. It is merely a conveniently sensed signal. How can the measured AE signals aid the process of predicting $\{O_p\}$? Clearly, much further research is required to answer this question satisfactorily.

Consider now the specific case of predicting cutting forces when AE constitutes the $\{O_s\}$. Dornfeld and others have suggested that the 'True Mean Square (TMS)' value of the AE signal is proportional to the work rates in cutting Likewise, ? have demonstrated from end milling investigations that there is a strong correlation between TMS and the measured cutting force components [6]. Thus it should be possible to gain much insight into the real-time process phenomena that influence cutting force magnitudes through real-time sensing of AE. Further, a learning system based on data obtained through $[M_p]$ may not need to have a very comprehensive MDB when the learning is augmented by AE. This is because the information obtained from the AE signals could be used to (somehow) calibrate the model or compensate for the errors in the model.

The above discussion suggests that it would be useful to direct a part of future analytical or computational modeling efforts towards developing an ability to predict $\{O_s\}$ (instead of merely predicting $\{O_p\}$):




$[M_s]$ = The process model aimed at predicting $\{O_s\}$ — unlike $[M_p]$ which predicts $\{O_p\}$

A Possible Nudge
Fig. 10

**10. A Case Study**

Several projects have recently been initiated at the author's laboratories to implement and test some of the approaches described above. The following provides a description of the progress made on one of the more promising projects.

The aim of the project (see Fig. 11) is to enhance an existing turning center so that it acquires the ability to (i) autonomously perform on-machine inspection of the parts machined, (ii) continuously learn from the dimensional errors it has discovered, (iii) anticipate the dimensional errors expected on the next part from the uncompensated part program, and (iv) compensate the part program appropriately.

The system attempts to compensate for machine tool errors (kinematic, positioning, etc.); thermal deformation errors; and the elastic deflections of the machine tool, workpiece, and cutting tool under the cutting load.

Preliminary laser measurements are used to determine the distribution of the machine tool errors and thermal deformation errors (as functions of cutting time). A BPN may be trained to learn the information thus obtained. Next, simple analytical models are developed for the compliances of the machine tool and the workpiece. Only semi-empirical models (after M. Kronenberg) have so far been adopted while modeling the radial component of the cutting force. The models when put together may contain over ten unknown model coefficients which change from one cutting situation to another.

On-machine inspection is performed by using the ‘Fine Touch’ principle [7]. “Fine Touch’ enables the use of the cutting tool itself as the contact probe by sensing changes in the signal from an electromagnetic coil placed around the tool (or the workpiece). The method is capable of detecting contact within a positional accuracy of 1 $\mu$m.

The major problem that arises is regarding the unknown coefficients of the force and compliance models which need to be calibrated for each new cutting scenario. To solve this problem, initial values for the compliance model coefficients are determined from the preliminary experiments and, for the cutting force model, from a simple static MDB. It is left to the neural network to integrate all the information.

The system has been implemented in mostly in the manual mode, i.e. the functions of many of the modules and the interfacing between the modules has been implemented manually. The results have shown that the dimensional errors, which are of the order of 80 $\mu$m for the uncompensated part programs, can be decreased to around 5 $\mu$m. This is an encouraging result since the random error itself of the machine tool is of the order of 3 $\mu$m.

**Acknowledgments**

The author would like to thank his colleague, Dr. Alexander Djordjevich, for his contributions to some of the ideas contained in this paper. He also thanks his research student, Mr. Liu Zhanqiang, for his implementation of the startegy illustrated in Fig. 11.

**References**

1. van Luttervelt, C.A., 1995, Working Group: Modelling of Machining Operations,” Working paper, STC Cutting, CIRP General Assembly, Enschede.

2. Armarego, E.J.A., Jawahir, I.S., Ostafiev, V.A., and Venuvinod, Patri K., “Modeling of Machining Operations,”, Working Paper, Meeting of the Working Group on ‘Modeling of Machining Operations’, STC Cutting, CIRP, Paris, January 1996.

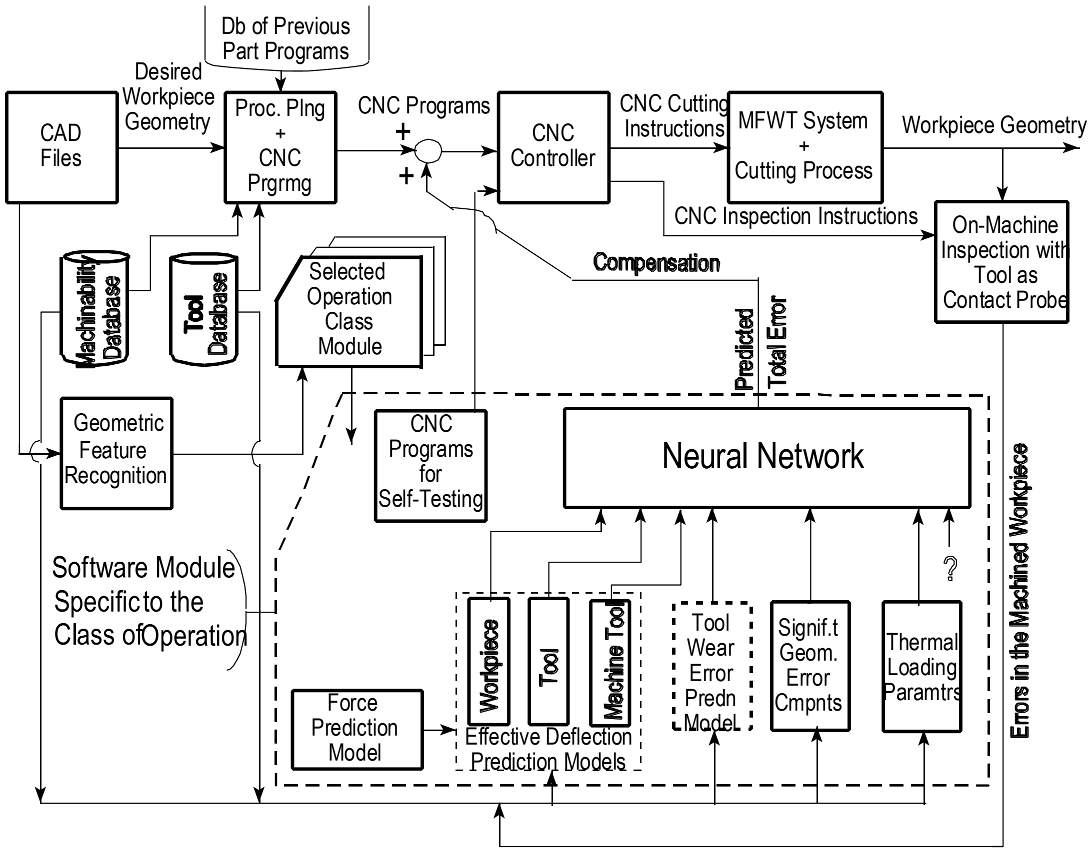
3. Venuvinod, Patri K., “Literature Database for Modellers of Machining Operations: Progress Report,” Working paper, Meeting of the Working Group on ‘Modeling of Machining Operations’, STC Cutting, CIRP General Assembly, Tianjin, August 1997.

4. Anonymous, “Model Your Metal Cutting,” Manufacturing Engineering, SME, May 1997, p. 48.

5. Kannatey-Asibu, E.J., and Dornfeld, D.A., “Quantitative Relationships for Acoustic Emission from Orthogonal Metal Cutting,” J. Eng. For Industry, Trans. ASME, Vol. 103, pp. 330-340, August 1981.

6. Man Liu, and Liang, Steven Y., “Monitoring of Perophera l Milling Using Acoustic Emission,” Trans. NAMRI/SME, pp. 120-127, 1990.

7. Ostafiev, V.A., and Venuvinod, Patri, K., “A New Electromagnetic Contact Sensing Technique for Enhancing Machining Accuracy, IMECE-97, ASME, November 1997, accepted.




A Learning System Augmented Through Modeling and Sensing
for the Compensation of Dimensional Errors Produced by Turning Centers
Fig. 11