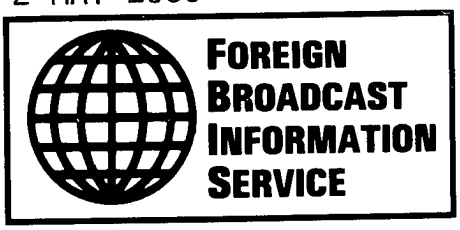
JPRS-JST-89-009

2 MAY 1989



JPRS Report

Science & Technology

Japan

JPRS-JST-89-009

2 MAY 1989

SCIENCE & TECHNOLOGY

JAPAN

CONTENTS

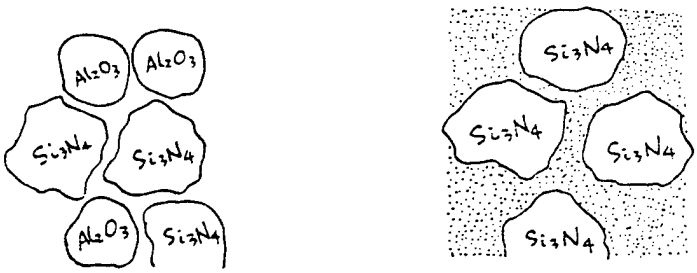
Properties of β-SiAlON Prepared by HIP

43067065a Tokyo 8TH HIGH TEMPERATURE MATERIALS KISO TORONKAI in Japanese 10-11 Nov 88 pp 52-55

[Article by Kazuya Yabuta, Hiroaki Nishio, and Hiroki Okamoto of Nippon Kokan K.K., and Kazumori Kishi and Masaki Umebayashi of Government Industrial Research Institute (GIRI), Kyushu]

[Text] 1. Introduction

Defects in β-SiAlON ($Si_{6-z}Al_zO_zN_{8-z}$, where z = 0 to 4.2) caused by non-uniform blending of raw material powders lower the strength of its sintered ceramic material. To prevent the non-homogeneous blend, Kishi et al. successfully improved the strength by approximately 70 kg/mm^2 for a ceramic material produced from the same powder blend as the previous one but by using aluminum alkoxide in place of alumina powder.[1] (Figure 1)



Solid-solid blend (previous method)

Solid-liquid blend (GIRI, Kyushu, method)

Figure 1. Illustration of Blending Methods for Homogeneous Mixture

In this report, the preparation of β-SiAlON with z = 0.5 from α-Si_3N_4 and aluminum isopropoxide through gas pressure sintering and hot isostatic pressing (HIP) is described, and its mechanical properties and oxidation characteristics are discussed.

2. Experimental Method

In a ball mill, 91 percent by weight of α-Si_3N_4 powder and 9 percent by weight (on Al_2O_3 basis) of hexane solution of aluminum isopropoxide were wet-blended and, after spray drying, preliminarily calcined in air at 900°C for 2 hours to make the raw material powder. After preliminary molding with a 30 MPa one-shaft press and 300 MPa cold isostatic pressing (CIP), followed by gas pressure sintering at 1,800 to 2,000°C for 5 to 10 hours in a N_2 atmosphere of 1 MPa, a β-SiAlON ceramic material was prepared with 1 to 2 hours' HIP under 190 MPa.

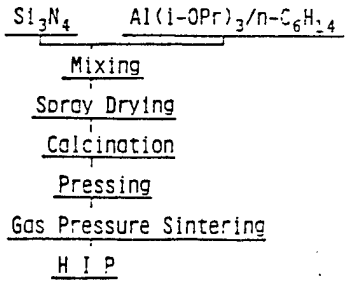


Figure 2. Flow Chart of Sample Preparation Steps

3. Results

(1) Mechanical Properties

According to the JIS specifications, normal temperature three-point and high temperature atmospheric bending tests were conducted. (Figure 3) In the graphs of Figure 3, SiC, Si_3N_4, α-SiAlON, β-SiAlON were all commercially available. All of them, except SiC, are confirmed to contain Y_2O_3. Normally, at 1,350°C these test pieces containing Y_2O_3 undergo transformation thought to be the softening of the interparticulate glassy phase. However, as can be seen in the graph, β-SiAlON prepared for the experiment breaks elastically even at 1,350°C. Its bending strengths were approximately 80 kg/mm^2 and approximately 75 kg/mm^2 at ambient temperature and 1,350°C, respectively, showing almost no deterioration at the higher temperature.

(2) Crystalline Phase

When Si_3N_4-Al_2O_3, the component group of the experiment, is used as the starting material, the single phase of β-SiAlON is not obtained. Always, oxygen-rich multiple phases have to be produced.

When the crystalline phase of a sample is traced with X-ray diffraction during gas pressure sintering, a crystalline phase, called X-phase, is found in a ceramic material sintered around 1,650°C. However, when sintered higher than 1,750°C, only two phases of the β-Si_3N_4-type and Si_2N_2O-type structures can be confirmed. Our understanding is that a low-melting liquid phase is formed during the process of sintering between the Al_2O_3 coating over the particles and the oxide films of SiO_2 and Si_3N_4, and an X-phase will elute when cooled

after sintering at 1,650°C or the Si_2N_2O-type crystalline phase will precipitate from the liquid phase and Si_3N_4 after sintering further at a temperature higher than 1,650°C. These phenomena can be explained without contradicting the well-known four element phase diagram of the Si-Al-O-N system.[2]

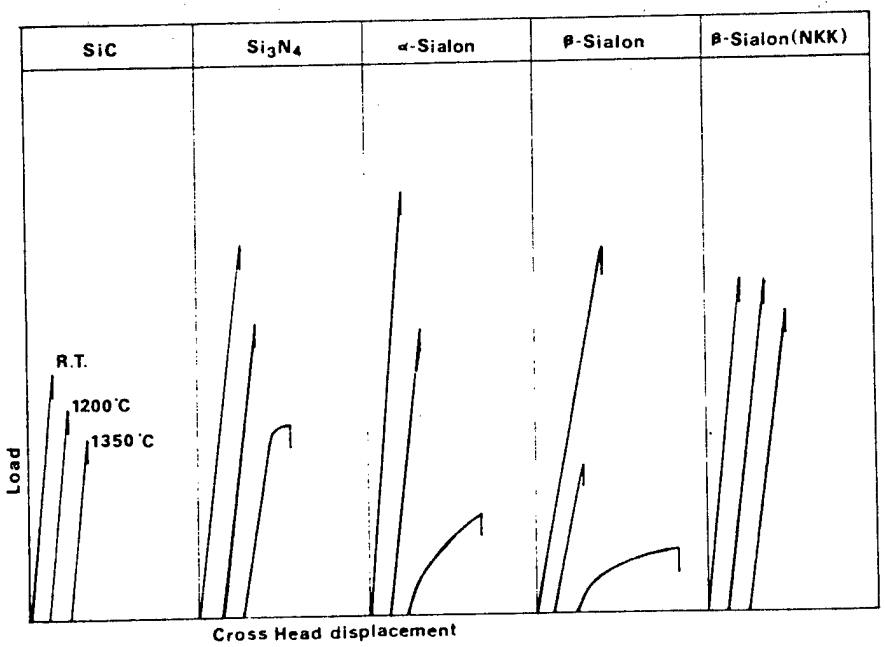


Figure 3. Load-Displacement Curves During Three-Point Bending Strength Measurement for β-SiAlON

(3) Observation With Transmission Electron Microscope (TEM)

A TEM photograph of β-SiAlON, prepared for this experiment, is shown in Figure 4 [not reproduced].

No interparticulate phase greater than 5 nm can be seen in the triple point of the interparticulate zone. The high values of high temperature strength are considered to originate mainly from the extreme micronization of the interparticulate phase. The high strengths also indicate that neither the Si_2N_2O-type crystalline phase (O-SiAlON), contained in the SiAlON ceramic material of this experiment, nor β-SiAlON weakens at 1,350°C.

Ohashi et al.[2] have already obtained a Si_2N_2O-type ceramic material, which maintains its high temperature strength, from Si_3N_4 and SiO_2 with CeO_2 as an aid. Homma et al.[3] reported about a Si_3N_4-type ceramic material with high temperature strength by sintering Si_3N_4 without any additives.

Therefore, it should be naturally understandable that, in the β-SiAlON ceramic material prepared in this experiment, both the major phase of β-SiAlON and the second phase of O-SiAlON show no decline in high temperature strength.

Also, all three particles shown in Figure 4 were confirmed by electron beam diffraction to have the β-Si_3N_4-type structure.

(4) Oxidation Characteristics

The oxidation characteristics of the five samples, for which the mechanical properties were measured, were examined and compared. Each sample was cut to a size of 10 x 20 x 10 (thickness) mm and the pieces were oxidized in an electric furnace at 1,500°C or 1,600°C for 5 hours. Their appearance after oxidation is shown in Figure 5 [not reproduced].

Because of a glassy phase which rises to the surface of each Y_2O_3-containing Si_3N_4, α-SiAlON, and β-SiAlON, the outlines of these samples are not clear, whereas β-SiAlON prepared in this experiment is covered only on the surface by an extremely thin oxidation film, showing an excellent anti-oxidation property.

To compare quantitatively the weight increase due to oxidation, test pieces were subjected to oxidation at 1,200°C for 100 hours and 1,350°C for 100 hours and the weight increases were determined. (Figure 6) As shown in Figure 5, β-SiAlON prepared in this experiment exhibited an anti-oxidation property almost equal to or slightly better than SiC.

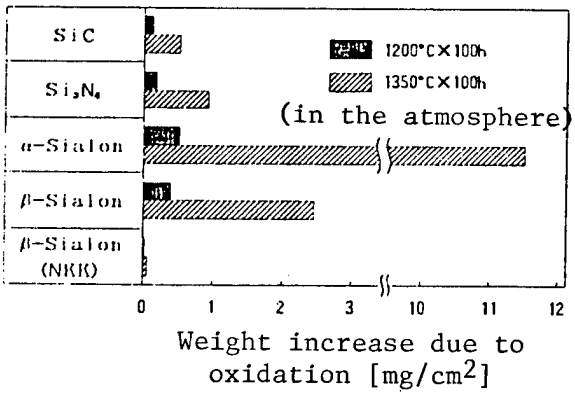


Figure 6. Oxidation Weight Increase for Ceramic Samples

It is not yet clear why the anti-oxidation property has improved, but basically it is conjectured that the oxidation film of comparatively stable and highly viscous SiO_2 is formed over β-SiAlON consisting of Si-Al-O-N and prevents the progress of oxidation. On the other hand, when Y_2O_3 is involved, it is surmised, the interparticulate component and SiO_2 form a Y_2O_3-SiO_2 compound or low-melting point glass, neither of which works effectively as a protective film at a high temperature.

4. Summary

Attaching great importance to homogeneous blending, a β-SiAlON ceramic material was prepared using aluminum alkoxide as a starting raw material. As a result, the ceramic product exhibited almost no decrease in strength even at 1,350°C

and had an excellent oxidation characteristic. These characteristics of the ceramic material appear to indicate that Si_3N_4-base ceramics can be fully used at a temperature exceeding 1,350°C.

References

1. Kazumori Kishi, Masaki Umebayashi, Eiji Tani, and Kazuo Kobayashi, Journal of Ceramic Society of Japan, 92, [4], 231-232 (1984).

2. K. H. Jack, J. Mater. Sci., 11, 1135 (1976).

3. Yuki Ohashi, Shuzo Kanzaki, et al., Collection of Draft Papers, the 1987 Annual Meeting, Ceramic Society of Japan, Paper 3A25.

4. Katsuhiko Homma, Fukusaburo Yamamoto, and Hiroshi Okada, Journal of Ceramic Society of Japan, 95, [2], 223-228 (1987).

Properties of Si_3N_4-SiC Whisker Ceramics

43067065b Tokyo 8TH HIGH TEMPERATURE MATERIALS KISO TORONKAI in Japanese 10-11 Nov 88 pp 95-99

[Article by Yasuhiro Goto, Takeyuki Yonezawa, and Akihiko Tsuge of Toshiba R&D Center]

[Text] 1. Introduction

The effects of whisker compositing on strength and fracture toughness and the relationships with sintering temperature were examined for the Si_3N_4-SiC whisker-base composite ceramics. Furthermore, as a means to make the whisker effects more conspicuous, the control of whisker orientation by extrusion-molding was examined.

2. Experimental Method

2.1 Examination of Sintering Temperature

To investigate the effects of sintering temperature, a sintering aid which enables densification at a relatively low temperature was examined. In this paper, experiments using a yttria-spinel ($MgAl_2O_4$)-base aid are discussed. The aid was added at a rate of 5 percent by weight in each case. The standard content of SiC whiskers was set at 20 percent and, under certain sets of conditions, the amount was changed from 10 to 40 percent to evaluate the effects of that change. Under each set of sintering conditions, the matrix without whiskers was sintered for comparison purposes.

Prescribed amounts of the sintering aid and SiC whiskers were added to silicon nitride, the raw material powder was prepared by ball mill blending and was sintered at a hot press. The sintering was done at 1,500 to 1,750°C, in a nitrogen atmosphere, at a pressure of 400 kg/cm^2, and for 60 minutes.

Resulting ceramic samples were tested for density and Young's modulus (supersonic method). Furthermore, JIS-size test pieces (3 x 4 x 40 mm) were fabricated and evaluated for fracture strength and fracture toughness. Fracture strength was measured by a three-point bending test with a span of 30 mm. Fracture toughness was measured with the SENB method by three-point bending, similar to the fracture strength test, with the addition of a

0.1 mm-wide notch in each test piece. The microstructure of each sample was observed with a scanning electron microscope (SEM) and its crystal component phase was sought by X-ray diffraction.

2.2 Orientation of SiC Whiskers

Molds of a thin sheet form were made by extrusion-molding, and the controllability of the orientation of whiskers was investigated. A binder for molding was added to a mixed powder containing 20 percent of whiskers and mixed in a (Henschel) mixer before forming a sheet 60 mm wide. Sheets were made in five thicknesses of 0.06, 0.1, 0.2, 0.5, and 1.0 mm, to study the relationship to whisker orientation.

Sheet-form molded materials, after drying, were subjected to X-ray diffraction to obtain the whisker orientation from the ratio of diffraction intensities. Also, the sheets were layered and sintered at 1,600 to 1,750°C, and resulting samples were evaluated in the same manner as described in 2.1 above.

3. Results and Discussion

3.1 Examination of Sintering Temperature

Figures 1 and 2 show the relationships of sintering temperature with density and Young's modulus, respectively. The graph of Figure 1 indicates that dense samples were obtained at temperatures above 1,550°C. Young's modulus increased due to the compositing with whiskers as predicted from the rules of compositing. The change of Young's modulus with sintering temperature for the composite parallels the change in the matrix.

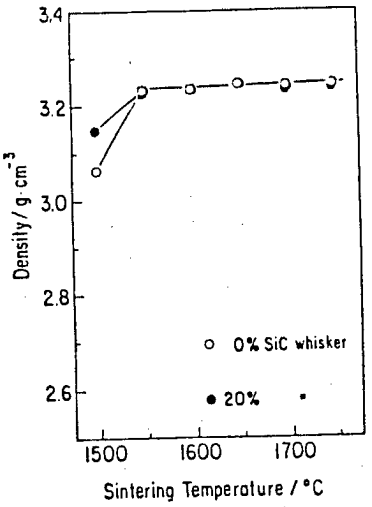


Figure 1. Relationship of Density and Sintering Temperature

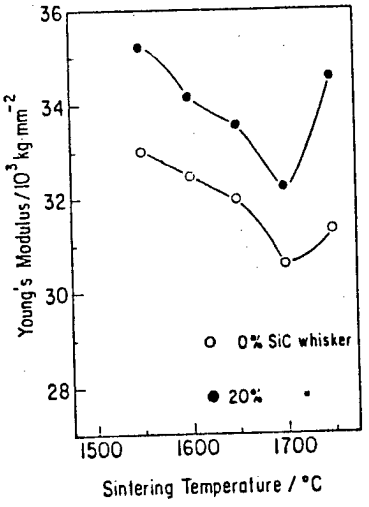


Figure 2. Relationship of Young's Modulus and Sintering Temperature

The bending strength and fracture toughness, K_{IC}, are shown in Figures 3 and 4, respectively. Both are seen to increase with the addition of whiskers in a low sintering temperature region. However, as sintering temperature reaches 1,650 to 1,700°C, both properties assume the same value levels as the matrix and do not exhibit any benefit of whisker addition at temperatures higher than that.

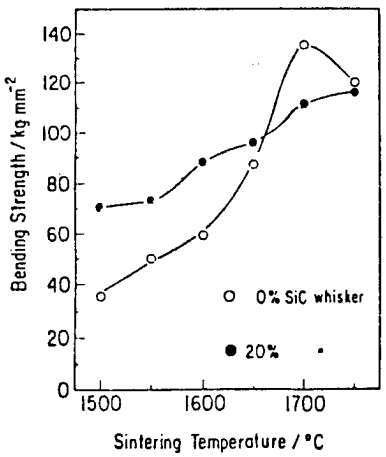


Figure 3. Relationship of Bending Strength and Sintering Temperature

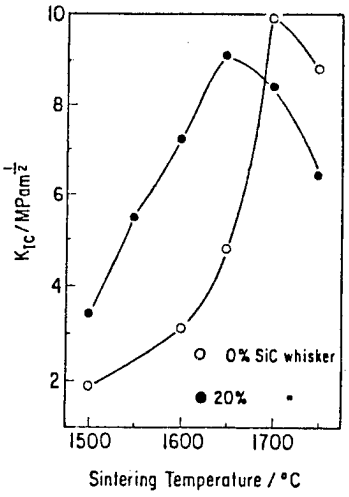


Figure 4. Relationship of Fracture Toughness and Sintering Temperature

The four graphs of Figure 5 show the relationships of density, Young's modulus, bending strength, and fracture toughness to the SiC whisker content when it is varied from 0 to 40 percent for a sintering temperature of 1,600°C. Because densification was not achieved at a whisker content of 40 percent, all other property values at 40 percent whisker content turned out to be low. Young's modulus should increase with the whisker content once densification is achieved. According to the graphs, both bending strength and fracture toughness show their respective maxima at a 20 percent whisker content.

The mechanical properties of the matrix are shown to be of lower values in a low sintering temperature region. The reasons could be the incomplete transition from α to β silicon nitride (Figure 6), where the microstructure of the matrix appears to consist of isotropic particles without pillar-shaped particles of β-silicon nitride, as shown in Figure 7 [not reproduced]. As the sintering temperature rises, the transition from α to β proceeds and the properties of the matrix improve due to the presence of the pillar-shaped particles in the microstructure. This effect is not readily manifested when whiskers are mixed, probably because changes in the whiskers themselves take place during the sintering process at a high temperature.

3.2 Orientation of SiC Whiskers

The orientation of SiC whiskers was evaluated by measuring X-ray diffraction of molded sheets after drying. In the X-ray diffraction peaks of β-SiC,

attention was focused on the peaks due to (111) plane and (220) plane to obtain a diffraction intensity ratio of I_{220}/I_{111}. Because the growth direction of whiskers is <111> and the plane (111) intercepts the plane (220), which is equivalent to the plane (220), the ratio I_{220}/I_{111} becomes greater as more whiskers are oriented strongly along the sheet plane.

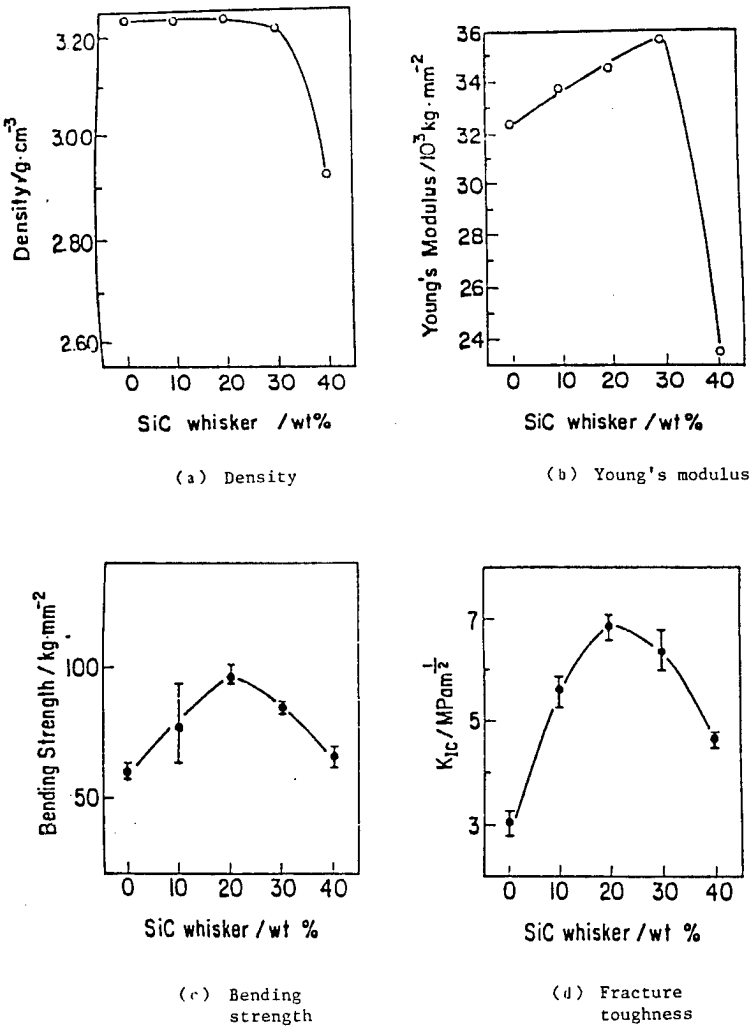


Figure 5. Relationships of Various Characteristics to Whisker Content (Sintering temperature: 1,600°C)

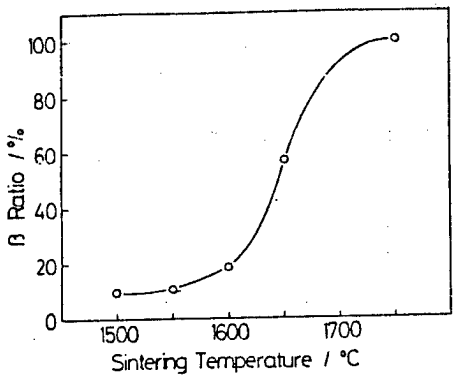


Figure 6. Relationship of β Phase Content and Sintering Temperature

The relationship between the ratio I_{220}/I_{111} and the sheet thickness is shown in Figure 8. When whiskers are totally randomly oriented, the I_{220}/I_{111} ratio is 0.6. On the other hand, as in Figure 8, the ratios are always greater than one regardless of the sheet thickness. Therefore, it is clear that SiC whiskers are oriented. Also, the extent of orientation is shown to be dependent on the sheet thickness; within the scope of this experiment the thicker the sheet, the more extensive the orientation. When a mixed dough containing SiC whisters is passed through a nozzle with a thin gap, it is thought, the resulting mold will contain highly oriented whiskers because they tend to lie flat. Therefore, it had originally been thought that molds had to be made rather thin to obtain a high degree of orientation of whiskers. Because of the complex relationship of many factors including the viscosity and fluidity of the dough, our results showed that the thinning would not necessarily improve the orientation. However, as the thickness increases further, less stress should be applied to whiskers during extrusion; thus, it is conjectured that the extent of orientation will decline after passing a peak at a certain mold's thickness. Figure 9 [not reproduced] is a photograph of a molded sheet as observed with a scanning electron microscope. The arrow indicates the direction of extrusion and many whiskers can be seen oriented in that direction. Thus, we confirmed that whiskers could be oriented through extrusion molding and that the extent of orientation could be controlled by adjusting the thickness of an extruded sheet.

Figure 10 shows the relationship between the fracture strength of samples, made by layering molded sheets and sintering, and the sintering temperature. For comparison purposes, the same relationship for samples with randomly oriented whiskers is also shown. Due to the influence of whisker orientation, strong anisotropy emerged. When test pieces were cut parallel to the orientation direction, they were stronger than test pieces with random whisker orientation. Particularly at a sintering temperature of 1,650°C, a high strength value exceeding 130 kg/mm^2 was obtained.

This research project was carried out as a part of the "Elemental Technology Development for a Ceramics Turbine for Coal Gasification," which was commissioned to the Fine Ceramics Technical Research Association under the Next Generation Industrial Key Technology R&D System of the Agency of Industrial Science and Technology, MITI.

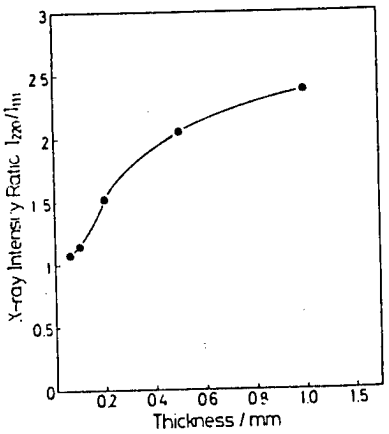


Figure 8. X-Ray Diffraction Intensity Ratio and Sheet Thickness

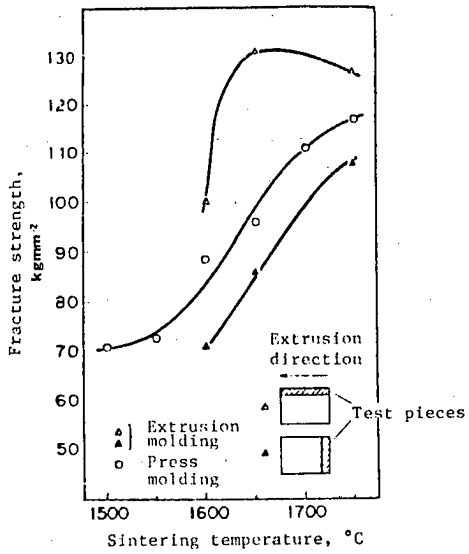


Figure 10. Relationship Between Strength and Sintering Temperature

Si_3N_4-SiC Nanocomposites

43067065c Tokyo 8TH HIGH TEMPERATURE MATERIALS KISO TORONKAI in Japanese 10-11 Nov 88 pp 100-104

[Article by Koichi Shinhara, Atsushi Nakahira, and Tsuyoshi Hirano of National Defense Academy, and Kansei Izaki and Tansei Kawakami of Mitsubishi Gas Chemical Co., Inc.]

[Text] 1. Introduction

Many attempts have been made to apply Si_3N_4 and SiC to turbocharger rotors and to other areas as high temperature structural materials. Of the two, Si_3N_4 is superior in room temperature strength, fracture toughness and heat shock resistance, but inferior to SiC in high temperature strength and oxidation resistance. Therefore, attempts are being made to improve these characteristics by compositing these two materials.

Recently, we successfully prepared a Si_3N_4/SiC nanocomposite by dispersing SiC particles of several tens of nanometers in diameter in Si_3N_4 particles by means of hot pressing Si-C-N composite powder obtained by the CVD method. The structure of this nanocomposite and the influence of SiC upon the improvement of various characteristics of the composite is reported in this paper.

2. Experimental Method

2-1 Synthesis and Evaluation of Si-C-N Composite Powder

As gaseous raw materials, hexamethyl-di-silazane and ammonia were used. These materials were vaporized and introduced into an Al_2O_3 reaction vessel maintained at 1,000°C for gas phase reaction to produce Si-C-N composite powder. By changing the amount of ammonia, several kinds of powders with different C/N ratios were prepared. The synthesized powders were stabilized by heat treatment in a nitrogen atmosphere. The powders were subjected to chemical analyses, X-ray diffraction (XRD), fluorescent X-ray, BET and scanning electron microscopy (SEM).

2-2 Preparation of Si_3N_4/SiC Composite

Sintering aids consisting of 6 percent by weight of Y_2O_3 (Japan Yttrium Co., Ltd.) and 2 percent by weight of Al_2O_3 (Sumitomo Chemical Co., Ltd.) were added to a synthesized powder, and the mixture was blended with ethanol as a dispersant, with Si_3N_4 balls, in a ball mill and placed in a polyethylene container. After drying, the mixture was dry-blended in a ball mill and passed through a No 50 mesh screen. The particles that passed through were placed in a reduced pressure atmosphere at 100°C to remove adsorbed moisture. The powder thus obtained was hot-pressed for 2 hours at 1,800°C in a N_2 atmosphere under a pressure of 34 MPa.

2-3 Evaluation of Characteristics

The mechanical properties of hardness, strength and fracture toughness were measured as follows. A micro-Vickers tester and a high temperature hardness tester were used to measure hardness from room temperature through 1,150°C, the three-point bending method was used to evaluate strength from room temperature through 1,500°C, and the IM and SEPB methods were used to determine fracture toughness.

Concerning thermal properties, the laser flash method was used to determine thermal conductivity from room temperature through 1,000°C, and a push rod-type thermal expansion meter was used to determine the linear expansion coefficient from room temperature through 1,400°C.

To observe microstructures, XRD, an optical microscope (OM), an SEM, a transmission electron microscope (TEM), and a high resolution electron microscope (HREM) were used.

3. Experimental Results

The composite powder consists of spherical microparticles of approximately 0.2 micron in diameter. The powder was found by XRD to be non-crystalline. Even after the heat treatment, no sign of crystallization was observed. Component analyses revealed that the powder was a highly pure Si-C-N composite powder containing an extremely small amount of metallic impurities. As a result of XRD, it was found that the ceramic material consisted of β-Si_3N_4 and β-SiC and, the Si_3N_4/SiC composite was obtained by hot pressing the composite powder. The relative densities of the composites did not depend on the C/N ratios of the composite powders but all were more than 98 percent. A gradual increase of hardness was observed with the increase of SiC volume fraction.

Figure 1 shows the effects of SiC volume fraction on strength and fracture toughness. Strength increased up to 10 percent of SiC volume fraction and remained nearly constant above 10 percent. On the other hand, fracture toughness increased with the increase in SiC volume fraction up to 10 percent and then decreased above 10 percent to approach the value of a Si_3N_4 single phase.

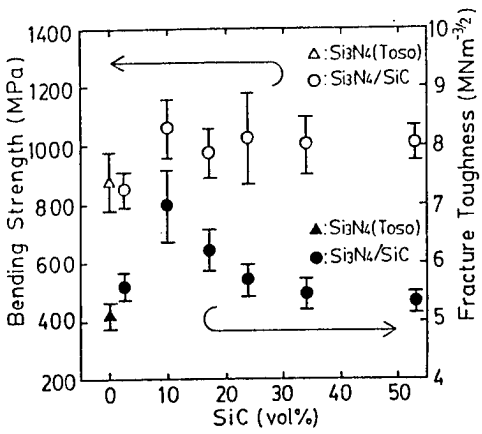


Figure 1. Effects of SiC Volume Fraction on Bending Strength and Fracture Toughness Values

Figure 2 shows changes in the microstructure with the change in SiC quantity. When a Si_3N_4 single phase was hot-pressed in the same way as the composite powder, the microstructure, as shown in Figure 2 (a), showed well-developed pillar-shaped crystals and an average aspect ratio of a relatively high 4.7, although some particles appeared to have grown abnormally, in some cases, to become large, crude particles of 3.0 microns in width and 8.0 microns in length. As the SiC volume fraction increased, the micronization of the structure advanced with the growth of an increased number of Si_3N_4 pillar-shaped particles. As seen in Figure 2 (b), at 10 percent SiC, the structure was micronized and pillar-shaped particles were well developed to improve the average aspect ratio to 6.0. There were neither the large, crude particles that were seen in the single phase structure, nor isotropic particles. However, as the SiC volume fraction further increased above 10 percent, the growth of pillar-shaped particles was increasingly impeded and more isotropic particles appeared, as shown in Figure 2 (c), at 34 percent SiC volume fraction, where the aspect ratio also decreased to 3.0.

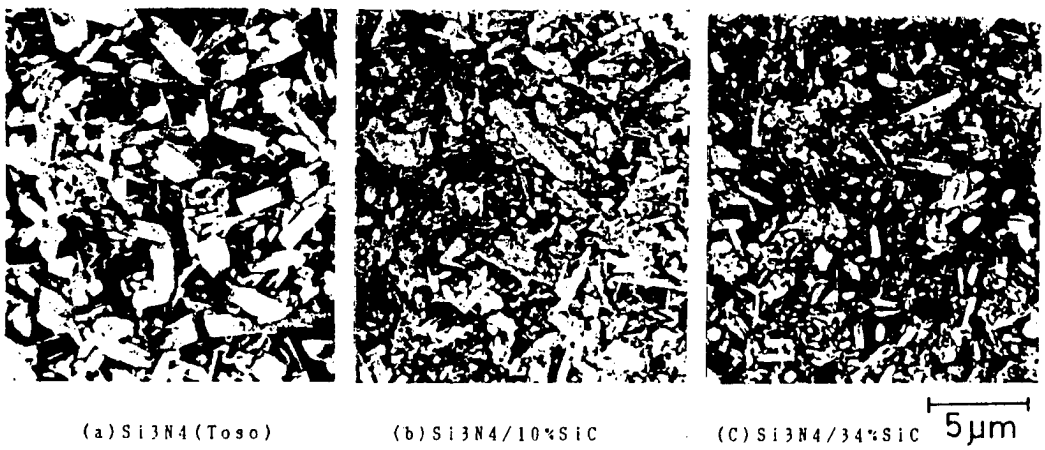


Figure 2. Effects of SiC Volume Fraction on Structure

The relationship of strength to temperature is shown in Figure 3. The addition of SiC suppressed the loss of strength in the high temperature region, and the sample containing 34 percent SiC maintained a strength of 1,200 MPa from room temperature through 1,200°C. The addition of SiC also suppressed the deterioration of hardness in the high temperature region.

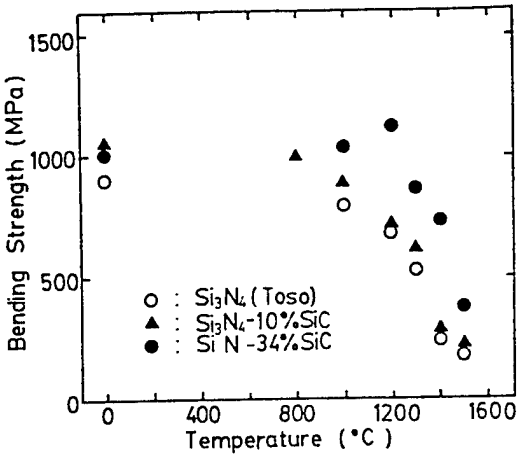


Figure 3. Temperature Dependence of Strength

Similar to the Si_3N_4 single phase material tested for comparison, thermal conductivity was negatively dependent on temperature, though the composite's conductivity values were one-half to one-third of that of the single phase. Based on the rules of compositing, the thermal conductivity of the composite had been predicted to increase with the addition of highly conductive SiC. However, as shown in Figure 4, a different result was obtained. On the other hand, the composite showed a linear expansion coefficient value similar to SiC.

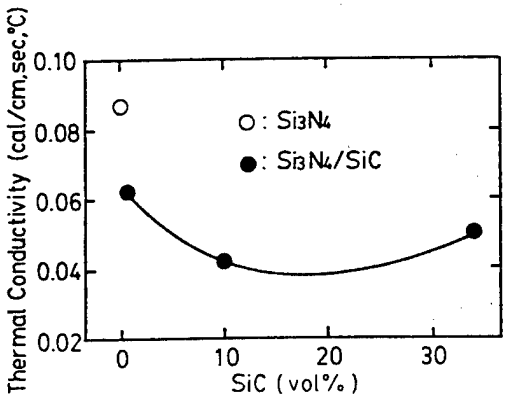


Figure 4. Effect of SiC Volume Fraction on Thermal Conductivity

It was learned through TEM and HREM observation that the composite was a nanocomposite, i.e., a material composited at the nanometer level, in which, as shown in Figure 5 [not reproduced], microparticles, several tens of nanometers in diameter, of SiC were dispersed not only around but also inside Si_3N_4 particles. Furthermore, no significant reaction phase was observed between Si_3N_4 and SiC particles within Si_3N_4.

4. Discussion

4-1 Effects on Structural Conditions

In the composite, SiC is thought to play the following two roles concerning the structural conditions of the composite: (a) to suppress particle growth and micronize the structure, and (b) to become a nucleus for the elution of β-Si_3N_4 from the liquid phase. Role (a) was clearly demonstrated by the micronization of the structure with the increase of SiC. It is conjectured for role (b) that pillar-shaped particles grow from the nuclei of SiC particles to form a three-dimensional intertwined structure, when SiC is less than 10 percent; whereas, in cases of high SiC content, the growth of pillar-shaped particles is suppressed because of too many nuclei. The conjecture is based on the fact that SiC particles are dispersed within Si_3N_4 particles and the growth of pillar-shaped particles are promoted when the SiC content of the composite is 0 to 10 percent but suppressed when the SiC content exceeds 10 percent. For this reason, the structural condition varies with the differences in the volume of SiC, and, it is thought, each mechanical property reflects these changes.

4-2 Effects on Mechanical Properties

Hardness increases with the addition of SiC almost according to the rules of compositing. Strength and fracture toughness increase due to the growth of pillar-shaped particles when the SiC volume fraction is from 0 to 10 percent, but as the SiC volume fraction exceeds 10 percent, fracture toughness declines because the growth of pillar-shaped particles is suppressed. However, it is conjectured that, as the overall structure becomes increasingly micronized, the sources of fracture will be restricted to a smaller area and, therefore, the material's strength will not decrease. Also, the increase of fracture toughness in a less than 10 percent SiC volume fraction region appears to have been caused by the toughness improving mechanisms such as the curvature or inclination of the leading edge of a crack. In this composite, in which SiC is dispersed not only around but within Si_3N_4 particles, it is quite possible that not only is the overall structure toughened but each crystalline particle is also toughened. Furthermore, softening of the interparticulate phase and the accompanying slow crack growth, which are the usual causes for the loss of strength in a high temperature region, appear to be inhibited by the presence of SiC particles in that phase.

4-3 Effects on Thermal Properties

When Si_3N_4 and SiC are simply mixed, the mixture's thermal conductivity should increase with the increase in its SiC content, according to the rules of compositing. However, this particular composite gave us results which were different from the predicted results based on the rules. The reason for this is conjectured as follows: because of the dispersion of SiC particles not only around but within Si_3N_4 particles, phonon scattering is increased not only at the interface between Si_3N_4 and SiC particles but also because of the strain on the Si_3N_4 lattice. Although thermal conductivity initially

decreases with an increasing quantity of SiC, conductivity hardly changes above 20 percent SiC content. The reason for the apparent lack of change in thermal conductivity could be explained as follows. With the increase in SiC content, the quantity of SiC dispersed within Si_3N_4 particles will increase to intensify phonon scattering to diminish thermal conductivity of each crystalline particle; however, at the same time, the number of SiC particles around Si_3N_4 particles will increase and these SiC particles will help increase thermal conductivity, instead of helping increase phonon scattering. Thus, the overall thermal conductivity of the composite will not change.

5. Summary

A Si_3N_4/SiC composite was prepared by hot-pressing amorphous Si-C-N composite powder synthesized by the CVD method. The effects of SiC on the composite's structural conditions and various properties were examined, with the following conclusions.

(1) In this experiment, a Si_3N_4/SiC nanocomposite was obtained in which SiC particles of several tens of nanometers in diameter were dispersed around as well as within the Si_3N_4 particles.

(2) The volume fraction of SiC significantly affected the structural conditions of the composite. In the 0 to 10 percent SiC region, the growth of pillar-shaped Si_3N_4 particles was promoted, fracture toughness was drastically improved to a maximum value at 10 percent, and the growth of large, crude particles, such as those observed in a Si_3N_4 single phase, was suppressed to micronize the structure to improve strength. As the SiC content exceeded 10 percent, the growth of pillar-shaped particles was inhibited and the fracture toughness value of the composite diminished to approach that of Si_3N_4 single phase while the composite's strength did not diminish because the micronization of the structure advanced.

(3) The composite's hardness increased gradually with the SiC content. It is conjectured that the suppression of the decline of the composite strength in the high temperature area was due to the presence of SiC particles around Si_3N_4 particles.

(4) The composite's thermal conductivity decreased from that of Si_3N_4 single phase from zero up to 10 percent of SiC content and it remained almost unchanged at SiC contents higher than 10 percent.

Fabrication of YBCO Fibers by Sol-Gel Method Using Metal Alkoxide

43067068a Tokyo 1988 FALL CONVENTION OF JAPAN SOCIETY OF POWDER AND POWDER METALLURGY in Japanese 9-11 Nov 88 p 29

[Authors: Yoshio Masuda, Kobe Steel, and Tsuyoshi Tateichi, Koberuko Scientific Laboratory]

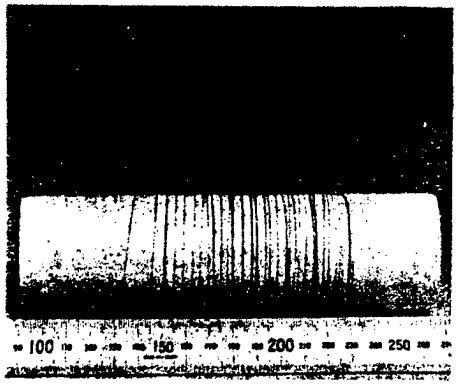
[Text] 1. Introduction: Various methods, such as the metal sheath method[1] and the suspension spinning method,[2] have been examined to fabricate $YBa_2Cu_3O_x$ superconducting fibers. We tried to fabricate superconducting fibers by a sol-gel method using metal alkoxides and acetic acid salts. We report our results on a fiber fabrication method and a process that forms the $YBa_2Cu_3O_x$ superconducting phase.

2. Experimental Method: The required amounts of $Y(O_iC_3H_7)_3$, $Ba(O_iC_2H_5)_2$, and $Cu(OCOCH_3)_2$ were dissolved in organic solvents such as 2-propanol, DMF, ethylene glycol, and naphthenic acid and were made into a uniformly mixed solution. Conditions of the solution suitable for fabricating fibers were examined by studying the properties of the sol or particles that were formed when the pH of the solution was altered using acetic acid, hydrochloric acid, or ammonia water. A solution, in which [the required materials] were mixed with acetic acid [chemically] equivalent to four times the total amount of the metals in the materials, was heated in air at 60°C to form a viscous transparent sol. Gel fibers were spun directly from this sol. The gel fibers were then dried in vacuum at 60°C, fired under oxygen flow at 900--950°C, and fabricated into ceramic fibers of the $YBa_2Cu_3O_x$ phase. The transformation of the gel into $YBa_2Cu_3O_x$ ceramic fibers was studied using TG-DTA, FT-IR, and x-ray diffraction, and microscopic conditions were observed using a scanning electron microscope (SEM). Also, the temperature dependence of magnetization was measured using a vibrating magnetometer (VSM) after ceramic fibers were made into a powder.

3. Results and Discussion: A sol that was highly uniform and could be spun well was formed from the solution made with acetic acid. Fibers longer than 5 m were formed from this type of sol, and the fiber diameters were between 10 microns and 1 mm. The appearance of gel fibers [thus formed] are shown in Fig. 1. Since a rapid weight reduction of about 70 percent was observed up to approximately 420°C in the TG curve, fibers had to be heated at the rate of 30°C/hour or less below 400°C. Otherwise, the shape of the fibers was

difficult to maintain because the fibers melted or bubbled. When the sol was heated to 300°C, carbonyl radicals appeared in the polymerized products and Cu and Cu_2O crystals were precipitated. When the sol was heated to 600°C, a carbonate, $BaCO_3$, and Y_2O_3 were formed, and the copper was oxidized as CuO. When the temperature was raised to 800°C, the carbon bonds almost disappeared, and $YBa_2Cu_3O_x$ and a small amount of $BaCuO_2$ were formed. At 900°C, only the $YBa_2Cu_3O_x$ phase remained. The magnetization of the $YBa_2Cu_3O_x$ fibers that were fired for five hours in an oxygen flow at 900°C was measured and the T_c (onset) of the fibers was found to be 92 K.

Fig. 1. Appearance of the gel fiber.



References

1. O. Kohno et al., Jap. J. Appl. Phys. 27, L77 (1988).
2. T. Goto and M. Kada, Jap. J. Appl. Phys. 26, L1527 (1987).

Chemical Processing of Zirconia Ceramic Composites by Colloidal Method

43067068b Tokyo 1988 FALL CONVENTION OF JAPAN SOCIETY OF POWDER AND POWDER METALLURGY in Japanese 9-11 Nov 88 p 66

[Authors: Shinichi Hirano, Takashi Hayashi, and Chikage Kato, Faculty of Engineering, Nagoya University]

[Text] 1. Introduction: Ceramics that contain zirconia are studied in many systems because they are strong and tenacious. However, it is difficult to disperse zirconia uniformly in a matrix using conventional methods.

We identified the conditions under which zirconia preferentially precipitates on the surface of silica spheres highly dispersed in a solution when zirconium alkoxide is hydrolyzed. These conditions were applied to synthesize mullite-zirconia composite particles. The mullite-zirconia composite particles thus formed were used to synthesize a ceramic in which zirconia is dispersed uniformly in a mullite matrix. The microstructure of the ceramic was observed and its mechanical properties (Vickers hardness H_v and fracture toughness K_{1c}) were evaluated.

2. Experiment: $Zr(O^iPr)_4$ and $Y(O^iPr)_3$ were measured by balance so that there was 4 molar percent of Y_2O_3-ZrO_2. These were stirred and refluxed in i-PrOH. Colloidal silica dispersed in n-BuOH was mixed and stirred into this alkoxide solution, and the mixture was hydrolyzed using water with a pH of 4. A similar operation was performed on a mullite/i-PrOH dispersed solution, and the composite particles formed were separated by ultrafiltration using water with a pH of four. Molding was done with the following two methods. (1) Uniaxial molding --> CIP processing. (2) Redispersion in water that contains a surface active agent and colloid pressing --> CIP processing. The material was sintered for five hours at 1600°C. The zeta potential of such particles was measured by micro-electrophoresis and their specific surface area by the BET method. The shape and form of the particles and the microstructure of the sintered products were observed using a scanning electron microscope (SEM) and a tunneling electron microscope (TEM). The H_v and K_{1c} of the product were measured by the IM method.

3. Results: The relationship between the zeta potentials and pH of the SiO_2 particles, $4YZrO_2$ particles, and the $SiO_2/4YZrO_2$ composite particles is shown in Fig. 1. The zeta potential of the SiO_2 particles is negative for all

ranges of pH, but the $SiO_2/4YZrO_2$ composite particles exhibit an isoelectric point at a pH of 7 similar to the curve of the $4YZrO_2$ particles, indicating that the surfaces of the SiO_2 particles are uniformly coated with $YZrO_2$. Metal alkoxides are preferentially hydrolyzed and precipitated on the surfaces of spherical SiO_2 particles, which are highly dispersed in a liquid. With this process, uniform $SiO_2/4YZrO_2$ composite particles whose surfaces were modified with zirconia were formed (Fig. 2). This processing was also applied to the mullite/$4YZrO_2$ system. The fact that the mullite/$4YZrO_2$ composite particles obtained by this method had their mullite particle surface modified by $4YZrO_2$ was confirmed through a dark field image observation with a TEM. The mullite/$4YZrO_2$ composite particle products formed by the colloid pressing --> CIP processing are densely filled. When the material was sintered at 1600°C, its density reached more than 98 percent of the theoretical density and exhibited a system in which $4YZrO_2$ was uniformly dispersed. The K_{1c} value of mullite alone was 2.2 $MN/m^{3/2}$, while that of mullite with 20 volume percent of $4YZrO_2$ exhibited high values exceeding 5.0 $MN/m^{3/2}$.

Figs. 1 and 2

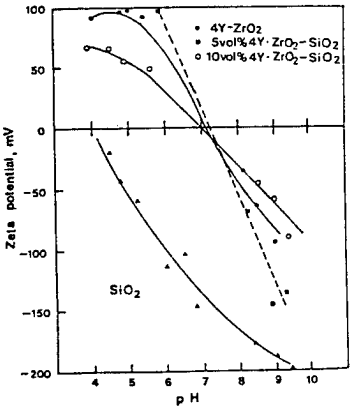


Fig. 1

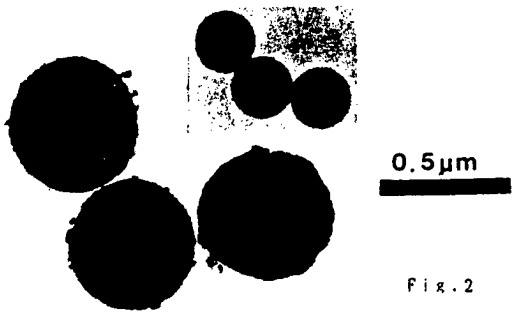


Fig. 2

Amorphous Alloy Powder of Heat Resistant Intermetallic Compounds by Mechanical Alloying

43067068c Tokyo 1988 FALL CONVENTION OF JAPAN SOCIETY OF POWDER AND POWDER METALLURGY in Japanese 9-11 Nov 88 p 132

[Authors: Hiroshi Kimura and Masayoshi Kimura, Mechanical Engineering Department, National Defense Academy]

[Text] Purpose: The authors are conducting a series of studies on the creation of Co, Fe, and Ni alloys in the form of solid-phase amorphous alloys and their high-density formation using a stirring type high-energy ball mill. In this study, we attempted to create uniform amorphous alloy powders of intermetallic compounds that contain Ti by mechanical alloying of powder mixtures of elements and compound powders. Such alloy powders are expected to become heat-resistant light structural materials.

Method: We used a stirring type high-energy ball mill with a processing capacity of 1 Kg powder at 200--500 rotations [per minute] and attempted to create amorphous powders by mechanically alloying or mechanically grinding a mixture of metallic elements or intermetallic compound powders. The solid-phase amorphous process was monitored in real time as a function of the milling time using the tank temperature and torque. A structural analysis of the mechanically alloyed amorphous powders was conducted by x-ray analysis, and the thermal stability was studied using a difference scanning calorimeter (DSC).

Results: Figure 1 represents the monitoring of the solid-phase amorphous process at various rotational speeds when an intermetallic compound powder of TiNi was mechanically ground. During the early stage of milling in the mechanical alloying process involving intermetallic compounds, a period (4 hours 20 minutes to 7 hours) with a constant temperature (96°C) appears. This period is similar to the plateau period during the mechanical alloying of powders of the distinct metal elements Ti and Ni. The period can be considered to be a latent period preceding a reaction.

The variation of the DSC curves when the samples were mechanically ground at 200 rpm and 300 rpm as a function of the milling time is shown in Fig. 2. At the point in the plateau period corresponding to 5 hours 20 minutes [of milling time], no crystalline peak indicating an amorphous reaction was

observed. Figure 3 indicates the JMA plot of the amorphous transformation (x) deduced from the heat generated by mechanical grinding. At a high speed rotation of 300 rpm, a two-stage process (n = 1.6 and 3.15) becomes evident. The particle diameter distribution of the amorphous NiTi powder [produced] by mechanical alloying and mechanical grinding is shown in Fig. 4. Mechanical grinding produced finer particles. A detailed analysis based on the [reaction] speed will be reported at the conference along with results for TiAl.

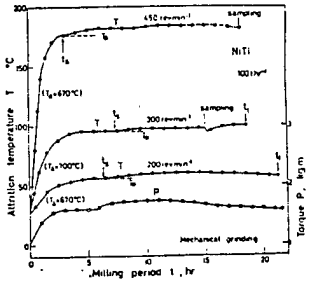


Fig. 1. Solid-phase amorphous process by mechanical alloying of TiNi intermetallic compound powder.

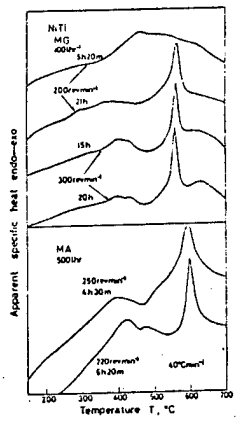


Fig. 2. Variation of the DSC curves when NiTi is mechanically alloyed and ground.

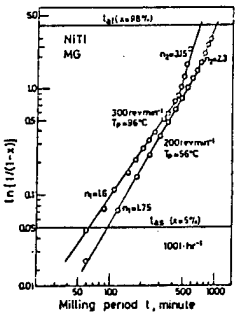


Fig. 3. JMA plot when a solid-phase amorphous transformation was caused by mechanical grinding.

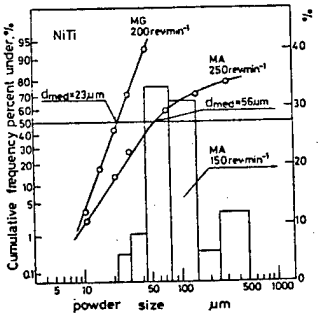


Fig. 4. Size distribution of NiTi amorphous powder produced by mechanical alloying and mechanical grinding.

Epitaxial Growth of SiC Thin Films on Si Substrate by Laser CVD

43067068d Tokyo 1988 FALL CONVENTION OF JAPAN SOCIETY OF POWDER AND POWDER METALLURGY in Japanese 9-11 Nov 88 p 185

[Authors: Takeshi Kuragaki, Hirohide Nakamatsu, Jie Tang, and Shichio Kawai, Industrial Science Laboratory, Osaka University]

[Text] 1. Objective: SiC, which has excellent heat resistance, is considered a semiconductor material that can operate at high temperatures. Here, we report an epitaxial growth of the 3C-type SiC at a relatively low synthesis temperature, 1150°C, by laser CVD [chemical vapor deposition]. A thin film was fabricated by CVD after a carbonized layer was formed on a Si substrate. The structure and crystalline properties of the carbonized layer and the epitaxial film were studied, and we report about the relationship between the epitaxial growth and the crystalline properties of the carbonized layer.

2. Experiment: Thin films were fabricated by an ultralow pressure laser CVD device. Si(111) was used as a substrate, and Si_2H_6 and C_2H_2 were blown against the substrate from nozzles as reaction gases. The pressure in the device was lowered to 10^{-4} Pa, the substrate heated by electron-beam impact, and the temperature measured with an optical pyrometer. The carbonizing reaction was carried out by heating the substrate to 1150°C after C_2H_2 gas was introduced. After the reaction gases were introduced, CVD was applied at 1150°C by irradiating the substrate with an ArF excimer laser at a pulse rate of 180 Hz while blowing the reaction gases against the substrate.

3. Results: First, a carbonized Si film was fabricated by varying the reaction time while the C_2H_2 gas pressure was set at 5.0×10^{-3} Pa. A diffraction pattern of a 3C-type SiC single crystal was detected by reflective electron-beam diffraction (RED), and a sharp peak was observed in the infrared reflection spectrum of the thin film at 794 cm^{-1} after three minutes of reaction. The peak for this film was sharper than that observed for a film after five minutes of reaction (Fig. 1). Moreover, when CVD was performed on these carbonized layers, the films with five minutes of carbonizing became twins and the films with three minutes of carbonizing became single crystals. Following this, we set the carbonizing time at three minutes and studied CVD conditions. If we set the C_2H_2 gas pressure at 2.0×10^{-3} Pa and let r be the ratio of the reaction gas pressures (Si_2H_6/C_2H_2), then single crystal thin

films grew at r = 0.5--1.3 (Fig. 2). At r = 1.5, the thin films were twins. The film thickness was about 0.1 micron and stable at r = 0.7--1.3, but the thickness became irregular at r = 0.5.

From these results, we confirmed that thin films became either single crystals or twins depending on very slight differences in the crystalline conditions of the films. No twins were formed when the reaction gas pressure ratio was between 0.7 and 1.3, but we noticed the growth of epitaxial thin films of a uniform thickness. The effect of laser irradiation will also be discussed.

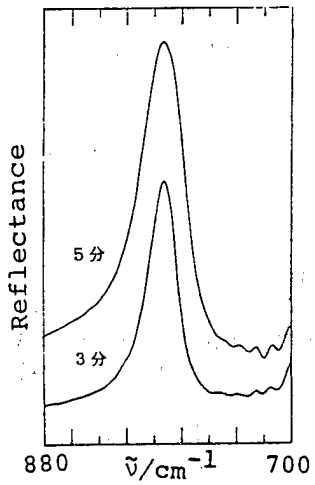


Fig. 1. Infrared reflection spectrum.

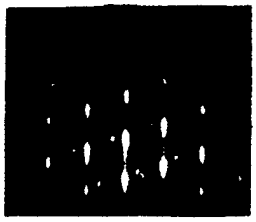


Fig. 2. Photograph of reflective electron-beam diffraction.

Al_2O_3/Si_3N_4 Nanocomposite

43067068e Tokyo 1988 FALL CONVENTION OF JAPAN SOCIETY OF POWDER AND POWDER METALLURGY in Japanese 9-11 Nov 88 p 192

[Authors: Atsushi Nakahira and Koichi Niihara, National Defense Academy, and Jun Okijima and Toshio Hirai, Metallurgy Laboratory, Tohoku University]

[Text] 1. Introduction: When compared with other ceramics, Al_2O_3 has low fracture toughness and bending strength. It also weakens at high temperature from slow crack growth. The authors have been trying to reduce these defects by nanocomposition. We dispersed microparticles of Si_3N_4 in Al_2O_3 and fabricated Al_2O_3/Si_3N_4 nanocomposites using hot pressing. We report here the microstructure and various mechanical and thermal properties of these nanocomposites.

2. Experiment: Al_2O_3 powder made by Asahi Chemical Industry and Si_3N_4 powder provided by Stark were used as raw materials. The powder components of the composite were mixed in a ball mill and hot pressed under a pressure of 28 MPa at 1500--1800°C in a nitrogen atmosphere. The microstructure of the samples produced were observed by SEM, TEM, and HREM [high resolution electron microscopy]. The fracture toughness of the samples was measured by the IM method, bending strength by 3-point bending tests, thermal conductivity by the laser flash method, and thermal expansion coefficient by a push rod dilatometer.

3. Results: The microstructure, such as the composition phase and particle size, of the composite materials formed varied greatly depending on the amount of Si_3N_4 added and the temperature of the hot press. As shown in Fig. 1, the sample that was hot pressed at 1600°C clearly exhibits a nanocomposite structure in which Si_3N_4 particles are dispersed among Al_2O_3 particles. The fracture toughness and bending strength of this Al_2O_3/Si_3N_4 nanocomposite are 5 $MN/m^{3/2}$ and 500 MN/m^2, respectively. These values are about 1.5 times greater than those of Al_2O_3. Bending strengths measured at high temperatures are shown in Fig. 2. The figure illustrates that our Al_2O_3/Si_3N_4 nanocomposites maintained a bending strength the same as the bending strength at room temperature until 1200--1300°C. We interpret this fact to be due to the suppression of slow crack growth by the nanodispersion of Si_3N_4.

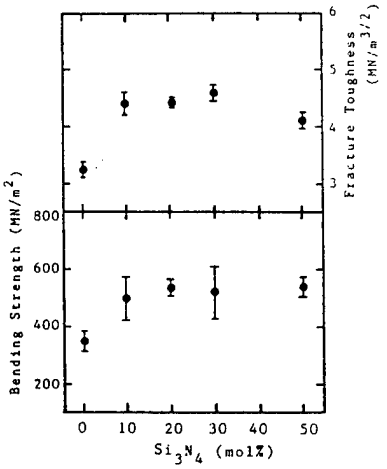


Fig. 1. Fracture toughness and bending strength of Al_2O_3/Si_3N_4 nanocomposites.

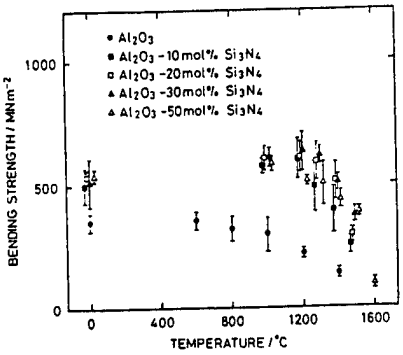


Fig. 2. Bending strength of Al_2O_3/Si_3N_4 nanocomposites at high temperatures.

Nanostructure, Thermal and Mechanical Properties of Si_3N_4/SiC Composites

43067068f Tokyo 1988 FALL CONVENTION OF JAPAN SOCIETY OF POWDER AND POWDER METALLURGY in Japanese 9-11 Nov 88 p 193

[Authors: Koichi Niihara, Takeshi Hirano, Atsushi Nakahira and Katsuaki Suganuma, National Defense Academy, and Kansei Izaki and Takamasa Kawakami, Mitsubishi Gas Chemical]

[Text] 1. Introduction: Research and development on Si_3N_4 and SiC as high temperature structural materials are being carried out. In recent years, organic silicon compounds have become popular as the raw materials for Si_3N_4 and SiC. Recently, we have succeeded in fabricating Si_3N_4/SiC nanocomposites in which SiC particles of tens of nanometers in diameter are dispersed among the Si_3N_4 particles by hot pressing the Si-C-N precursor powder obtained by CVD. In this talk we report on the nanostructure and thermal and mechanical properties of these composites.

2. Method: The Si-C-N precursor powder was synthesized using a gas-phase reaction in a reaction tube made of Al_2O_3 and maintained at 1000°C by placing hexamethyldisilane and ammonia in the tube. Several types of powder with different C/N ratios were synthesized by varying the amount of ammonia, and the products were stabilized by thermal processing in an N_2 atmosphere. As a sintering auxiliary, 6 weight percent Y_2O_3 and 2 weight percent Al_2O_3 were added and hot pressed for two hours at 1800°C under a pressure of 34 MPa in a nitrogen atmosphere. As for the mechanical properties, hardness was measured by a high-temperature durometer from room temperature to 1150°C, and bending strength was measured by the three-point bending method from room temperature to 1500°C. Fracture toughness was measured by the IM method and by the SEPB method. Thermal conductivity was measured by the laser flash method from room temperature to 1000°C, and the linear expansion coefficient was measured by a push-rod dilatometer from room temperature to 1400°C. The phase of formation was identified by x-ray diffraction, and the microstructure was observed using an optical microscope, a scanning electron microscope, a transmission electron microscope, and a high-resolution electron microscope.

3. Results: By using x-ray diffraction, we found that an Si_3N_4/SiC composite was formed by hot pressing the Si-C-N precursor powder and that the resulting sintered product consisted of beta-Si_3N_4 and beta-SiC. The bending strength [of the sample] increased by up to 10 percent as the volume fraction of SiC

was increased, but the strength remained roughly constant beyond that. The fracture toughness increased by up to 10 percent as the volume fraction of SiC was increased. However, when the fraction exceeded 10 percent the fracture toughness decreased, approaching the toughness value of single-phase Si_3N_4. Figure 1 shows the temperature dependence of thermal diffusivity. As in the case of single-phase Si_3N_4, the temperature dependence of our composite was negative but its actual value was approximately one half that of single-phase Si_3N_4. The specific heat hardly changed at room temperature, and its temperature dependence was positive in both cases. Nevertheless, tendency for the single-phase value to increase was observed when the temperature was raised. Because of this, the thermal conductivity of the composite was about 1/2 to 1/3 of that for the single-phase case. We believe that these properties are somehow related to the microstructure peculiar to the composite.

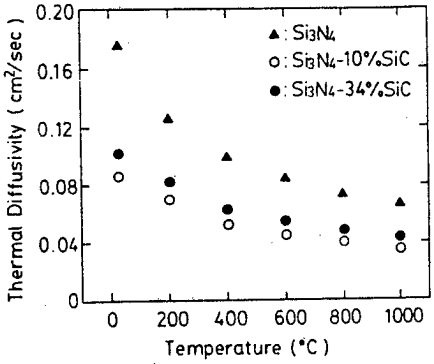


Fig. 1. Temperature dependence of thermal diffusivity.

Dispersion and Sintering Properties of Al_2O_3-SiC Whisker Composite Ceramics

43067068g Tokyo 1988 FALL CONVENTION OF JAPAN SOCIETY OF POWDER AND POWDER METALLURGY in Japanese 9-11 Nov 88 p 194

[Authors: Toshio Ishii, Magnetic Materials Laboratory, Hitachi Metals, and H. K. Bowen, Massachusetts Institute of Technology]

[Text] 1. Objective: Recently, Al_2O_3-SiC whisker composite ceramics have become popular as tenacious materials to overcome the brittleness of ceramics. These composite ceramics, however, do not sinter well and hot pressing is used instead. As a result, it is difficult to fabricate products with complicated shapes. The objective of this study is to increase the density of Al_2O_3-SiC whisker composite ceramics by a conventional sintering method, thus enabling one to fabricate products into practical shapes.

2. Method: From preliminary experiments, we found that SiC whiskers form a three-dimensional network in the Al_2O_3 matrix when a conventional shaping technique is used. The whiskers obstruct the uniform contraction of the Al_2O_3 during sintering and thus lead to the formation of many cavities. To obtain a high-density sintered product, it is important to uniformly disperse the SiC whiskers in the Al_2O_3 matrix and also to orient them within a given plane.

For this purpose, we have tried various basic polyelectrolytes as the dispersion agents for both Al_2O_3 powder and SiC whiskers. These polyelectrolytes form macromolecular ions when dissolved in water. The macromolecular ions then cover the surfaces of the SiC whiskers and make the surfaces charged. When raw materials are subjected to wet pressing, whiskers repel each other, disperse themselves uniformly in the Al_2O_3 powder, and strongly orient themselves perpendicular to the press.

The polyelectrolytes we have tried as dispersion agents are polyethylene imine (PEI), polyvinyl pyrrolidone (PVP), polyacrylic amide (PAA), and Darvan (registered trade mark) C. For comparison, we also measured the dispersion properties of polystyrene-poly(methyl methacrylic acid) (PS-PMMA).

3. Results: The dispersion properties of various dispersion agents acting on SiC whiskers manufactured by ARCO Metal are shown in Fig. 1. Using the data in Fig. 1 and the viscosity of the solutions, we found that the water solution of PEI worked particularly well as a dispersion agent for SiC whiskers in

Al_2O_3 powder.

We present in Fig. 2 the density of a product when it was sintered for 20 minutes at 1710°C in an Ar atmosphere after the sample was mixed with a 4 weight percent PEI water solution and wet pressed. Also, when we sintered a sample that contained 15 weight percent of SiC whiskers at 1725°C, we achieved a 95.5 percent density in the sintered product. Through observations of the sintered products using an SEM and the anisotropy of the rate of contraction during sintering, we found that the SiC whiskers were dispersed uniformly in the Al_2O_3 matrix and were strongly oriented perpendicular to the press direction.

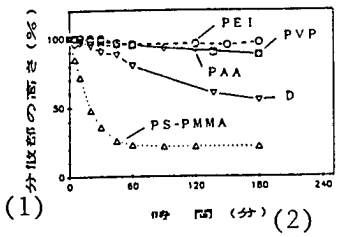


Fig. 1. Dispersion properties of whiskers when various dispersion agents are used.
Key:
1. Height of the dispersed part
2. Time (minute)

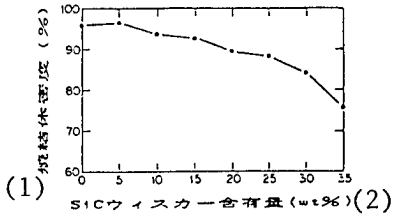


Fig. 2. Relationship between SiC content the sintered product density.
Key:
1. Density of sintered products
2. SiC whisker content

Mechanical Properties of Hot Pressed Alumina-SiAlON Composites

43067068h Tokyo 1988 FALL CONVENTION OF JAPAN SOCIETY OF POWDER AND POWDER METALLURGY in Japanese 9-11 Nov 88 p 195

[Authors: Kazumasa Takatori and Osami Kamigaito, Toyoda Research Center]

[Text] Objective: To improve alumina-based ceramics, we fabricated alumina-SiAlON composites and evaluated their thermal and mechanical properties.

Method: Al_2O_3, Si_3N_4, and AlN were mixed in the amounts listed in Table 1 using a wet ball mill, dried, ground, and used as the starting raw materials. The mixing ratio of Si_3N_4 and AlN was chosen so that, in theory, beta-SiAlON ($Si_{6-z}Al_zO_zN_{8-z}$) with $z=2$ was synthesized. The calculated amount of the synthesized product is also shown in Table 1. Hot pressing was performed in a nitrogen atmosphere under 20 MPa pressure at 1550--1700°C for one hour. The property evaluation was carried out by identifying the crystalline phase and observing the microstructure. Also, 4-point bending strength was measured at room temperature, 1200°C, and 1400°C, and the Young's modulus and K_{IC} were measured at room temperature. An oxidation test was conducted at 1400°C for 100 hours.

Table 1. Raw material composition.

Sample No.	Composition (weight percent) Al_2O_3	Si_3N_4	AlN	Synthesized SiAlON (weight percent)
A100	100	0	0	0
NA99	99	0.87	0.13	1.3
NA95	95	4.37	0.63	6.6
NA90	90	8.73	1.27	13.2
NA80	80	17.5	2.54	26.3
NA60	60	34.9	5.09	52.7
NA40	40	52.4	7.64	79.0

Results: (1) Dense sintered products of alumina and SiAlON were obtained by hot pressing. The particle growth of alumina was severely constrained due to the presence of SiAlON.

(2) The sintering mechanism is considered to be liquid-phase sintering. The microstructure of the product indicated that the SiAlON formed a net-like structure as the particle boundary phase for the alumina particles.

(3). The Young's modulus decreased as the amount of SiAlON was increased, indicating that an additive property was valid.

(4) The bending strength at room temperature reached a maximum when

the composition was about 30 percent SiAlON; this maximum strength is twice that of

alumina alone. Bending strength at high temperatures also

increased noticeably compared to that for alumina alone.

(5) The K_{IC} of the composites reached 4.5 $MPam^{1/2}$, which is 30 percent higher than that of alumina alone.

(6) The composites formed mullite on their surfaces from oxidation. The increase in oxidation grew as the amount of SiAlON was increased.

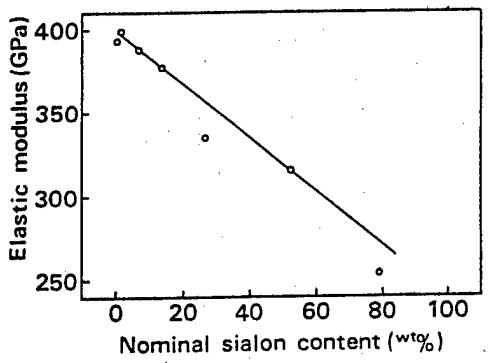


Fig. 1. Dependence of Young's modulus on composition ratio.

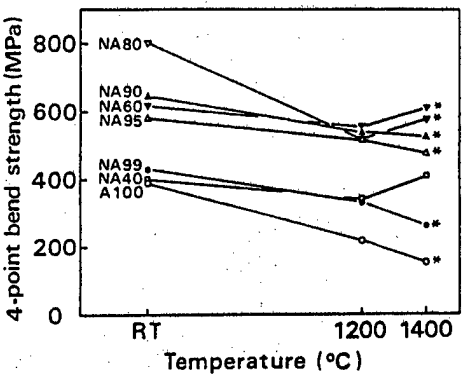


Fig. 2. Temperature dependence of bending strength.

AI Technology in Manufacturing Facilities and Equipment

Domain Shell for Diagnosis in Chemical Plant

43064091a Tokyo 1988 JOINT CONVENTION RECORD OF INSTITUTE OF ELECTRONICS AND INFORMATION ENGINEERS OF JAPAN in Japanese 1988 pp 1-155 - 1-158

[Article by Yuzo Malsuko, Toyo Information Systems Co., Ltd.]

[Text] 1. Introduction

An expert system (ES) is classified as either synthetic or analytic depending on the type of task it seeks to solve. Diagnostic ES's are famous examples of the later type.

Numerous diagnostic systems have been developed to date, including many which are now commercially operational. Among these systems are those developed through the use of our company's system design support tools called the Brain and Super-brain.

Evaluation of these programs reveal that software routines such as data input and user interface functions, common to all ES programs, are coded independently by many of the ES developers. Obviously this is unproductive, and wasteful since these efforts could have been directed to solving problems directly related to the ES tasks themselves. Thus, we, the tool developers, must respond by simplifying input methods and quickly incorporating these general functions into future releases of the development tool kits. In the United States, ES developers are seeing rapid developments in the tool market. Automated Reasoning Corp. has released the diagnostics tool "IN-ATE", Digital Equipment Corp. has developed an application shell for diagnostics to be announced at AAA-88, and Gensym Corp. now has a domain shell for real-time expert system G2-class diagnostic problems.

This article describes how an ES development tool (a domain shell for diagnostics in plant operation) may be used to design a computer based error diagnostic expert system which will assist in the operation of a chemical plant.

Besides the basic function of providing operation diagnostics, a full-fledged domain shell for plant diagnostics must be equipped with the capability of determining what to monitor and where to monitor, and the means of connecting to the physical plant.

2. Status of Development of Domain Shells

It was 1970 when the medical diagnostic system MYCIN was developed using the LISP language. MYCIN was followed by DENDRAL, a chemical compound determination system. Encouraged by the success of DENDRAL, systems developers started pouring substantial effort into ES development.

Advances in the design of general-purpose ES development tools in the 1980's led to a reduction in ES development costs which in turn led to furthering of the ES development effort. Many of that time foresaw that these ES development tools would be quickly followed by mature expert systems. Alas, the distance between the tools and real expert systems was too long of a jump for one step. The speed of advances in AI technology was slow and the level fell short of expectations. Few ES programs appeared that were commercially viable at all. This period of gestation however did help tool developers identify essential tasks common to many expert systems as development tools were actually applied in designing ES programs. It was found that for a given domain of ES tasks, an effective and generalized shell program, the domain shell, may always be designed.

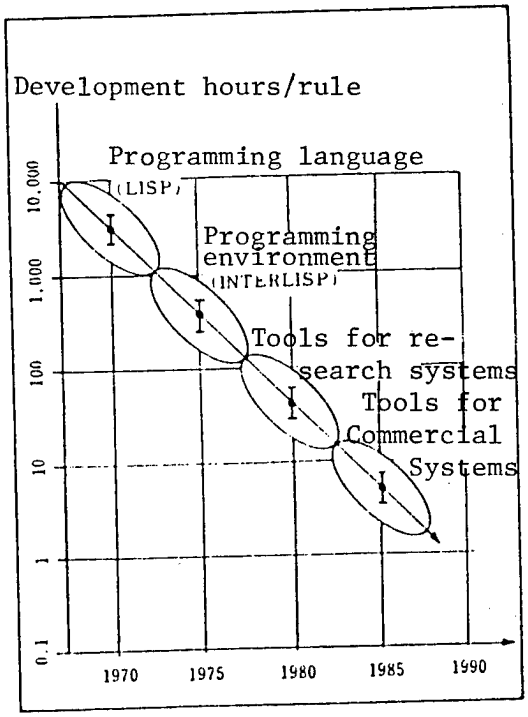


Figure 2-1 Reduction in cost for the development of AI systems [excerpted from: Expert Systems by P. Harmon and D. King]

Domain shell programs differ from one another in structure and inner workings due to the area of application of the expert system.

As for the domain shell for diagnosis of plant operation, our main topic for discussion here, the area of application is plant operation while the type of task is malfunction diagnosis.

3. Summary of the System

The system incorporates a knowledge base which consolidates all the knowledge regarding the system's error diagnostic tasks. Based on this knowledge, the system conducts monitoring of plant operation and diagnostics of malfunction states. The subject is a chemical processing plant dedicated to exclusive processing of gas or liquid, running under a normal condition. At the present stage of development, the system is linked to the plant off-line via file accessing.

3-1. Hardware Environment

Figure 3-1 shows a diagrammatic configuration of the basic hardware used

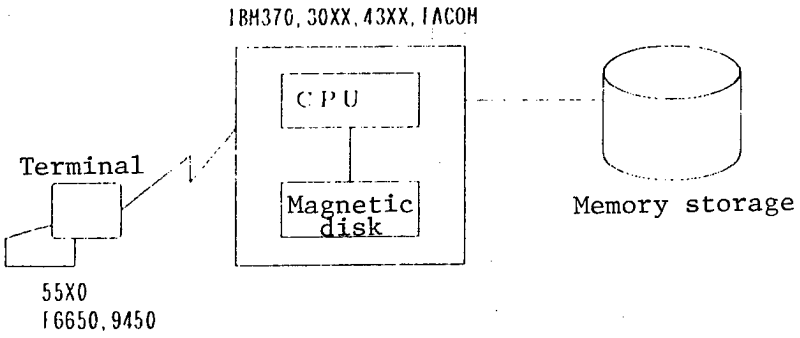


Figure 3-1 Basic hardware configuration

3.2. Software Environment

The basic software used are as follows:

OS: MVS (IBM)
OS IV/F4 (FACOM)
Language processor: UTILISP
Terminal software: 3270PC with Japanese language support,
DOS with Japanese language support (IBM)
Region: 4 Mbytes (minimum)

3.3. System Configuration

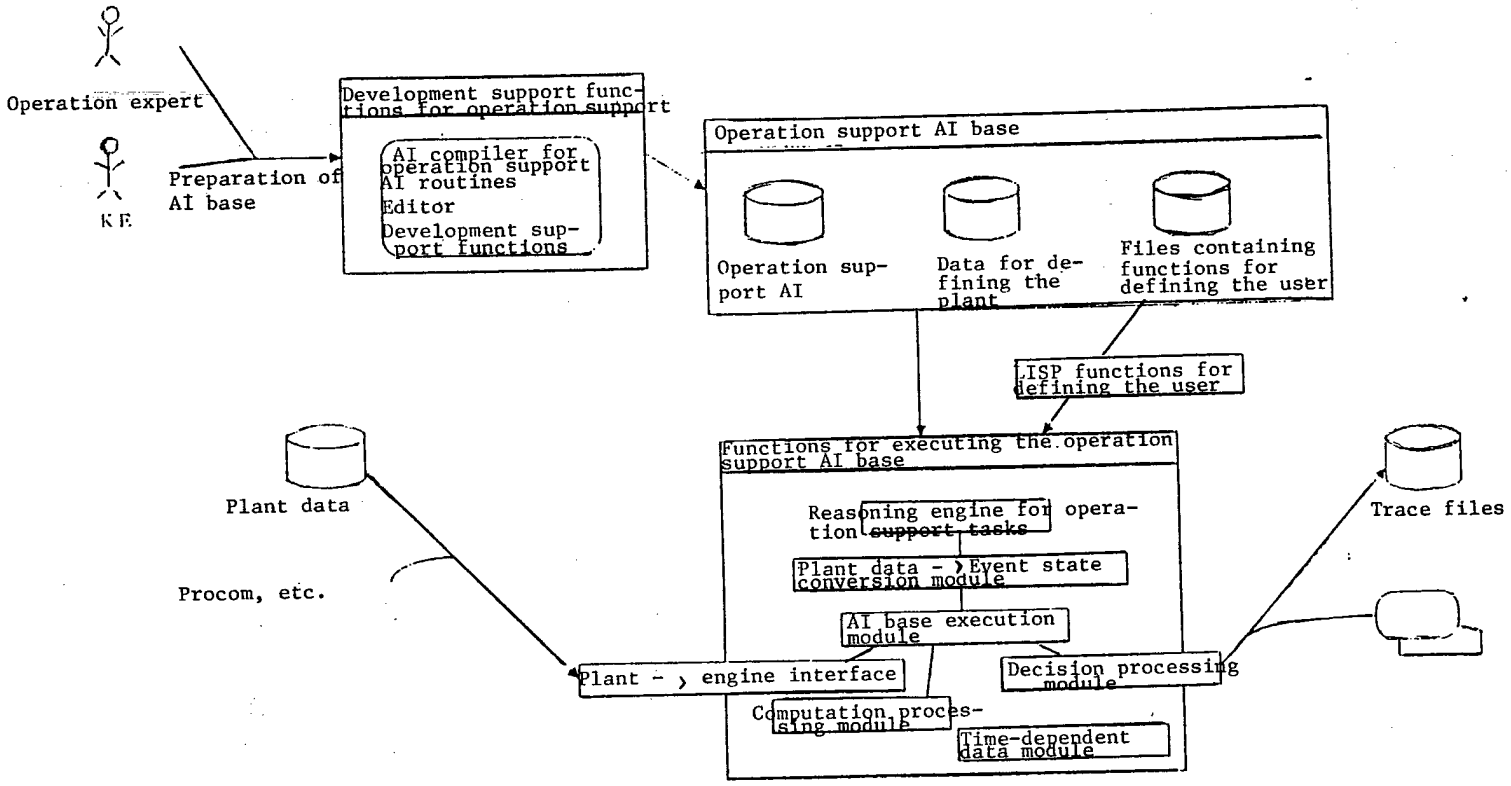


Figure 3-3 System organization

The system consists of system development kits (i.e., editor, compilor, plant data display, and data checking routines), a reasoning engine, and a knowledge base processing module, including computational processing routines.

The interconnections among these components are illustrated in Figure 3-3.

4. Summary of the Functions

Generally, an analytic expert system, designed for diagnostic purposes, performs the following basic tasks and problem solving functions:

	Basic Tasks	Problem Solving Functions
	1. Hierarchical classification of abnormal events	1. Event-driven decision making
D	2. Hierarchical expression of system structure	2. Target-driven decision-making
I A P	3. Inclusion of problem interpretation	3. Uncertainty-based decision-making
G R N O	4. Selection of monitoring points	4. Hypothesis-based decision-making
O B S L	5. Determination of cause of error	5. Emphatic reasoning
T E I M	6. [Determination of system] efficiency using rough model	6. Raising levels of efficiency and completeness
C S	7. [Determination of system's] completeness using detailed model	7. Analysis of cost effectiveness

(Citation from: Report of Research in AI Processing Systems by ICOT-JIPDEC AI Center)

In designing the new system in question, all of the abovementioned basic tasks except the coverage of the "completeness" issue were addressed. Development is aimed at design of a commercially viable system. Its problem solving functions are constructed around an event-driven decision-making algorithm. These routines are designed in such a way as to concurrently perform [the reasoning process] and the scheduling of input data selection and access cycles. This design consideration led to the creation of a system capable of efficient decision-making.

The following are the main system functions:

(1) Functions regarding editing of knowledge and data entry:

--Function for defining the plant model
--Function for translating plant data to event data
--Function for reading data through dialogue with user
--Tool for defining plant data

(2) Knowledge-based reasoning functions:

The plant may not be shut down just because there has been a small error. The new system is designed in such a way as to repeat the reasoning action based on the knowledge base [in order to insure that the imminent malfunction does not pose a sufficiently high risk]. The old Super-Brain did not have this capability.

(3) Functions common to domain shells:

--Time-dependent data processing functions
--Computational functions for accounting of parts and materials
--Plant emulator

Description of the procedure of knowledge-to-logic conversion and data entry follows next.

4.1. Editing of Knowledge and Data Entry

(1) Functions for Defining the Plant Model

The user may select the structure of the plant model freely without specifying the type of model in advance. In defining the plant, three levels of definition are used. First, the subject plant is defined as a whole, then each individual facility is defined (such as the distillation tower), and finally each instrument is defined (for instance, the flow meter).

The new domain shell adds a tool for defining plant data in order to facilitate knowledge input. This is our next topic for discussion.

Figure 4-1 illustrates the flowchart for the plant data definition displayed by two programs. The preparation of plant data begins by selecting either the option for a new definition or the LOAD option. The content of the input data is selected by either pressing the PF8 key (a function key on the 550) or the PF3 key. The data entered through the selection screen is output to a file by selecting either the SAVE option or the knowledge base format. As an example, the operation of the device definition screen and the conversion of the input data into the knowledge base format are discussed.

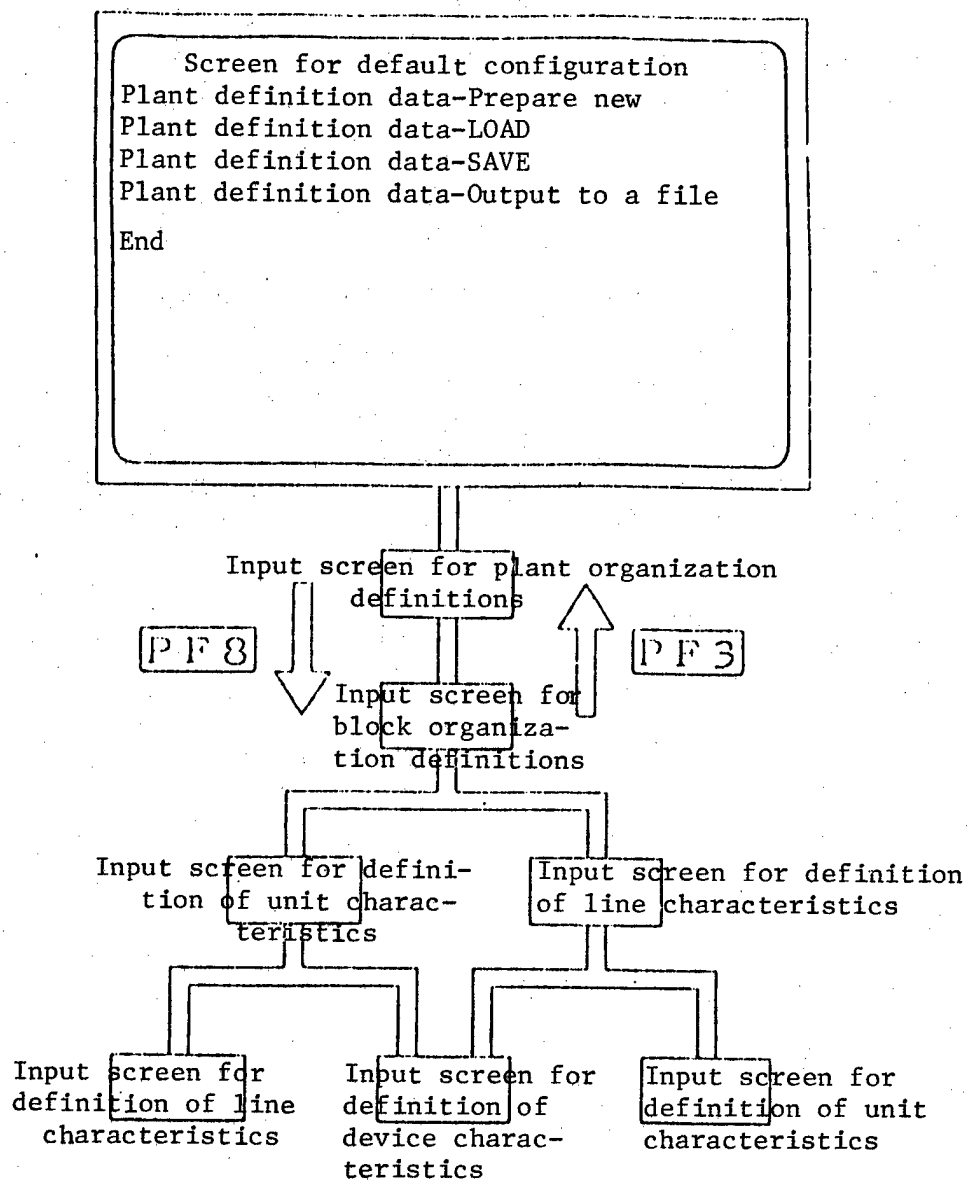


Figure 4-1 Flowchart for the plant data definition, displayed by the tool program

Initialization Screen and Its Functions

In the screen for initialization of the plant data definition, the user may select from the options of New Preparation of Plant Definition Data, LOAD, SAVE, File Output, and End.

1 New Preparation of Plant Definition Data:

This option is selected when new plant definition data needs to be prepared. First, the abbreviated plant name of the new data is entered in the input field. Next, "1" is entered in the numeric input field, which is followed by pressing of the execution key. The display switches from the initialization screen to the input screen for the plant organization definition.

2 Plant Definition Data LOAD:

This option is selected when a previously prepared file of plant definition data, retained by the SAVE option, needs to be retrieved. First, the file name is entered in the input field. Next, "2" is entered in the input field. Next, "2" is entered in the numeric input field, followed by pressing of the execution key. As soon as the file is loaded into memory, the screen switches to that file for input of plant configuration definitions.

3 Plant Definition Data SAVE:

This option is selected when the currently prepared plant data needs to be saved. First, the file name is entered in the input field. Next,"3" is entered in the numeric input field, followed by pressing of the execution key. The program saves the data just prepared in a file with the specified file name.

4 Plant Definition Data File Output:

This option is selected when outputting the definition data as a source file is desired. First, the name of the source file is entered in the input field. Next,"4" is entered in the numeric input field, followed by pressing of the execution key. The program saves the prepared data in a file with the file name given, in the knowledge base format of the domain shell.

E End:

This option is selected to terminate the session for plant definition data preparation. "E" is entered in the numeric input field, followed by pressing of the execution key. The plant data definition tool program terminates the data definition session.

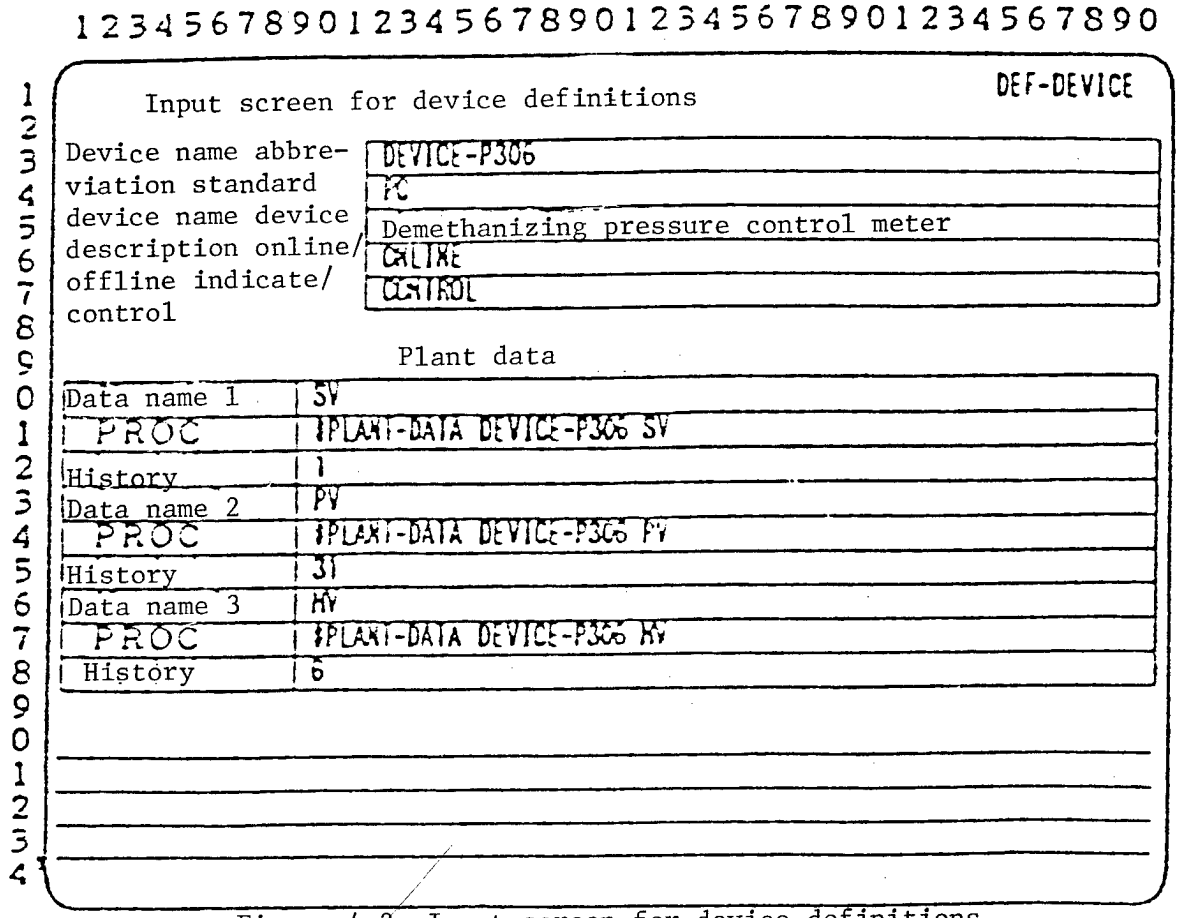


Figure 4-2 Input screen for device definitions

Device Definition Screen and Its Functions

The screen where device definitions are input [See Figure 4-2] is used to specify the type and attributes of a device which is used in the subject plant. To display this screen, the cursor is first brought to either the input field for the abbreviated device name, which is located in the input screen for unit definitions, if the device belongs to the unit, or the input field for abbreviated device names, located in the input screen for line definitions, if the device belongs to a line. This is followed by pressing the PF8 key.

When the screen in question is displayed, the abbreviated device name DEVICE-P306 appears in the input field for the abbreviated device name as this particular device happens to always be defined.

The user may enter, in the input field for plant data, the definitions for the plant data which is to be read during the decision-making session.

In this particular example, the screen displays DEVICE-P306. This device is defined as a PC (standard device name) to be used as an on-line control device for the demethanizing pressure control meter. According to the screen description, DEVICE-P306 is accompanied by 3 plant data items, SV, PV and MV each of which describes a separate plant definition.

There is no screen that follows the device definition input screen. By pressing PF3, the previous screen may be recalled.

```
(DEF-DEVICE
    DEVICE-P306
Demethanizing pressure control meter
    (DEVICE-TYPE  PC)
    (PLANT-DATA
        (SV (PROC (#PLANT-DATA DEVICE-P306 SV)) (HISTORY 1))
        (PV (PROC (#PLANT-DATA DEVICE-P306 PV)) (HISTORY 31))
        (MV (PROC (#PLANT-DATA DEVICE-P306 MV)) (HISTORY 6)))
    (ONLINE/OFFLINE  ONLINE)
    (INDICATE/CONTROL  CONTROL))
```

Figure 4-3 Input data converted to an AI base format

5. Conclusion

At Toyo Information Systems, in addition to the domain shell for plant operation diagnosis, we are developing a scheduler-type domain shell. We believe that widespread commercial use of the domain shell is certain to come, but in order to make it a success, we need to resolve a couple of problems. First, more experience needs to be gained in identifying generalized algorithms in programs for a particular domain or ES task. Secondly, the types of domains and ES tasks need to be restricted to a workable few for which a sizable market exists. Although an infinite number of ES tasks are possible (i.e., design, control, problem interpretation, and problem solving), there are a very limited number of domain shells that will be commercially marketable.

The development [of the new domain shell] was conducted on a multi-client basis. Some of our clients asked for the added capability of diagnosis based on the model's kernel (deep knowledge). We could not incorporate this in the domain shell discussed, but we are hoping to address this issue in future development opportunities. The issue of

reasoning based on deep knowledge has been recognized as a very important subject in the field of expert systems. By building deep knowledge inside a domain shell program, we can raise its level from a mere tool to the next-generation ES shell. It is a challenging problem, and we are making a commitment to addressing this issue up front.

Expert System for Diagnosis of Defective Electronic Devices

43064091b Tokyo 1988 JOINT CONVENTION RECORD OF INSTITUTE OF ELECTRONICS AND INFORMATION ENGINEERS OF JAPAN in Japanese 1988 pp 1-159 - 1-162

[Article by Tomoco Kumamaru and Kenji Adate, Toshiba Corp.]

[Text] 1. Introduction

While electronic components are used as the building blocks for a much larger system, they themselves come in small packages. The actual internal circuits have even smaller dimensions, beyond the human eye's capability to distinguish. Furthermore, extra precision is required in the task of diagnosis if such a device breaks down. This requires detailed work using diagnostic instruments, backed up by expert knowledge regarding the physics of the malfunctioning state as well as the diagnostic technique for identifying the cause of the malfunction. Today, the engineering of artificial intelligence is being actively sought in numerous fields. Expert knowledge is built into computers, and provides such machines as tools for supplying right information to people who are in need of such expert knowledge.

At Toshiba, we considered the idea of applying artificial intelligence to the field of electronic components, and through the use of an in-house developed expert system development tool, TDES3 (a Tool for Developing Industrial Expert Systems), we have created an expert system for diagnosis of malfunctions in electronic devices. Below, we would like to present a summary of the development system focusing on the indispensable pre-development tasks of reorganizing and translating knowledge into a programming format.

2. Organization of Expert Knowledge of Diagnosis

In general, human knowledge regarding a specific even is more than the information directly related to the event. However, all of this knowledge cannot be squeezed into the expert system. If this were tried, it would lead to an uncontrollable knowledge explosion. To avoid this, we considered organizing the knowledge. We identified "what domain of knowledge goes into the expert system". For instance, regarding the subject of error diagnosis, to be discussed below, we extract and put into the system that part of the knowledge that's directly relevant to the diagnosis.

Incidentally, broadly speaking, a diagnostics job covers more than the diagnostic task itself. It also encompasses pre-incident evaluation tasks such as forecasts and predictions, which are constantly being conducted even when no malfunction has occurred. As for the diagnostic task itself, it may be an analysis of a specific device (exclusive diagnosis of the malfunctioning unit or a statistical analysis for a mass production product.

As for the system currently under discussion, we limited the domain of ES knowledge to that for analysis of a specific device. Furthermore, we divided the knowledge into two classes: knowledge of part malfunctions, and knowledge of the diagnostic tasks. The former was organized as knowledge for defining diagnostic rules, and the latter as a guide for prescribing physical diagnostic steps to be taken by the operator.

2.1 Organization of Knowledge of Malfunctions

Knowledge of the malfunction concerns explanation of the physical state underlying the malfunction event, which relates the types of stress that might damage the device to a particular malfunction state (failure mode in the component. Usually, confirmation of the failure mode doesn't mean immediate identification of the stress which caused the malfunction. The reason is that a failure mode only describes a rather simplistic phenomenon, such as a short circuit or an open one, and does not provide sufficient knowledge to investigate causes of the stress. This is why we need to map the diagnostic knowledge to sections of the device that can be affected by certain failure modes. The map is then filled in with instances of damage and possible causal stresses. These factors are then correlated for causality and sorted based on rules of physics and experience.

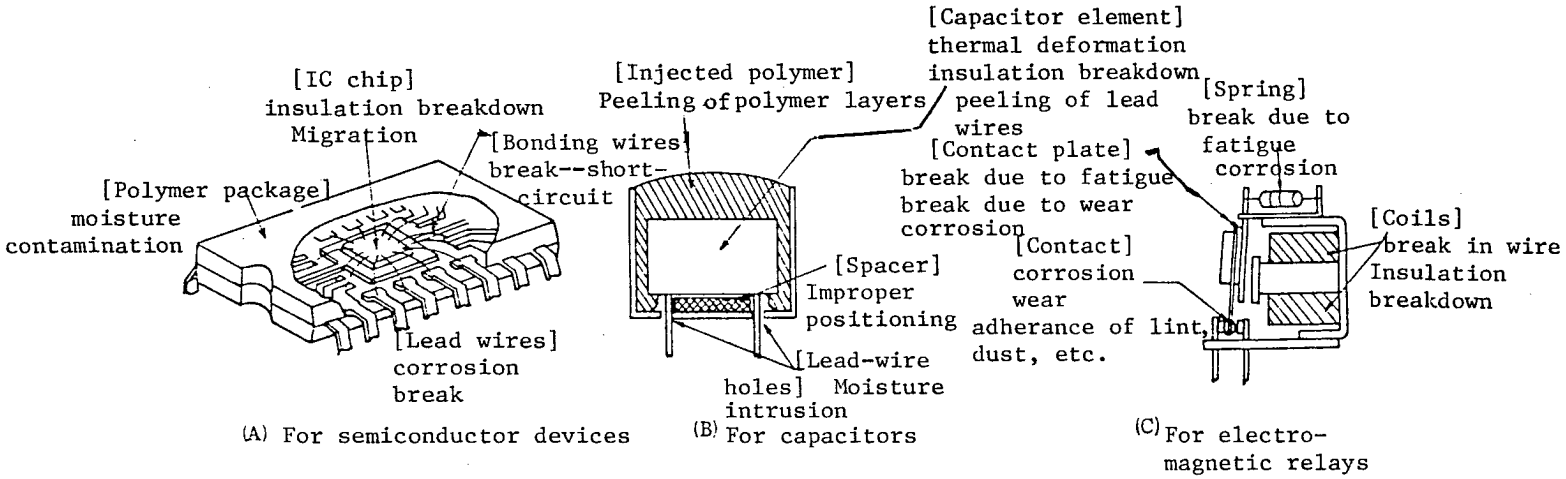


Figure 1 shows examples of possible causes of malfunctions in some electronic components. Each device has unique sections associated with different types of malfunctions, and the degree of possible sustained damage depends on the type of applied stress. The diagnostic task then is to monitor this malfunction condition. In order to investigate the cause of the failure, we also need knowledge of the device's production process as well as various sources of stress which arise due to the device's usage and the environment it's used in.

2.2 Organizing Knowledge for Diagnostic Tasks

The knowledge needed for diagnostic tasks regards the procedure for diagnosing the handling diagnostic instruments. According to general diagnostic procedures, the malfunction is located by restoring the item through an unpackaging technique, the state of the malfunction is scrutinized in detail using an observation technique, and, if any foreign matter is found at the location of the malfunction, the intruding substance is identified using a testing procedure. There are a number of commercially available diagnostic instruments for this type of failure analysis. The knowledge needed to use such an instrument is given in the instrument's manual. What's needed then is to incorporate this type of general knowledge into the actual diagnostic procedure, and organize it as specific piece of working knowledge which details "how to unpackage the device, what section to study, and how to study it."

Let's suppose that the elected diagnostic job is suitable, its proposed content is consistent with the working knowledge, and that the cause of the malfunction has been identified. Figure 2 is an electron microgram of the cross section of a capacitor, shown in Figure 1 (B), cut using a precision low-speed cutting device to reveal the location of the spacer section. The failure mode was assumed to be an insulation failure between the electrodes. The possibility of a short circuit in the component was considered small. With these assumptions made, we investigated the polymer insulator for signs of failure. The result was this: as expected, the polymer material had a pore which was caused by moisture intrusion.

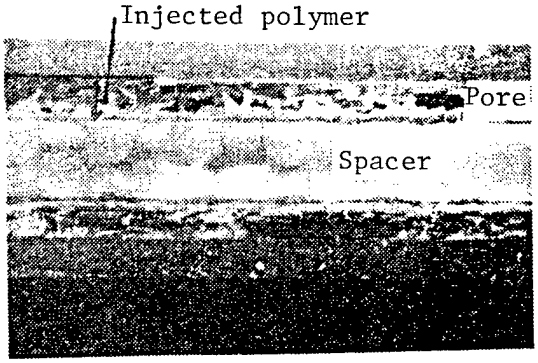


Figure 2 Electron microgram of a defective section

3. Knowledge Expressed by TDES3

Ways of expressing knowledge in a format that does not depend on the type of computer in use and allows for more versatility of expression are being considered. TDES3 relies on a syntax set unique to AI processing to express knowledge. It comprises schemata, facts, rule modules, and procedures. The following shows how the syntax is employed:

(1) Schemata

A schema is a structured set of facts. A schema may have multiple slots which are used to classify the object represented. Facts referring to multiple instances of classification may be handled in units of schemata. With TDES3, schemata are expressed in key words describing the types of devices, diagnostic conditions, malfunction states, etc. Figures 3 illustrates an example of a schema. In the figure, the term "is-a" indicates that the slot is passed from the semiconductor device to the diagnostic condition.

```
(schema   Semiconductor Device)

     (Type of Device       Initial Stage)       ;Slot
     (Failure Mode         Initial Stage)       ;Slot

(schema  Diagnostic Condition)

     (is-a Semiconductor Device                 ;Passing of slot
     (Diagnosis            Initial Stage)       ;Slot
```

Figure 3 Examples of schemata

(2) Facts

A fact describes an established event or condition. Since there were no established facts known at the initialization of the new system, all events and states were expressed in terms of schema.

(3) Rule Modules

A rule module is a meaningful set of rules. It is used to raise the level of reasoning efficiency by limiting the execution of the reasoning action within the selected module. In the new system, groups of rules regarding failure events and causes were organized into rule modules.

(4) Rules

In contrast to facts and schemata which are used to describe facts, rules are used to identify the cause and effect relationship between these facts. Figure 4 shows use of a rule. As for the format for coding a rule, we adopted a method of first placing the schema in a conditional section, testing the schema's slot value for a match with the content of working memory, and writing out a specific execution procedure selected in accordance with the result of the conditional test. The figure depicts the definition of a rule for a case where the slot values of the schemata for the semiconductor device and for failure conditions are identified by means of diagnostic work, and for a case where there is the possibility that the cause will change, depending on the environmental state.

```
(rule     Cause 7
          (schema   Semiconductor device
                    (Type of device:    Transistor)
                    (Failure Mode:      Short circuit)
          (schema:  Failure State       (State:    Contamination in chip))
          (schema:  Environmental State (Environment:  7x1)

=>

          (if  (==?x1  High humidity environment)  then
               write t  "The cause for the malfunction is `a contamination
                         of the chip by ionic species.'  They were
                         generated by an intrusion of moisture.")

          else
               (write t "The cause for the malfunction is a `contamination
                         during production processing at the
                         manufacturing plant.")

           ))
```

Figure 4 An example of a rule

(5) Procedures

A procedure is used when and where efficiency can be increased by resorting to the procedural language. For the new system, we used them for the display of construction drawings of the electronic devices. The procedures were incorporated in such a way as to enable simultaneous processing of the reasoning and the display by making use of a scheme of multi-task calls.

4. Functions of Expert System and Its Operation

4.1 Organization of System's Functions

Toshiba's workstation AS3000 is used to run the system. By adding multi-screen and multi-processing capabilities into the computer design, its functions were made more versatile, and through use of a menu-driven user interface, its operation was made simpler. Figure 5 illustrates how the functions of the system are organized. The left half of the diagram refers to the relationship between the function blocks of the reasoning engine of TDES3. The engine performs the following tasks: (1) Cross references are made between all facts and conditionals specified by the rules, (2) based on results of the comparison, one of the competing pairs from the set of facts and conditionals is selected, and (3) the rule that this pair is associated with is used to test the failure condition. The engine repeats this cycle until the decision-making process comes to a dead end or there is nothing left in the set. The right side of the diagram describes the function blocks of the diagnostic program. Rule modules are defined in terms of functions. As for its system-level functions, there are two: diagnostic support, and support for investigation into the cause of the malfunction.

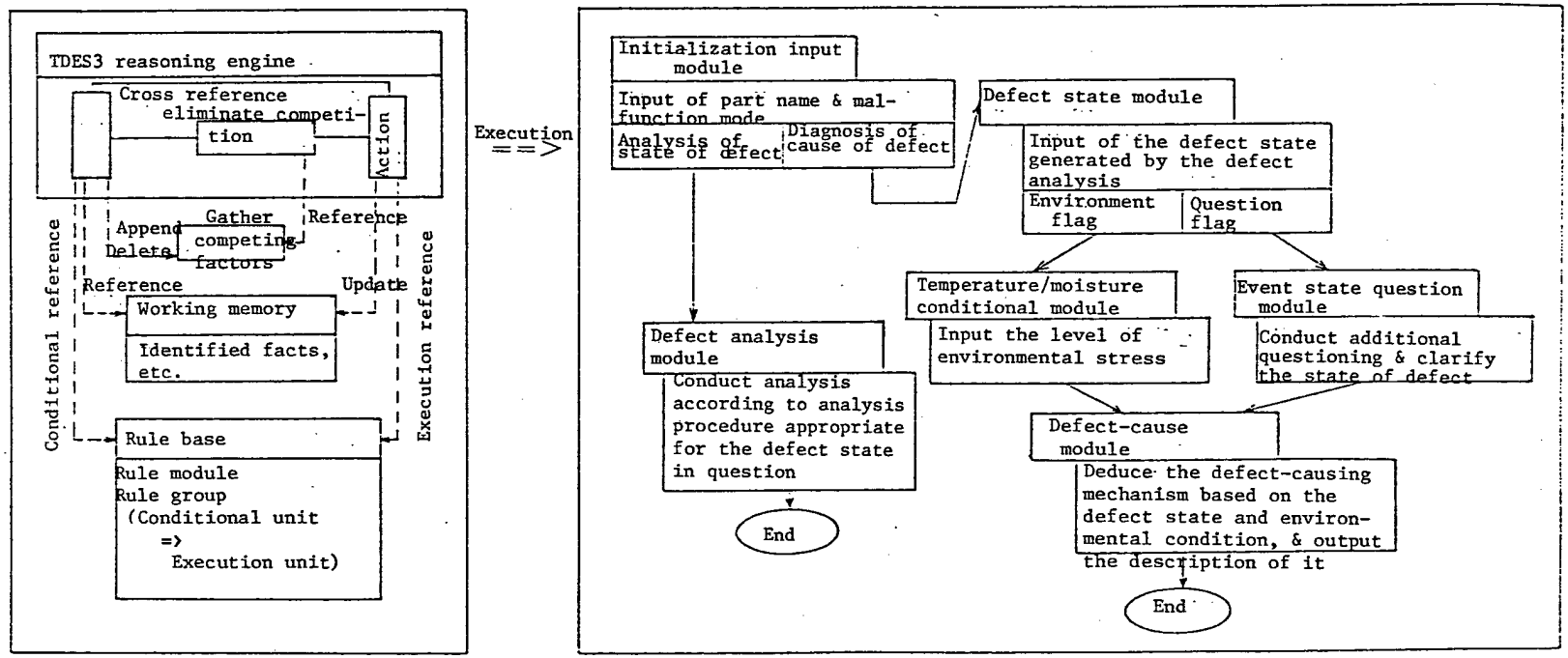


Figure 5 Organization of expert system functions

4.2 Example of the System's Operation

As an example to illustrate how the system operates, we selected a diode, which is an electronic device, and considered a case of open circuit failure. A scenario of the action taken by the system is depicted in Figure 6. To assist in the diagnostic effort and hence to improve the crew's work efficiency, the system provides guidance regarding the procedure for possible on-site inspection that may produce diagnostic evidence and provides instructions for remedies. As for support during investigation of the cause, the system generates messages on the screen concerning the cause of the malfunction and the mechanism, which is deduced from the observed evidence and circumstances which led to the malfunction.

(figure 6 on following page)

5. Analysis

Generally, diagnosis encompasses a repetition of the reasoning cycle, which is needed to trace the direct cause, and even a repetition of the circumstantial factors that induced it, from what's available in the malfunctioning device. If the failure cannot be located or if the circumstances are now known, the cause may never be found. Like the procedural setup shown in Figure 6, before a technician can select a diagnostic procedure, he has to carefully check each of the possible event descriptions that fit the actual malfunction. Equipped with vast knowledge about the physics of malfunctions and accumulated experience in dealing with malfunctions, an expert system can be a very powerful tool for determining the right diagnostic procedure in a short amount of time. On the other hand, there are instances where an expert system may not work at all. Many times at a working plant, the cycle of the occurrence of a malfunction in a device or an instrument, and the investigation into the cause, followed by modification, prevents repeated occurrence of the same malfunction, throwing out the chance of ever coming across an event for which previously acquired knowledge may become meaningful. In other words, we need to consider ways to handle cases where the real cause of the malfunction differs from those imagined by the experts. Although a general solution for this type of problem may never be found, to fill this possible void, we are trying, at least for electronic devices, what we call the good-unit diagnosis where a trouble free device is used as a reference in order to diagnose the defective one. Also, from time to time we combine this technique with a test for reproducing a malfunction by applying accelerated levels of stress to the subject device.

6. Conclusion

Often with a limited staff, the task of troubleshooting multiple types of parts becomes hard to plan and coordinate because of the constant need for adjustment on the part of the staff in order to respond to

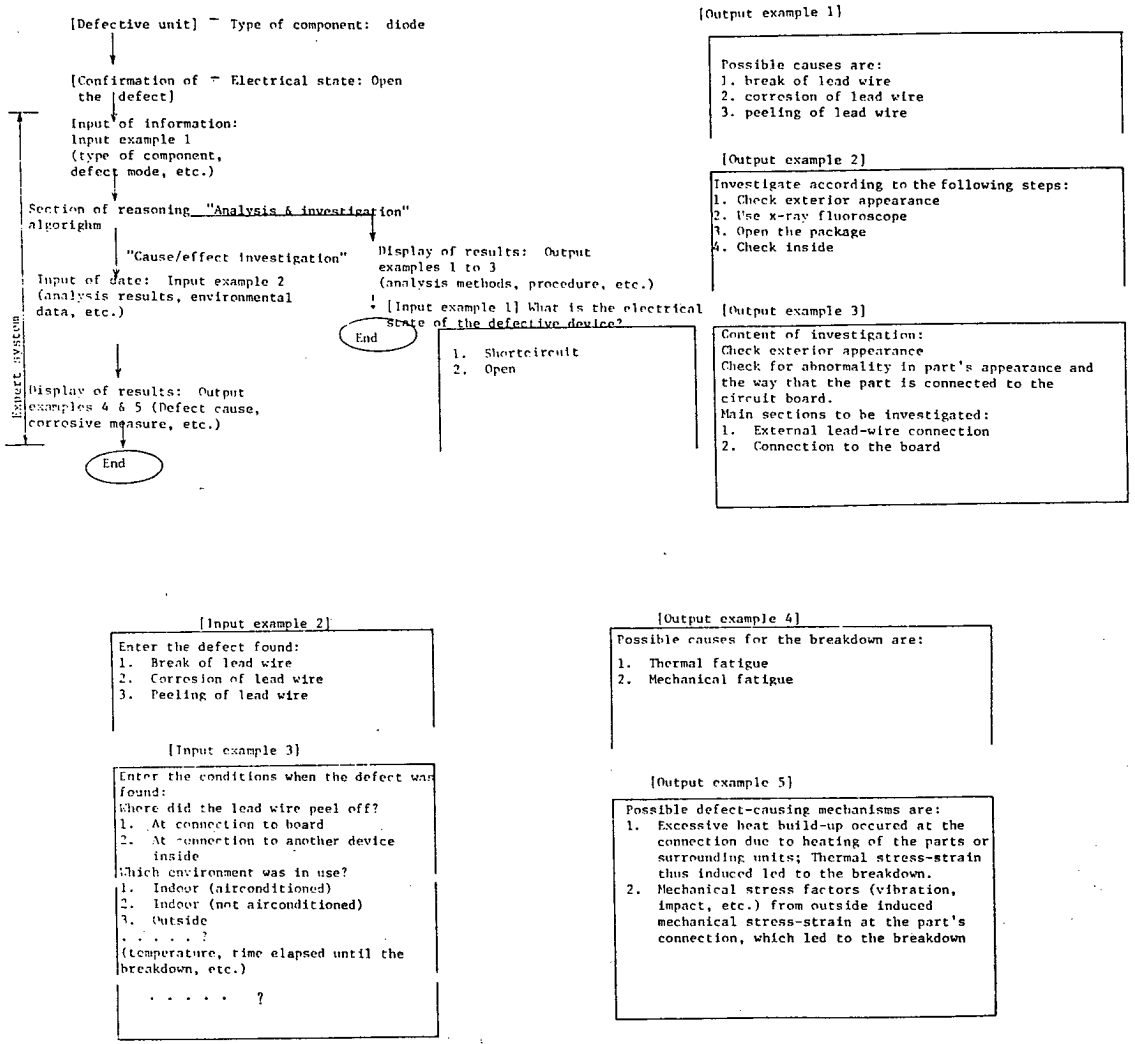


Figure 6 An example of the system in action; "Case of an analysis & investigation"

differences in diagnostic procedures and failure evaluation standards from one part to another and from one failure mode to another. The diagnostic system discussed above may be taken advantage of in reducing the work load of the troubleshooting crew who may otherwise have to rely on a time-consuming process of trial and error.

In the future, we are considering expanding the knowledge base and extending the coverage of subject devices. At the same time, we are hoping to utilize the knowledge base for improving the reliability of design of electronic instruments which are made of electronic devices tested by our ES program.

Expert System for Diagnosis of Malfunctions in Electric Machinery

43064091c Tokyo 1988 JOINT CONVENTION RECORD OF INSTITUTE OF ELECTRONICS AND INFORMATION ENGINEERS OF JAPAN in Japanese 1988 pp 1-163 - 1-166

[Article by Kinjiro Ito, Hiroyoshi Wada, Shinichi Fukumoto, Fuji Electric Engineering Co. Ltd.]

[Text] 1. Introduction

In recent years, the need for specialists with a high level of technical expertise is growing quickly due to an increase in the number of malfunctions occurring in aging electrical machinery, and as the trend for larger machines is picking up.

Recently, the number of incidences of malfunctions in electrical machinery is growing quickly due to excessive use of these machines, beyond their indicated lifespan, as well as to the industry-wide trend for design of larger electric machinery. The effect of such failures on the job being performed as well as on the defective machines themselves are complex requiring many experts, highly trained in the technology. To compound the problem, an alarming number of highly qualified professionals are aged and retiring. Also, the training of new personnel is not easy. These circumstances are prompting development of expert systems which possess expert knowledge and can substitute for retiring personnel. The situation surrounding the issue of the diagnosis of electrical machinery is no exception. In this report, we will discuss recent expert systems we constructed for diagnostic purposes for primary electric devices such as power transformers, rotating machinery, and electronic computers. The development utilized a general purpose ES development tool.

2. Diagnosis of Equipment and the Expert System

As far as machines are concerned, there is always a definite cause for a malfunction whose symptoms we can monitor. Furthermore, the cause and the observed event must be connected via a causality relation which is deterministic and logically defined.

On the other hand, although the experience-based knowledge that experts have is often ambiguous and complicated, it is still a fact that trained workers do correctly identify the cause of malfunctions by basing their judgment on the observed indications. So, through systematic study of the subject of the diagnosis, we can construct models of effective solutions, conceived through the interaction of experience and theory. As for methods of failure diagnosis, there are several options: diagnostic techniques based on system characteristics and internal structure; techniques based on pattern identification as in signal processing; and techniques that look at the cause and effect relation between machine units and parts or between the associated logical elements. However, most of the available techniques are based on the approach of causality identification.

Generally speaking, when used in a big plant environment, a diagnostic ES is often constructed from a custom-made ES development tool by special knowledge engineers. But, should the projected size of the system be relatively small, a development kit which is readily available off the shelf, may as well be used, and the system can be developed by the experts themselves in order to bring down the costs of the ES development.

3. Summary of the ES Development Tool

We used the tool program called COMEX (Compact Knowledge Based Expert System) developed by Professor Haruki Ueno of Tokyo Denki University. It consists of a knowledge editor, a reasoning engine, a knowledge base, and a data base. Besides these main units, it has various service routines available to assist in the coding of a complex ES program. The hardware for running the system was a personal computer with an MS-DOS system. The final program can be executed without being aware of the language used for development, which is FORTRAN.

The knowledge editor has dialogue-driven functions. They enable preparation, display, and revision of procedural guides which store knowledge collected by experts. As for the reasoning engine, its functions cover: inquiries by the user; access of the data base; retrieval of responses from the user; reasoning based on the response; display of the result of the reasoning process with message (underscored for better visibility, if highly probable); display of advice consistent with the result reached by the reasoning process; explanation of the reasoning process; and management of the data base. COMEX's other features focus on compatibility with the concept of a multi-layer knowledge base. One of these features is the ability to call the reasoning engine as a subroutine, and another is the ability to use the add-on programs supplied for sharing of data. In addition, by making use of an appropriate interface, the system can retrieve external data on demand and pass reasoning results to an off-line system.

As for the knowledge structure, at the very bottom layer, the observable events represented by numerical data (facts) and non-numerical data (characteristics) are input. At the next higher level, there exists a layer of intermediate hypothesis which is made up of intermediary facts expressed by logic elements like AND and OR. The information at these two layers is used as a conditional item to be referenced by the criteria frame. In the criteria frame layer, each one of the hypotheses for conclusion is tested, resulting in acceptance or rejection of the logic. A criteria frame is a kind of knowledge model expressed in terms of a standard framework (comprised of judgment conditions, confirmation conditions, and exclusion conditions). The strategy is to test each hypothesis for failure to fit the 3-stage [conditions], and screen it out if it indeed does fail. With a sound hypothesis identified this way, a conclusion is reached.

4. Examples of Applications in the Diagnosis of Electrical Devices

Application of a diagnostic ES means coordinated monitoring of multiple evidential signs from a subject machine, which may be hard to do using just a single sensor device. It then has to identify the state of the machine correctly, and, if an error has occurred, estimate the cause of error in terms of location and extent of the malfunction. What this means is that since the diagnostic technique depends on a particular electrical machine, the actual knowledge representation and reasoning process need to be designed differently.

4.1 Diagnostic ES for Power Transformers

Figure 1 depicts a diagrammatic outline of the automated diagnostic expert system developed for diagnosis of transformers using an automatic in-oil gas analyzer and COMEX. This system is an on-line ES which can automatically diagnose abnormalities inside the transformer for the type of malfunction and its location by basing its decision-making action on data collected concerning the amount and type of hydrocarbon gases dissolved in the transformer insulation fluid, which is monitored on-line by the automatic analyzer. Also, since diagnosed data can be stored in the data base, effective management of the transformer's historical profile can be maintained which may help identify any cyclic variation in the machine's operating state.

If heat is generated inside an oil-filled transformer due to an insulation breakdown or localized heating, the part of the insulation fluid (or insulation paper) in and around the heated area may be thermally affected, inducing a chemical dissociation which in turn can cause release of a hydrocarbon gas. Gases thus generated dissolve in the insulation fluid, spreading throughout the inside of the transformer, carried by convection flows. Therefore, by monitoring the concentration and composition of the gases found in fluid samples extracted from the transformer, the performance of the transformer may be monitored to identify any abnormalities.

This type of diagnostic technique can be run without shutting down the transformer, and can still detect an impending abnormality at an early stage of development. Therefore, it has been adopted in many instances as the method of maintenance for oil-filled transformers.

The expert system developed was built around these pieces of diagnostic knowledge and conclusion models while its source of information for prescribing remedial measures for the troubleshooting crew was a knowledge base.

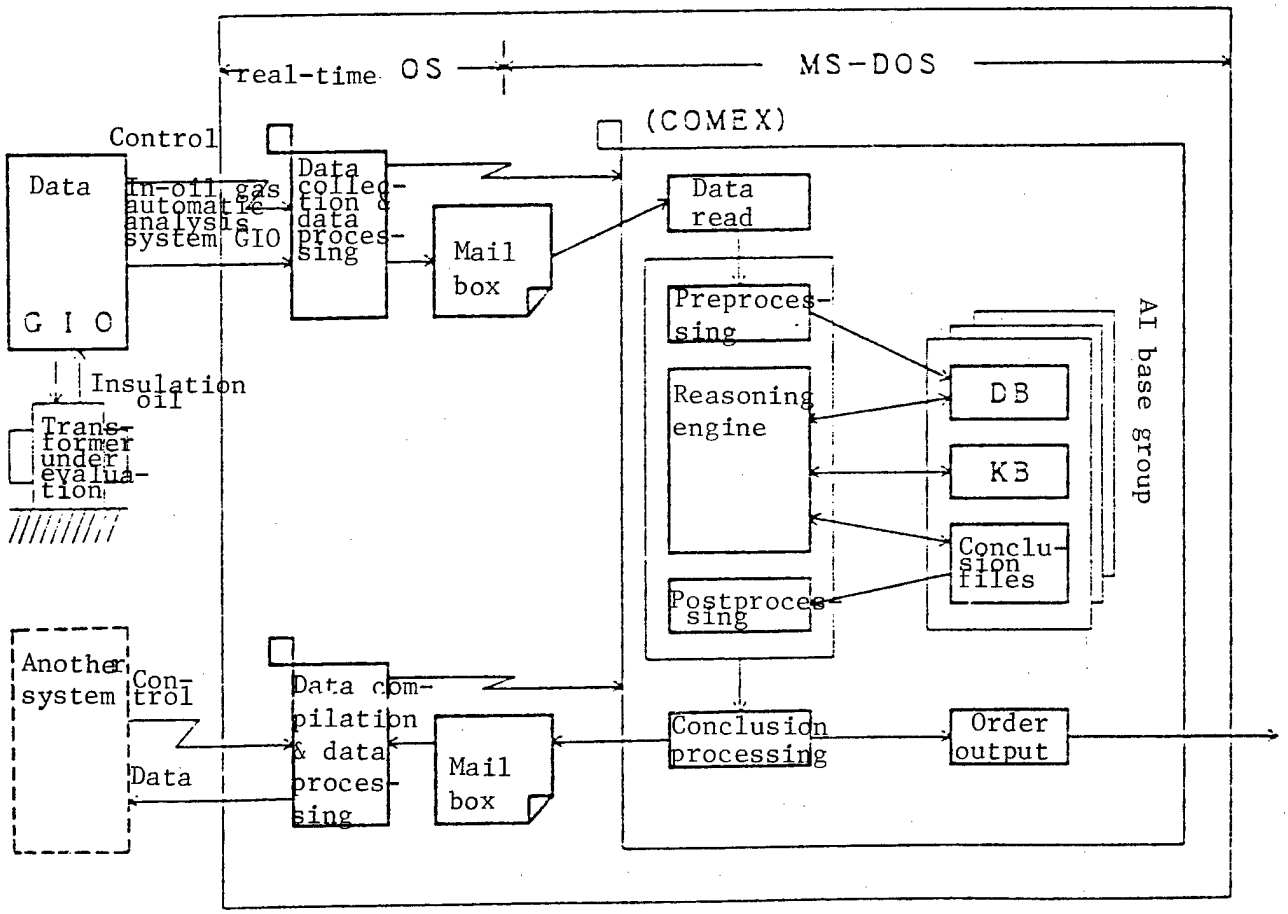


Figure 1 ES for automatic diagnostics of power transformers

The knowledge base was built by defining observed [conditions] (15 items), hypotheses for conclusion (26 items), and intermediary hypotheses (132 items). For each hypothesis for conclusion, a criteria frame was constructed. An example of conducting a reasoning session based on the constructed knowledge model is cited in Figure 2. The observed items correspond to the data from in-oil gas analysis and property tests done on the insulation oil. Except for the turbidity and four odor of the fluid, the figures shown indicate known values.

The conclusion of the diagnosis states that the gases found in the oil are at an abnormally high level, and that there is a good possibility for excessive heat build-up accompanied by burning of the solid insulator material. Furthermore, the system identifies the main culprit to be C_2H_2, and asserts the presence of an excessively heated section due to a bad electrical contact or current leak.

To evaluate the credibility of the knowledge models of the system, we compared the conclusions drawn by the system and those by experts for 945 real cases involving physical devices. According to the results, 84 percent of the diagnoses made for either warning or abnormal states were correct, and with those for normal states included, the figure rose to 98 percent.

Also, the system produced almost identical conclusions regarding the types, locations and extent of the abnormalities as those of the experts.

[figure 2 on following page]

4.2 Diagnostic ES for Insulators in High-Pressure Rotating Machinery

An insulation material in a high-pressure rotating machine can gradually lose insulation strength due to various stresses it encounters during the machine's operation. In many cases, the process is complex involving multiple types of stress, because it is rare for only a single type of stress to be entirely responsible for an eventual insulation breakdown

There are two types of information one may obtain to characterize the insulator material by running tests. One is numerical information on measurable quantities such as insulation resistance, dielectric dissipation factor, current growth rate, partial electrical discharge, etc. The other is information concerning patterns such as voltage curves from electric dissipation, and partial electrical discharges as well as electrical discharge waveforms.

Also, the mode of insulation degradation depends on the insulation system in a particular subject machine and on its design specifications. Therefore, any type of malfunction tends to be a complicated event caused by competing factors.

Skilled insulation specialists sort these pieces of information into two categories: one category for repairable damages due to reversible processes of moisture absorption and another category for irreversible processes due to aging. By drawing on this information as well as information from visual inspection of the material's appearance and the material's history profile, these experts can conduct comprehensive diagnosis of insulator materials for the location and extent of the defect.

NO. 4 Code Name : DEMO-CASE1 ********** 760825

Comment: Electrical Device Joint Research, Vol. 36, No. 1;
Subject No. 1; Specs: 33/6.6KV 10MVA

Specifications: Voltage 33.00 KV, Capacity 10.00 MVA

Date of analysis: 8-25-76 (7-18-76)

Gases found in the oil			Quantity from this measurement		Quantity from last measurement
Hydrogen	(ppm)	H2	2700	(	860)
Methane	(ppm)	CH4	3760	(	1670)
Carbon monoxide	(ppm)	CO	400	(	10)
Carbon dioxide	(ppm)	CO2	1800	(	690)
Ethylene	(ppm)	C2H4	4720	(	2050)
Ethane	(ppm)	C2H6	790	(	30)
Acetylene	(ppm)	C2H2	80	(	40)
Total amount of combustible gases	(ppm)	TCG	12150	(	4660)
Rate of increase in amount of combustible gases	(ppm/Year)	TCGV	76847	(	0)
Data for insulation oil characteristics					
Water content	(ppm)	WOT	24.60	(	0)
Dielectric breakdown voltage	(KV)	BDV	56.00	(	0)
Volumetric resistivity	(*E12 Ω cm)	VOR	35.00	(	0)
Total acid number	(1/100 mgKOH/g)	KOH	2.50	(	0)
Presence of turbidness in insulation oil		CONT	0	(	0)
Presence of foul odor from insulation oil		SUML	0	(	0)

***** Results of ***** diagnosis

1) GAS1.3 Evaluation of in-oil gas state ; <=300KV <=10MVA --> A

(A:Abnormal, B:Warning, C:Lack of Data)

2) OHT Overheating --> A

(A:High, B:Medium, C:Low)

3) INSI Burn of solid insulation material --> B

(B:Large probability, C:Small probability)

4) PTC2H4 C2H4 Main control unit: presence of heat build-up due to bad contact & leakage current --> C

(C:Exists)

5) OIL0 Evaluation of quality of insulation oil --> C

(C:Normal)

Figure 2 Actions taken by ES in transformer diagnostics

In contrast to human counterparts, our ES program was armed with expert knowledge stored in a knowledge base format. It evaluated the cause and extent of gradation at 3 levels: "determined", "suspected", and "possible".

Although the diagnosis of the cause of degradation, represented by such a 3-level evaluation, occupies an important position within the overall diagnostics of the insulator, it is also essential to provide an overall rating of the insulation system in terms of risk assessment, i.e., "danger", "warning".

This consideration led us to adopt, in the new system, a double knowledge base scheme for supporting the reasoning session. The first knowledge base contained information for drawing conclusions concerning degradation causes, and the second for assessment of the danger of system breakdown.

NO. 3 Code name : 24-K642-1003-540920-SS0041*
Comment : 250KW 6000V 200185K S39

***** Data *****

Insulation resistance (MΩ)	R	1400.00
Polatization index	PI	1.50
Kick current	KC	None
Dielectric dissipation factor (1KV) (%)	DTO	6.60
Rate of change in dielectric dissipation factor (%)	DT	3.00
Voltage characteristics of dielectric dissipation factor	TP	TIP
Rate of increase in AC current (%)	DI	4.60
Point of rapid increase in secondary current (KV)	PI2	4.90
Maximum electric discharge (PC)	QM	50000.00
Corona waveform	QR	NOT Random
Dust	DUS	Unknown
Tracking	TRC	Unknown

***** Conclusion *****

Cause of fatigue

1) Excessive aging is certain
2) Moisture contamination is suspected
3) Dirt contamination is possible

Possibility for breakdown: High overall

Figure 3 Actions taken by ES in insulator diagnostics

Figure 3 shows the summary chart for an example of a diagnosis conducted on a 6KV induction motor. According to the conclusions listed in the chart, the cause of degradation is almost certainly the aging of the material. Moisture absorption and fouling are also suspected. The overall assessment indicates that the insulator is in a critical state, requiring urgent measures to be administered.

We conducted evaluation of this prototype ES program by comparing its diagnoses with opinions by experts on 220 cases of insulation evaluation. According to the results, the average rate of success of the program diagnosing the degradation correctly is 79 percent (with 100 percent being the rating for a perfect match with the expert's evaluation). As for the overall risk assessment, the success rating was 72 percent, which was similar in magnitude to the former figure.

4.3 Diagnostic ES for Rotating Machinery

Although extensive research work has been conducted in diagnoses based on vibrational methods for rotating machinery, the research results have not been organized into an easy-to-follow systematic order. As the subject of the diagnosis--performed in terms of an evaluation of the detailed data obtained by diagnostic instruments, and in terms of the machine's malfunction history and the first-hand experience of experts--belongs to the area of skilled specialists, it is also a welcome subject for the design of an ES program.

The program we discuss here embodies diagnostic procedures similar to those followed by trained technicians in day-to-day routine work. The system is capable of two types of diagnostic procedures. These are the "5-sense/simplified diagnosis", and the "detailed diagnosis". They are depicted as flow diagrams in Figure 4.

Generally, technicians versed in rotating machinery diagnosis refer to matrix-type diagrams of factors which relate monitored events or diagnostic parameters with malfunction modes (possible conclusions). But, when it comes to actual estimation of the cause of the malfunction, such a diagram alone is not sufficient, and they take into consideration other factors such as the degree of contribution of each even to the suspected failure mode, the overall likelihood for that mode to occur given a certain combination of events, and the trustworthiness of the collected data. In our ES, the knowledge base was built around all these pieces of information from experts, and they were expressed in terms of criteria frame theory.

The knowledge base consists of hypotheses for conclusion (24 items), observable conditions (5-sense/simplified: 39 items; detailed: 88 items), and intermediary hypotheses (5-sense/simplified: 89 terms; detailed: 111 items).

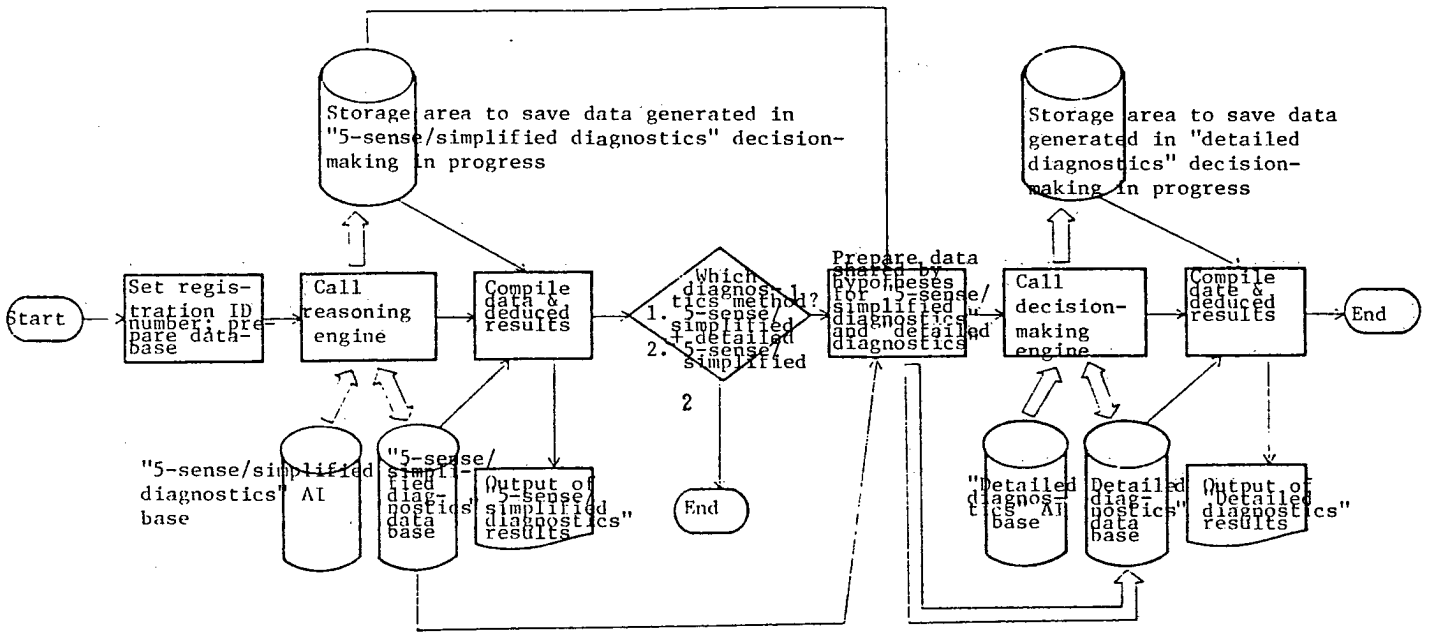


Figure 4 Flowchart of actions taken by ES in rotational device diagnostics

As for the decision-making process, first, a relatively limited volume of information--machine specifications, data obtained through the five senses, and simplified diagnostic information--is input to conduct a rough assessment based on a forward-feed reasoning. Next, this result is considered as the hypothesis on which to base the conclusion, and backward-feed reasoning is conducted based on observable conditions selected from many possibilities stored inside in the system. This way, a highly accurate diagnosis is made possible. In actual use of the system, the user performs the 5-sense/simplified diagnostics as part of a daily routine check. If any abnormality is detected, additional observed data, collected primarily by frequency analysis of machine vibrations, can be fed into the system to run a detailed diagnosis.

This ES program is a prototype model for general purpose usage, which incorporates into the design as many as possible of the causal factors that commonly affect many rotating machines.

We are considering ways to improve diagnostic accuracy, and the design of the man-machine interface. Also, we need to provide more knowledge within the system regarding generation of remedial measures. These issues will be addressed during future field tests.

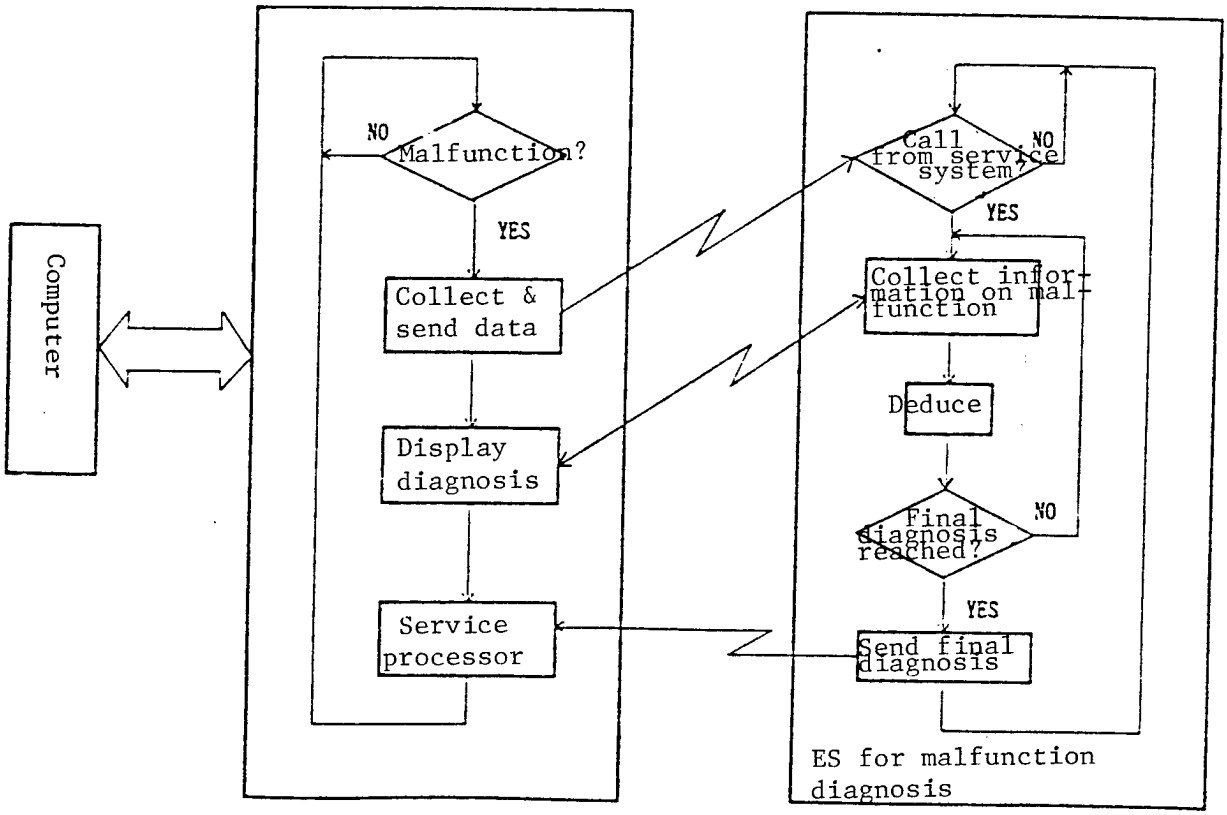


Figure 5 Computer malfunction diagnostics using ES

4.4 Diagnostic ES for Computers

Industrial computers play central roles in production processes, and therefore are indispensable. Therefore, maintenance of the computer is a very vital issue. This is especially so considering the trend of larger and more sophisticated computer systems. Usually, [when a system error occurs,] a computer is capable of storing RAS (reliability, availability, and serviceability) information which may be retrieved later to evaluate the malfunction. But conventionally, the use of this RAS function assumes mediation by the human operator at some stage of the troubleshooting process, and so, the diagnostic accuracy and speed really depends on this individual's expertise. The expert system we developed for diagnosis of computer malfunctions is an on-line system with remote-control capabilities. Based on the built-in knowledge base which contains expert methodologies, the system performs hierarchical reasoning actions in accordance with the RAS informaiton as well as the vast volume of collected data which is sorted out for suspected malfunction modes. As shown in Figure 5, the system's processing flow consists of several steps: first, a preliminary diagnosis is conducted based on the transmitted description of the abnormality; then, based on this preliminary result, the types of information needed are identified and collected, while a diagnostic session is conducted; and finally, this second step is repeated until the final conclusion is drawn, and the malfunction is identified. Therefore, by installing service processors in the computers on the plant floor, and connecting them to the diagnostic system via transmission lines, a very efficient scheme can be realized for management of the plant computers from a single support center. The ES program can substitute for skilled technicians in providing proper diagnostics, which may be accessed from any of the computers, should an error occur

5. Conclusion

5.1 Recommendations for ES Development

In recommending COMEX-based diagnostic ES programs, we would like to highlight the advantages drawn from our own experiences in developing similar programs: (1) Non experts can achieve quality diagnosis with an expert-like competency, and the system's reasoning process and assumptions are easily understandable. (2) an expert himself can build a system. Also, it is easy to administer revisions to the knowledge base through provided editing utilities. (3) A system thus created can be readily interfaced to an off-line system or an existing system. It is possible to organize knowledge in a structured format. (4) The system is run on a PC, and the Japanese language may be used. The system also has a good user interface.

5.2 Future Perspectives

The following issues deserve future research attention:

(1) Development of ES for Comprehensive Diagnosis"

Most of the systems reported above are designed for assessment of the cause of the malfunction. For measures to be taken to rectify the error, the user has to refer to the accompanying manual. But in actual practice of a troubleshooting measure, it is necessary to consider not only the evaluation of the machine's particular state but also other factors like how vital a role the subject machine plays in the concerned process; its history of malfunctions; and the company's plan for maintenance and inspection of products. To develop an expert system whose primary goal is to achieve this type of comprehensive product maintenance and inspection, the decision-making scheme will have to be even more complex, and the knowledge base in an even more extensive hierarchical structure. Also, considering the fact that a plant is an entity made of numerous machines and technologies, development of many more types of ES, tailored for these different objects, as well as development of a comprehensive network consolidating these ES modules are important issues for future research.

(2) Improvements in the Design of the Knowledge Base:

One of primary purposes of defect diagnosis is to predict a future malfunction. In many instances, once the diagnosis is drawn, corrective measures are taken before the imminent breakdown or failure of the malfunctioning unit, and therefore, it is not often possible to verify the deduced results. In this respect, on-site experience in evaluating degradating phenomenon prove to be a source of especially precious knowledge. Thus, it is essential to conduct an immediate follow-up on such occasion in order to improve the design of knowledge base.

We would like to thank Prof. Haruki Ueno of Tokyo Denki University, and colleagues at Fujitsu Facom Control Co. Ltd. for their generous counsel; as well as Mr. Kawano of Idemitsu Petrochemical Co. Ltd. for providing valuable information without which we could hardly have completed our development project.

Expert System for Diagnosis of Garbage Incinerator Operation
-- Processing of Fuzzy Factors in Knowledge --

43064091d Tokyo 1988 JOINT CONVENTION RECORD OF INSTITUTE OF ELECTRONICS AND INFORMATION ENGINEERS OF JAPAN in Japanese 1988 pp 1-167 - 1-170

[Article by Hajime Ase, Tetsuzo Tsukioka, and Masamichi Tate, Nippon Kokan K. K.]

[Text] 1. Abstract

Refuse materials that are produced by city life are processed in an incinerating plant to reduce waste volume and remove pollutants. The incinerating facility has also come to serve as a source of heat supply

for power generation, a notion popularized after the oil shock. Also, the design of the waste incinerator must comply with the recently tightened anti-pollution regulations which further curtail discharge of hazardous substances in the exhaust gas into the environment.

Viewed from the standpoint of control engineering, the combustion process, which plays the core role in the incinerator operation, may be considered a multi-variable interaction--multiple controlling quantities interact with multiple states. Furthermore, as external interference, the refuse materials themselves, constantly disrupts the process, it is difficult to grasp the characteristics of the process. Therefore, inexperienced operators often find it difficult to maintain stable operation of the plant. To solve this problem, automatic combustion control systems were developed. With such a system in use, operation of the plant furnace may be automated under normal circumstances. However, because of restrictions placed on data monitoring capabilities, a situation which the system in its automated mode may not be able to handle can break out. In such a case, an expert operator standing by can perform troubleshooting procedures. We have developed an expert system for the refuse incinerator. It provides assistance in operation of the plant when it is being run by an inexperienced operator. By applying AI technology, we created a computer-based system which possesses expert knowledge of plant operation. The knowledge is stored in what is called a knowledge base.

As a technical discussion of the expert system has already been included elsewhere in this article, in regard to the evaluation of the adopted expert knowledge, the construction of models, and the evaluation of the knowledge base structure and overall system performance, we would like to briefly outline the highlights in Section 2. From the point of view of AI applications in manufacturing facilities, we will discuss in Section 3 techniques--the one employed in our expert system as well as those investigated--for processing fuzzy factors contained in knowledge since the idea can be applied to practical production facilities other than garbage incinerators.

2. Summary of Incinerator and Expert System for Operation

Figure 1 shows a diagrammatic organization of the garbage incinerator plant and the expert system for plant operation. The figure's left section represents the plant. Waste materials carried by trucks are temporarily stored in the waste-disposal pit. After being stirred, they are crane lifted and thrown into the hopper. Inside the hopper, waste materials are transported on the fire grate by conveyer. As they are transported by the grate's conveyer, the wastes are dried by preheated air drafts coming from below and then burned. Finally, only burned ash remains and is dumped out at the exhaust. The combusted exhaust gas reaches as high as 900°C. It is collected by the boiler, and the stream that's generated is sent to the turbine unit to generate electrical power. The combustion in the incinerator is controlled by varying the fire grate speed, the volume and temperature of the combustion draft, the amount of cooling air, the valve adjustment for the secondary smoke-pipe dumper, the kiln rotation speed, the volume of the draft under the fire grate, and the amount of water vapor.

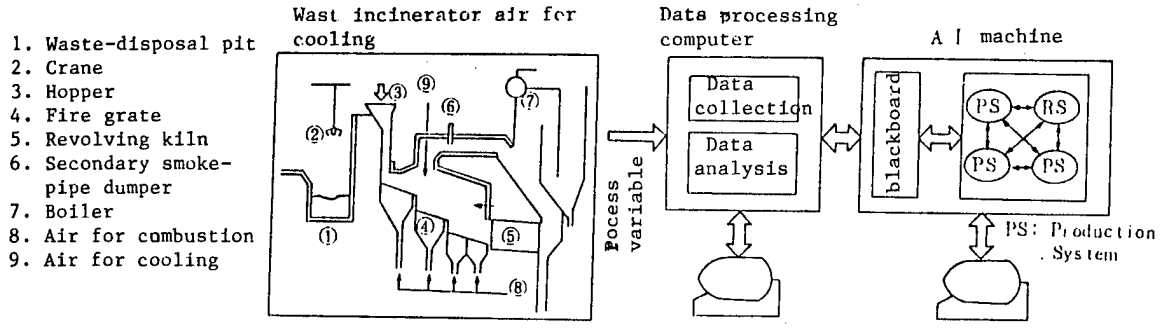


Figure 1 Expert system for waste-incinerator operation

The expert system possesses knowledge both for the prediction of malfunctions and the cancellation of malfunction predictions. Its structure is designed in such a way that the knowledge it contains may be divided and reorganized according to different types of abnormality states, and collaborative-style reasoning, connecting tasks which are grouped knowledge entities, are made possible. This reasoning is based on consideration of the transitions among operation nodes and their priority order among operation nodes. The system's unique structure also allows for hierchical structuring among tasks. In addition, the system comes with a blackboard for the sake of achieving collaboration between tasks, and for storing common data used in the reasoning session by various tasks modules. As shown in Figure 1, the system is organized into two major units; the decision-making unit and the data processing unit. The former unit is even equipped with communication functions to couple these units.

3. Processing of Fuzzy Factors in Knowledge

3.1 Fuzzy Factors in Knowledge for Incinerator Operation

Often the operator's knowledge about the operation of a plan facility, which is continuously running, is based on his predictions concerning the chronological state transition of the operation. For the operation of the incinerator, the knowledge needed for predicting abnormal conditions is also a prediction of the state transition itself. In addition, knowledge for devising possible moves in case a malfunction is detected is always substantiated by tacit conviction upon the prediction of a state transition--conducting this mode of operation against that mode will rectify the predicted abnormal state.

Knowledge of such a state change prediction cannot be fully evaluated and expressed in terms of known the formulae of AI science since it embodies reference to an ambiguous subject. In other words, this type of knowledge contains fuzzy factors. This situation is especially valid in cases where the causality relationship has to be determined for a great number of input and output factors.

3.2 Support for Incinerator Plant Operation

In response to the occurrence of an abnormal event during incinerator operation, the expert system for plant operation provides instructions to the effect, say, for increasing the processing volume or decreasing it in order to rectify the abnormality. The rule used in this particular instance of decision-making action may be generally expressed as:

IF abnormal state y_i is predicted because process state x_i is in a_i THEN change processing volume u_k to b_k. (1)

If there is just one predicted abnormal state, or even if there are multiple abnormal states, as long as the corresponding measures for them do not interfere with each other, proper instruction can be generated according to rule 1. However, for [other] multiple abnormal states, it is possible that prescriptions for varying processing volumes may affect each other. For instance, consider two rules:

IF abnormal state y_1 is predicted because process state x_1 is in a_1 THEN change processing volume u_1 to b_1 (increase by $\hat{-}b_1$), (2)

IF abnormal state y_2 is predicted because processing volume x_2 is in a_2 THEN change processing volume u_1 to b_2 (decrease by $\hat{-}b_2$), (3)

and suppose that a situation induces application of both. Rule 2 instructs and processing volume u_1 to be increased by $\hat{-}b_1$, while rule 3 tells to decrease u_1 by $\hat{-}b_2$. As there is competition, use of these rules alone cannot determine the correct amount of u_1.

It is still possible to resume the reasoning process by resorting to rules specially prepared for conflicting situations like the above. But, as our studies with experts revealed, a skilled operator does not look for solutions in terms of IF-THEN rules in such a situation. Experts stress the fact that "It depends on the circumstances." Our investigation of their approach characterized by this statement identified the following points:

(1) A skilled operator predicts how the current state will change in an assigned time if a certain operation is administered.

(2) He selects an operation that best rectifies as many of the known abnormal states as possible.
(3) In case it is impossible to avert interference, he selects an operation which minimizes the ill effect.
(4) He assigns an ill effect, small in extent, to a low priority abnormality item.

These findings may be summarized this way: do not just rely on simplistic experience-based knowledge, and make decisions according to vast knowledge of the prediction of the state transition. We have constructed a model using the Fuzzy theory for the above actions 1-4. The following is the algorithms thus determined:

Step 1: Solve for the set of possible operations (u_k) which act upon the predicted abnormal state 1.

Step 2: Determine by using the following formula how the degree of risk associated with the predicted abnormality 1 changes in time if the operation u_k is performed:

$$\tilde{h}_\ell(y_\ell) = \sup_{x_\ell \in X_\ell} \left[h_\ell(x_\ell), h_{R_k}(x_\ell, y_\ell) \right] \qquad (4)$$

x_ℓ, y_ℓ : State variables of predicted abnormality 1; X_ℓ: Present, Y_ℓ: Future
X_ℓ : Space where state variables x_1, y_1 exist
R_k : Factor affecting transition of level of risk involved in operation u_k
$h_\ell(x_\ell)$: Membership function for present level of risk involved in predicted abnormal condition 1
$\tilde{h}_\ell(y_\ell)$: Membership function for future level of risk involved in predicted abnormal condition 1
$h_{R_k}(x_\ell\ y_\ell)$: Membership function for identifying transition of x_1, to y_1, if u_k is performed

By denoting fuzzy sets by x, and y, which respectively describe present and future levels of risk, we may say that formula 4 is really an expression of the fuzzy synthesis $y = R_k * x$. Applying this notation, we can express a situation with multiple operations as a multi-operation fuzzy synthesis like:

$$y = R_{k1} \circ R_{k2} \circ \cdots\cdots \circ R_{km} \circ x. \qquad (5)$$

Step 3: Obtain an operation or a combination of operations for the predicted abnormality 1 by comparing the present risk and the future risk. Once we know the present value of state variable $x_1 = x_1^\circ$, we can obtain the present and future risk levels as $h_1(x_1^\circ)$, $h(\tilde{h}_1^\circ)$. The following evaluation is performed:

$$\tilde{h}_\ell(x_\ell^\circ) \leq h_\ell(x_\ell^\circ) \quad ; \text{For all predicted abnormal conditions 1's} \tag{6}$$

$$\tilde{h}_\ell(x_\ell) \leq \varepsilon \quad ; \text{For the lowest priority 1} \tag{7}$$

$$\|\tilde{h}(x^\circ)\| \rightarrow \min. \quad ; \text{Minimize with respect to } \{u\} \tag{8}$$

Inequality 6 tests which of the predicted abnormal conditions may be rectified. Inequality 7 tests if the level of risk involved in an abnormal item with a lower priority may remain small, within an allowable range, in the future. Formula 8 determines a final compromising action.

3.3 Numerical Examples

Let us cite a couple of examples with specific numbers. Let

$$\{1\}=\{1,2\}, \quad \{u_k\}=\{u_1, u_2, u_3\}.$$

Let the membership functions be as shown in Figure 2.

$h_1(x_1)$

x_1^1	1.0
x_1^2	0.9
x_1^3	0.8
x_1^4	0.6
x_1^5	0

$h_{R1}(x_1, y_1)$

	y_1^1	y_1^2	y_1^3	y_1^4	y_1^5
x_1^1	0	0	0	0	1.0
x_1^2	0	0	0.1	0.2	0.5
x_1^3	0	0.1	0.2	1.0	0
x_1^4	0	0.4	1.0	0.6	0
x_1^5	1.0	0.8	0.4	0.1	0

$h_{R2}(x_1, y_1)$

	y_1^1	y_1^2	y_1^3	y_1^4	y_1^5
x_1^1	0	0.6	0.8	1.0	1.0
x_1^2	0	0.8	1.0	0.5	0
x_1^3	0	1.0	0.5	0	0
x_1^4	0	0.5	0	0	0
x_1^5	1.0	0	0	0	0

$h_2(x_2)$

x_2^1	1.0
x_2^2	0.8
x_2^3	0.5
x_2^4	0.2
x_2^5	0

$h_{R1}(x_2, y_2)$

	y_2^1	y_2^2	y_2^3	y_2^4	y_2^5
x_2^1	0	0.4	0.8	1.0	1.0
x_2^2	0	0.8	1.0	0.5	0
x_2^3	0	1.0	0.5	0	0
x_2^4	0	0.5	0	0	0
x_2^5	1.0	0	0	0	0

$h_{R2}(x_2, y_2)$

	y_2^1	y_2^2	y_2^3	y_2^4	y_2^5
x_2^1	0	0	0	0	0.5
x_2^2	0	0	0	0.2	1.0
x_2^3	0	0	0.2	0.8	0.8
x_2^4	0	0.3	0.8	0.8	0.5
x_2^5	1.0	1.0	0.8	0.5	0.2

$h_{R3}(x_1, y_1) = h_{R1}(x_1, y_1)$, $h_{R3}(x_2, y_2) = (\delta_{x2y2})$

Figure 2 Membership function

Example 1: Let $x_1^{\circ}=x_1^3$, $x_2^{\circ}=x_2^2$.

Then the degrees of risk changes as follows:

	u_1	u_2	u_3
1	0.8->0.6	0.8->0.9	0.8->0.6
1	0.2->0.8	0.2->0.2	0.2->0.2

Although no operation is possible which can reduce risk for both 1 and 1 .

Example 2: Let $x_1^{\circ}=x_1^2$, $x_1^{\circ}=x_2^4$.

Then the degrees of risk changes as follows:

	u_1	u_2	u_3	u_2 & u_3
1	0.6->0.4	0.6->0.8	0.6->0.4	0.6->0.4
1	0.8->1.0	0.8->0.5	0.8->0.8	0.8->0.5

In this example, single operation alone cannot reduce both of the risks in 1_1 and 1_2, and inequality 7 cannot be satisfied. But by performing a combination of operations u_2 and u_3, the risk can be minimized.

3.4 Special Considerations in Developing an Algorithm

Many algorithms are possible for the procedure mentioned in Section 3.2. The processing volumes may be determined in accordance with the fuzzy control method. The normal degree of risk used in formula 8 may be either the maximum or the iterated sum. In this section, we will discuss one approach, based on formula 4, that we took in trying to design a processing pattern as close to a human thought process as possible.

To simplify the argument, we choose just one state variable and one processing variable.

Example 3: $X=\{x_1,x_2\}$

$h(x_1)=0.5$, $h(x_2)=0.6$, $hu(x_1,y)=0.9$, $hu(x_2,y)=0.7$; $y \in X$

Case 1: Process Based on Fuzzy Theory

Use of formula 4, $h^{\sim}(y)=6$, means the risk level $h(x_i)$ does not affect the final action at all. But, by throwing out this information, we may draw a conclusion inconsistent with an intuitively obtained action by a skilled operator on site.

Case 2: Process Based on Probability Theory

Provided the elements in X are exclusive of each other and independent, h(x), hu(x,y) may be converted into probabilistic values by using the following formulae:

$$p(x) = h(x)/K_1 \quad ; \; K_1 = \sum_{x \in X} h(x) \tag{9}$$

$$p(y/x) = hu(x,y)/K_2 \; ; \; K_2 = \sum_{y \in X} hu(x,y) \tag{10}$$

If, however, K =0 (K =0), we set K =|X| (K =|X|), where |X| is the determinant of the set X. Using equations 9 and 10, equation 4 can be rewritten as:

$$p(y) = \sum_{x \in X} [p(x) \cdot p(y/x)] \tag{11}$$

Equation 11 is applied to example 3 to get p(y)=0.49. This is the result for the case where λ = 1. with $h(x_1)=0.5 \times \lambda$ and $h(x_2)=0.6 \times \lambda$. But as long as the condition $0 < \lambda \times h(x) \leq 1$ is satisfied for any arbitrary value of lamda, the probability will turn out to be the same. Let the state of knowledge be expressed $S:(h(x_1) \times \lambda, h(x_2) \times \lambda)$. For instance, there is a tremendous difference in embodied information between, say, S_1:(0.5, 0.6) and S_2: (0.25, 0.3). But, we still end up getting the same probability. The reason is that we have omitted a rule (concept) which can indicate the level of incompleteness of the knowledge.

Case 3: Process Based on Dempster-Shafer Theory

Probability theory assumes that we have complete knowledge as given by $p(x_1)+p(x_2)=1$. In contrast, by using the D-S theory, we can treat the knowledge as being incomplete. Here, 2 expression spaces are defined in terms of functions Bel, Pls (See Figure 3). These functions calculate basic probability, and are defined as follows:

$$\mathrm{Bel}(B) = \sum_{A \subset B} m(A) \tag{12}$$

$$\mathrm{Pls}(B) = \sum_{A \cap B \neq \phi} m(A) \tag{13}$$

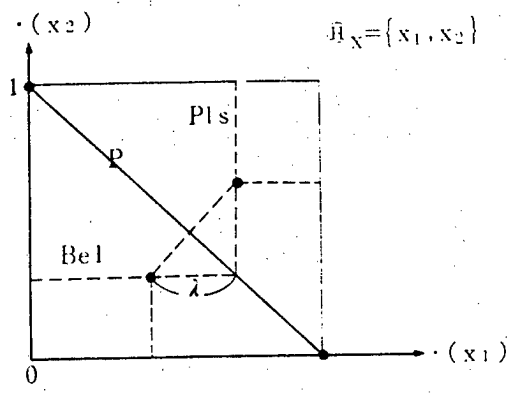


Figure 3 D-S theory-based spatial representatiaon of knowledge

where m is

$$m : 2^{\Theta} \rightarrow [0, 1]$$
$$m(\phi) = 0 \qquad (14)$$
$$\sum_{A \subset \Theta} m(A) = 1$$

According to the D-S theory, m determines all the information. The equation $\lambda = 1 - \sum_{x \in X} Bel(x)$ gives the parameter ($0 \leq \lambda \leq 1$) which indicates the level of incompleteness of the knowledge in question. With $\lambda = 0$, the knowledge is complete, and its value matches that of the probability. $\lambda = 1$ means we have no knowledge at all.

We describe the result of applying this model to example 3 below. Since with $\Theta_X = \{x_1, x_2\}$, we have $\sum_{x \in \Theta_x} h(x) = 1.1 > 1.0$, we can equate the h (x) to Pls. Then,

$h(x_1) = Pls(\{x_1\}) = 1 - Bel(\{x_2\}) = 1 - m(\{x_2\})$, $h(x_2) = Pls(\{x_1\}) = 1 - Bel(\{x_1\}) = 1 - m(\{x_1\})$

And,

$m(\{x_1\}) = 1 - h(x_2)$, $m(\{x_2\}) = 1 - h(x_1)$, $m(\Theta_x) = 1 - (m(\{x_1\}) + m(\{x_2\}))$

Also, we expand the D-S theory, and define the conditional basic probability as:

$$\sum_{y \in \Theta_Y} m_{XY}(\{y|x\}) = 1 \qquad \text{for } \forall x \in \Theta_X$$

$$m(\{y\}) = \sum_{x \in \Theta_X} m_{XY}(\{y|x\}) \cdot m(\{x\}) \qquad (15)$$

From this, the membership function's hu which give the transition state can be expressed as:

$$hu(x_1, y) = m_{XY}(y|x_1)$$
$$hu(x_2, y) = m_{XY}(y|x_2) \qquad (16)$$

Thus, by substituting these in equation 15, we obtain

$$m(\{y\}) = 0.71,\ m(\Theta_Y) = 0.29$$

The three cases discussed above are summarized in Table 1.

	$\{y\}$	$\{\tilde{y}\}$	Θ_Y
Case 1	0.6	—	—
Case 2	0.49	0.51	—
Case 3	0.71	—	0.29

From this table, it is obvious that even if the fuzzy factors are represented by values [0,1] alone, the result will differ greatly depending on what type of process for the interpretation of data is elected. Although it is beyond the present scope of the subject being considered to determine which case is more realistic, we would like to draw the reader's attention to the D-S theory discussed in case 3. This theory goes a step beyond probability theory by adding a means to measure the degree of incompleteness of knowledge. As a theory for handling ambiguous situations subject to subjective human actions, it is a very promising work. But, since the results given by this theory also contain subjective elements, it will be necessary to first establish standards to evaluate such results before the theory can be widely applied to practical cases.

4. Conclusion

We focused on methods for processing fuzzy factors found in knowledge from the standpoint of AI applications in plant facilities. These methods were developed in conjunction with our development of an expert system for waste incinerator operation. We have seen that in the area of monitoring and control technology, it is often necessary to augment a standard production system with a problem-solving system which considers the characteristics of the subject. The methods discussed in this article have been tested for effectiveness in the design of such a system.

Application of AI technology in a Sewage Processing Plant

43064091e Tokyo 1988 JOINT CONVENTION RECORD OF INSTITUTE OF ELECTRONICS AND INFORMATION ENGINEERS OF JAPAN in Japanese 1988 pp 1-171 - 1-174

[Article by Kazuo Maeda, Mitsubishi Electric Corp.]

[Text] 1. Introduction

A water treatment plant has to process sewage from many sources containing complex arrays of microorganisms--chief contributors to the making of sewage--and since the operation is not usually backed by a sufficient number of practical on-line telemeters, delays in investigating the characteristics of the process have occurred. Therefore, traditionally, successful management of a water treatment plant relies heavily on experienced operators. A skilled operator can look at the so called complexion of the treatment process--inspecting coloration and odor of the inlet water, treated water and sludge--and adjust the plant operation very efficiently. In recent years, however, the issue of securing skilled operating personnel is surfacing as a large problem, because as the average age of a worker is becoming progressively higher at the large plants, small to medium plants are having difficulties in recruiting enough operators.

To alleviate the problem, plant managers have started to look at a different solution, namely an expert system which incorporates expert knowledge of plant operation, distilled and organized into an AI format. Now, we are witnessing the beginning of ES development in the field as more prototype and practical systems are being developed. As it appears that the present status of plant operation management and the basic framework of the expert system happen to mix well, time is ripe for packaging of the know-how. We know that there will be a very receptive market.

In this article, I will first evaluate the needs of water treatment technology. Then, by citing specific examples, I will discuss the working of the expert system for water treatment operation. Finally, I will comment on future issues for next-generation expert systems.

2. Requirements of Water Treatment Technology

2.1 Requirements for Design of Treatment Process and Operation Support

Recent research works have significantly raised the level of understanding of the ecology of microorganisms. For instance, research with statistical methods exemplified by the cluster analysis has led to successful classification of several distinguishing [ecological] patterns. Also, it has led to identification of many factors related to the transition of the [ecological] state. These could not have been found by conventional microscopic tests alone. In the meantime,

although procedures for minor loop controls such as DO (dissolved oxygen) control have just about been established, as far as high-level operation controls--adjustment of parameters and setting of target values--for the same systems are concerned, we still have to rely heavily on field operators. Therefore, at this stage, design of systems that combine theory-based numerical process with on-site heuristic knowledge is highly desirable. Such techniques can lead to improvements in productivity for determining the procedural specifics of the process and controlling it.

2.2 Manual Operation sin Water Treatment Management

Critical-time tasks such as start-up operations, adjustments to the plant operation mode, and plant emergency shut-downs often have to be done manually, and therefore, successful execution depends on how experienced the field operator is. Also, it is very difficult to refer to the operation manual when an accident does happen, since manuals are written per instrument, and they usually come in hundreds of pages.

Development of systems which can help prevent mistakes by operators on manual tasks is highly desired as it will lead to effective prevention of manual-task related accidents.

2.3 Treatment Plant Diagnosis and Corrective Measures

Expert knowledge regarding diagnosis of the process and prescription of corrective measures is complex and diverse. Normal operation aside, if an emergency situation occurs, it is not easy even for a very experienced operator to make the right decision quickly. A computer-based support system is highly desirable.

(figure 1 on following page)

3. Examples of Operation Support Systems

3.1 Support System for Treatment Plant with Knowledge-Based Reasoning Capabilities and Electronic-Assisted Manual Mode

As shown in figure 1, this new system features a reasoning engine which accesses a built-in numerical data base, knowledge base, and algorithm base for reasoning information to work out the operation support procedure, and it also features an electronic-assisted audio-visual operator interface unit.

With a decision-making system which incorporates a production system (PS), the data stored in the abovementioned working data base is regarded as the initial state. The decision-making module operates on this data and generates instructions for the job to be executed. Specifically, if the content of the data in the working data base matches the assumption for a rule, the content of the rule's result

section is updated. If a match with a rule is not found, another appropriate rule is looked for and an attempt at a match is tried again. Sometimes, the numerical simulator is turned on as well in order to evaluate alternative remedies.

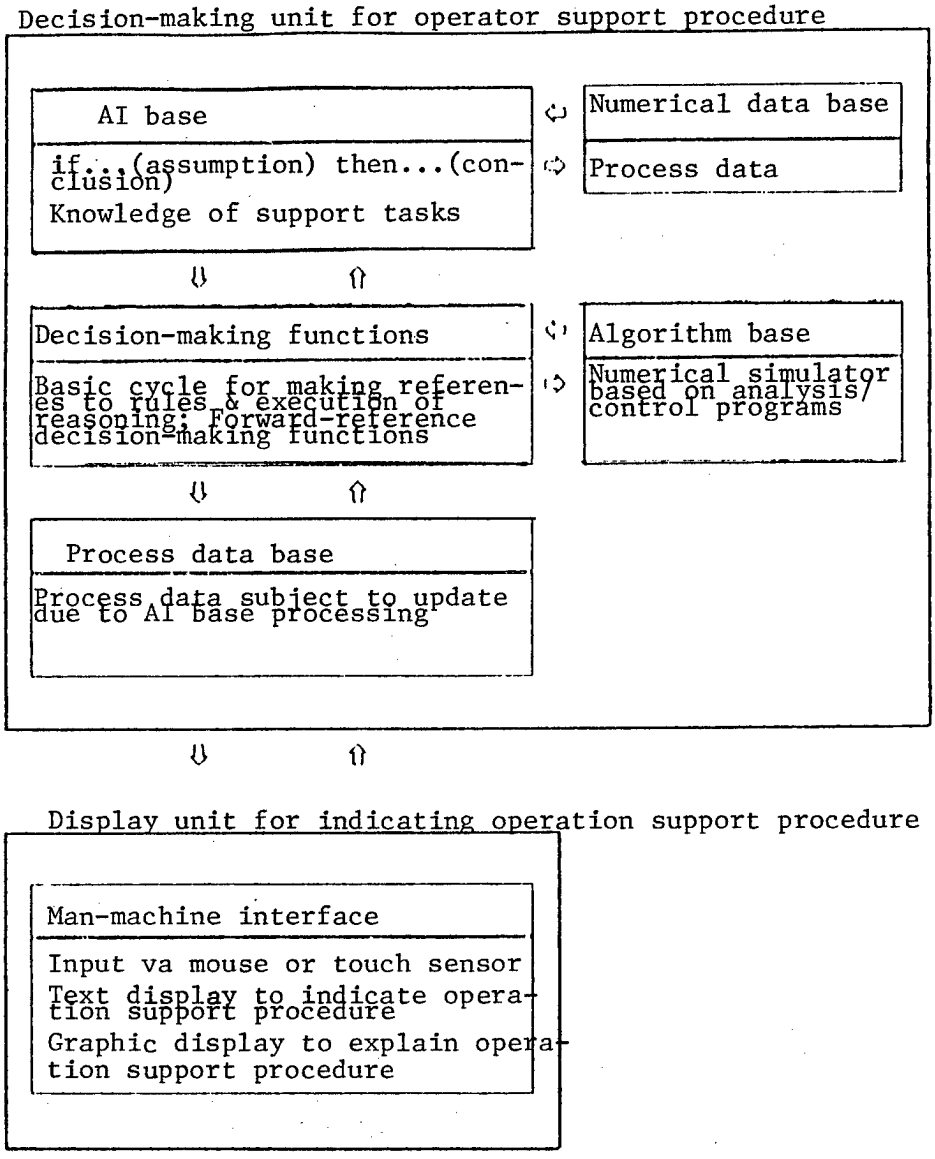


Figure 1 General configuration of an operation support system for a water treatment plant

As for the interface unit, it displays the instructions written to the working data base for the generated job procedures in a dialogue-type format with appropriate graphics on the screen.

The interface with the outside world is achieved via a unit for input of on-line data from the monitoring systems, as well as via a graphics interface. The operator may use the mouse or touch panel to give instructions.

Let me cite an example of a system in use for the task of deducing the cause of a malfunction and prescribing corrective measures in accordance with the water's quality control, which is a part of the activated sludge process--the sewer treatment's central process. In table 1, typical procedures for troubleshooting abnormal conditions are tabulated. Table 2 is an example of a decision table. Figure 2 illustrates the outlined procedure. Not only are the information on the process, and the current values fixing the management mode considered, but the history of occurrence of abnormal conditions is accounted for in determining if an abnormal state is impending. Should an abnormality occur, an alarm is issued to alert the operators. Following detection of this abnormality, the expert system starts working on the knowledge base for abnormal signs to deduce the cause for the abnormal state, and to prescribe instructions for remedies. Finally, verification as well as evaluation are conducted by the operator. The system presents the instructions to the operator in the manner of conducting a dialogue through the computer. Specific examples of a dialogue are shown in Figure 3.

Tests were conducted with prototype versions to improve the user interface capabilities of the system. Results showed that use of a mouse helps facilitate menu selection, and touch panels help ease input operations. With graphical pictures incorporated, more intuitive presentation of instructions was made possible. Also, since a tree display format was used to quickly grasp the description of the ongoing reasoning process, an inexperienced operator without any knowledge of a computer could demonstrate expert-like handling of plant operations.

Basic flow of actions by the support system

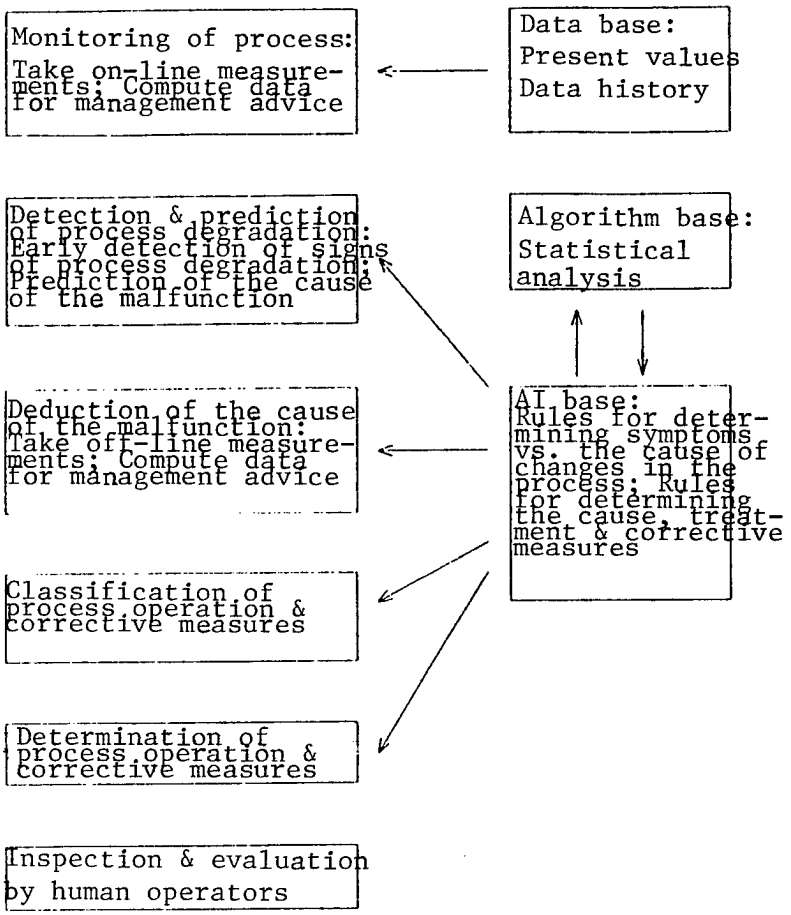


Figure 2 Flow of basic actions supporting water treatment operation management

Table 2 Decision table for process management; troubleshooting list for degradation management
(Note 1)
(Part 1: Discharged water COD:BOD or more)
(Related to sewage intrusion)

Diagnosis-led action		Deduced abnormal factors	All Initial level of pollution load due to rain water intrusion
Monitoring of process	Inlet water	Amount of water	(1) ‹ Increase
		COD	(2) ‹ Worsening
		SS	
		pH	
		Temperature of water	
		Health-hazardous substances; substances difficult to break down; Nutrient ratio; BOD; Discoloration; Foul odor	Rain after successive sunny days (10㎜/h)
	In-plant water	Amount of water	
		COD	
		SS	
		Discoloration	
	Initial sedimentation	COD	
		SS	
	Gas exposition tank	Returned sludge level	
		SV	
		Settling velocity	
		Rr	
		Amount of iodine consumption	
		Iron ion	
	Final sedimentation	COD	(1) ‹ Worsening
		SS	(2) ‹ Worsening
		pH	
		Transparency	
		Settling velocity №2, №3	(2) ‹ Worsening
		Discoloration	
		Foul odor	
		VSS	
		Soluble BOD	
	Visual inspection	Sludge surface	
		Sludge spread	
		Sludge flotation	
		Bulking	
		Sludge decay	
		Discoloration	
		Whipworm	
		Sludge disintegradation	
Indication of process management action to take		SVI BOD/SS load Anti-pollution ordinance; Air magnification; Gas exposing tank residence time	
Process diagnosis: cause & condition			Singularly excessive high-concentration COD due to initial rain intrusion
Guidelines for setting corrective measures & actions			MLSS Increase in air blowing volume
Process operation	Configuration values	MLSS	(2) slight increase
		DO	(1) ditto
		Return volume	
		Extraction volume	
		Air circulation volume	
	Operation costs	Air circulation power cost Pollution processing cost	

Note 1: The reference numbers in the table indicate the level of priority for performing the inspection or remedy.

3.2 System for Operation Support of an Activated Sludge Process Using an Adaptive Production System

This system also adopts a production system. It incorporates an adaptive function in the design of the rule control. The adaptive function is a relatively simple routine designed to automatically correct the coupling coefficient between production rules. With this built in, the system can dynamically vary the priority order of rules. Also, the system's search for procedures for [plant] operation and the diagnosis of the [plant] process can be accelerated. Furthermore, conclusions can be predicted in a forward-feed manner.

According to the results of tests with prototypes, with the addition of the rule adaptive routine, dialogue sessions with the user became more flexible and natural.

3.3 Operation Support AI System for Heavy-Load Activated Sludge Process

This system also uses a PS, and incorporates both forward-feed and backward-feed reasoning functions. As a yardstick to measure the degree of fuzziness, the idea of measuring the level of certainty in rules is adopted. The system consists of a rule data base, reasoning engine, rule editor, etc.

The system was put to use at a working plant to check operation integrity. The results show that more detailed management of the plant was achieved than using a conventional technique. Early detection of a worsening of the sludge condition was made possible, and compared to use of a normal method, the new system could maintain stable operation of the plant with a load of up to 4 times as large.

4. Problems in Present Systems and Issues for Future Systems

4.1 Problems in Present Systems

(1) Although extensive arrays of rules which do meet conditions but are chained to each other without much of a meaningful relationship are available, they more or less represent a shallow argument. We need to assign deeper meaning to the chain itself.

(2) Facilities are absent for performing common sensical and flexible reasoning. Also, if the actual situation differs even slightly from the assumed condition, the system fails to respond correctly.

(3) While expert knowledge is obtained through holding interviews and apprenticeship-like sessions, it is still very laborous to translate into an AI expression experiences felt and sensed.

(4) The dialogue mode for the question and answer session is monotonous and boring. Some type of capability for learning is desired.

(5) On-line real-time requirements are not satisfied.

(6) System's capabilities to talk to [subject] machines are limited.

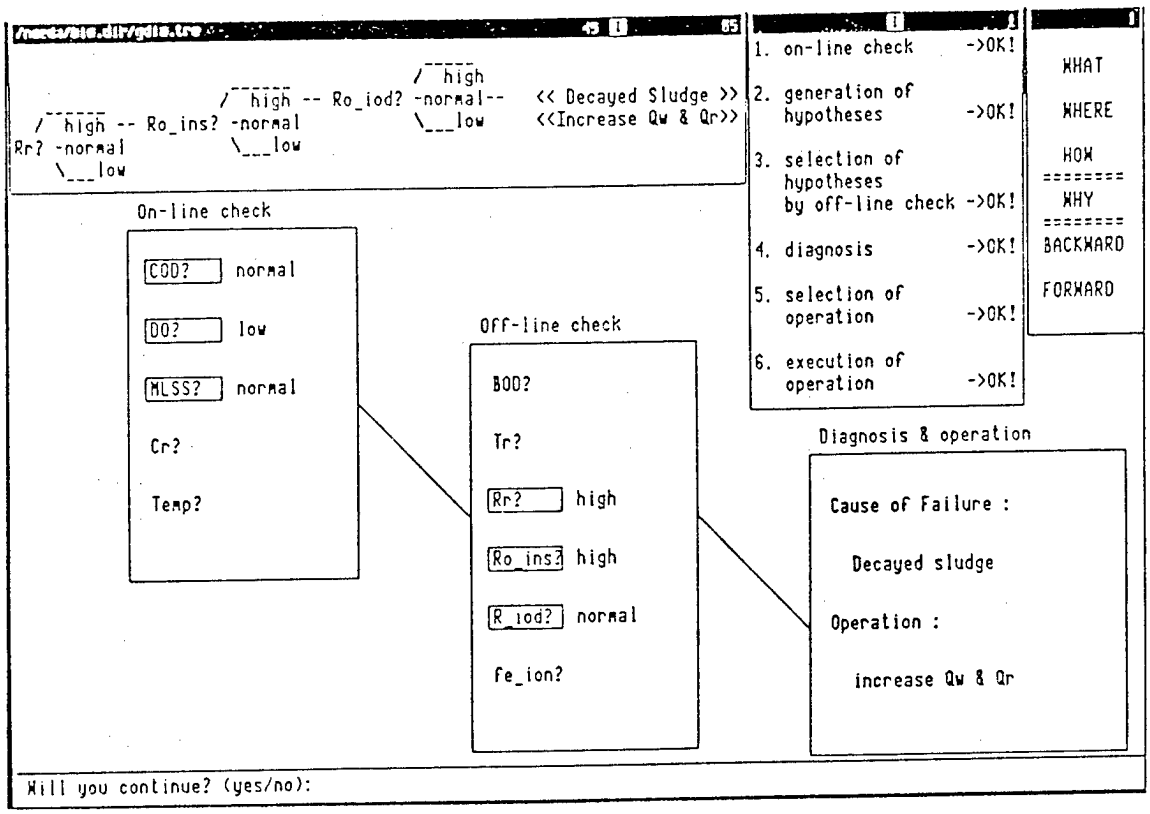


Figure 3 An example of a screen frame from a dialog with the water treatment plant operation support system

4.2 Issues for Future Systems

(1) A Utility must be made of information on the subject's structure and function, and of the vast knowledge such as established knowledge on dynamic phenomenon. Facilities for modeling support and simulation capabilities must be expanded.

(2) The flexible attitude of sticking to the template matching method must be stopped, and methods of analogy drawing and generalization should be implemented.

(3) More powerful facilities for editing knowledge rules--support functions for knowledge collection--are needed. More sophisticated routines for detecting inconsistency among rules are needed. More extensive decision tables are needed. These tables serve as a source of knowledge. Better functions for conversion between trees and groups of partial rules are needed.

Table 1 Troubleshooting list for abnormal conditions

Abnormal condition	Cause	Remedy
Abnormal levels of turbidity and SS in released water	Intrusion of plant sewer; rotting of sludge and deposition; Piled sand; Dense flow and shortcircuited flow; excessive return load; decrease in coagulation and precipitation strengths; maladjusted MLSS; and dispersion, disintegration, swelling and floatation of sludge	Readjust MLSS/DO; stabilize incoming water flow; dump decayed/ swollen sludge; and limit intrusion of plant sewer
Abnormal level of COD in released water	Maladjusted MLSS; nitrification; and dispersion, disintegration, swelling and flotation of sludge	Readjust MLSS/DO; and stabilize incoming water flow
Abnormal load of incoming sewer	Intrusion of plant sewer; and excessive returned water load	Limit intrusion of plant sewer/returned water load; and study temporary discharge/sludge treatment process
Abnormal level of DO	Too much or too little MLSS; excessive variation in volume of sewer water inflow; insufficient air draft; blinding of air filter; intrusion of plant sewer; intrusion of decayed sludge; excessive return load; change in organism profile; abnormal distribution of DO levels; nonconformity in initial sludge settlement; and maladjustment of channeling volumes of incoming sewer	Readjust MLSS; stabilize incoming water flow; readjust air draft; inspect air distribution unit; limit intrusion of plant sewer, decayed sludge; and returned water load; and conduct microsection tests
Abnormal level of MLSS	Too much (or too little) MLSS; excessive change in water volume; and swelling of sludge (micro organisms, matrix)	Readjust MLSS; stabilize incoming water flow; and prevent swelling of sludge

(4) Much talks has occurred concerning the design of a system that self learns and self organizes things, but for the time being, little progress is really expected. So, instead of talk, we'd better direct more effort to assessment of skilled operator's actions. Copying his action pattern should come first.

(5) Systems coded in symbolic processing languages may be better translated into more practical C or FORTRAN. Also, parallel reasoning methods should be introduced.

(6) A user interface that is intuitive and natural for the user should be actively sought. It should incorporate audio-visual and graphics capabilities.

5. Conclusion

We are witnessing major advances in the design of support systems for operation of sewage treatment plants. Not only the AI engineering, but also fruits of system engineering and pattern recognition technology are being incorporated into expert systems. Soon, these systems will be able to respond intelligently, mimicking human cognition. An expert system will certainly be identified as a [decisive] tool for understanding the subject process.

In this article, I discussed the requirements of water treatment technology, the present status and problems as well as future issues of expert systems for plant operation.

Understandably, preparation of decision tables and trees for the treatment process is laborous and time consuming. But it has to be done, and at this stage, the only way to collect the right information and enrich the knowledge base is to sit with plant operators and quality control staff and conduct painstaking surveys with them. Soon, however, the issues of improving knowledge base editing facilities, and adding functions for users to append and revise the knowledge base themselves will be most important.

Trends in Production Process Technology

Advances in CIM and Technology of CAD/CAM

43064080a Tokyo KIKAI TO KOGU in Japanese Nov 88 pp 2-7

[Article by Kazuo Yamazaki, Mechatronics-CAD/CAM Lab, Toyohashi University of Technology]

[Text] In the midst of today's turbulent business environment, the ability to identify market needs accurately and quickly, launch timely product development efforts, and economically manufacture and distribute quality products is basic to a company's ability to sustain a competitive edge in the manufacturing business. Consider also the socioeconomic situation of our time; the trend toward declining numbers of workers in the manufacturing industry continues as the labor force continues to be drawn to the service industry. As the standard of living improves, and people become more affluent, they demand a wider array of products and services to satisfy their individual wants and needs. Under these circumstances, manufacturering concerns have no choice but to find ways to efficiently develop, produce and supply far greater ranges of products and to quickly adapt their manufacturing structure to rapidly changing consumer demand. In order to perform these quick-tempo corporate operations effectively, companies must first establish an internal structure conducive to a rapid assessment of internal and external conditions.

Past advances made in the field of factory automation were generally based on the idea of improved efficiency for each manufacturing task. Today, however, this type of per-task automation is being superseded by the far more extensive concept of comprehensive automation, which seeks to integrate [the entire organizational system]. As this type of organizational consolidation gets underway and its full effect becomes apparent, the automation we will be talking about will no longer be limited to the individual tasks of factory production, but will affect all essential components of the manufacturing business--research and development, production, marketing and sales.

The means of integration takes the form of an information network. Its purpose is to provide management with information and data to manage the physical resources (i.e., production facilities, equipment for research and development, support systems for marketing and sales), and to provide information links between these divisions in order to achieve better organizational coordination. In recent years, this networking has been pushed by manufacturers to a level where most of their resources are now linked to

computers in one way or another. With such a computerized network, the distribution of information can be achieved at a speed and level of throughput that would be unimaginable in terms of the human effort required. The idea of networked resources came into being in the late 60's when manufacturers started looking to computers to gain better control of the manufacturing process. In the late 70's, the concept was developed into a plant integrating system called computer integrated manufacturing (CIM). What this CIM aimed at is, simply put, the acceleration of plant operations; it processes production-related information, and, based on the results, it monitors and controls the manufaturing facilites without human intervention/involvement as much as possible. While conventional office automation systems merely process information and route the results as output, which is again a form of information, CIM seeks to control plant equipment using this processed data. The essential difference between the two, therefore, is that the latter has to incorporate the technology for converting data into mechanical action.

1. Developments in CIM

Computer automation was first applied to tasks where a computer link meant immediate benefit. These were usually sales and marketing related data processing tasks. Let's stay away from these business-oriented tasks, and concentrate on manufacturing automation. Manufacturing operations basically involve two main task groupings: product design and development, and production. These task groupings, under plant organization, consist of stages of research and development, design, production control, production scheduling, production, and plant management. In the past, when a computer was introduced to provid support for these tasks, it resulted in the creation of separate task-oriented dedicated computer support systems. These customized systems grew independent of each other with their spheres of operation limited to within the task-defined bounds. Of these support systems, the ones that were introduced earliest and whose use became most widespread are the CAD (computer aided design) and CAM (computer aided manufacturing) systems.

Figure 1 illustrates the functions of CIM and the indispensible components that make up the system (those tasks concerned with sales and marketing are omitted). In the figure, the outermost shell contains the physical devices used in CIM, the inner shell contains the [software] systems, the next one contains the classified list of functions, and the center one contains the core technologies. The category of core technologies refers to the networking system and the relational database system. The former interconnects the various independent devices and systems that have been separately developed and put to use, while the latter centrally supervises the information resources, and on demand, provides requested information. Simply linking the numerous systems indicated in the figure alone does not create CIM. Critical to the organization of CIM is the "life-like" networking of the various components. By "life-like" networking, I mean a provision of well-attended two-way communications between different device and system units.

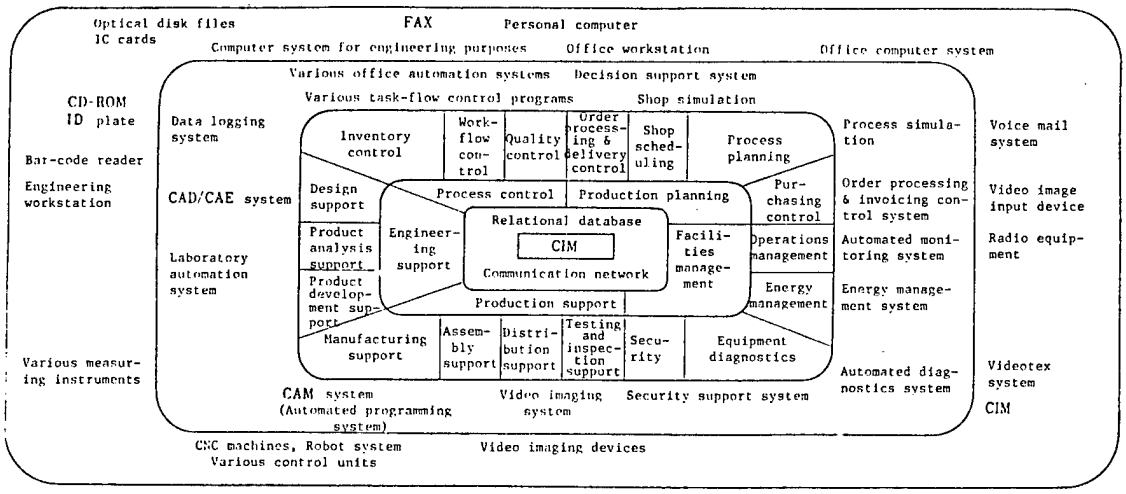


Figure 1. Functions of CIM, and its hardware/software configuration

CIM, as a computer-integrated system, coordinates tasks ranging from order processing to delivery while preparing vital information using computer facilities.

Prof. Spur of West Germany's Berlin Institute of Technology defined the functions of CIM according to the following 7 categories:

(1) sales and customer service
(2) planning of production programs
(3) product design and development
(4) design of production stages
(5) allocation of resources (equipment and consumption supplies and materials)
(6) production planning
(7) manufacturing management

In order to perform the above functions and at the same time maximize the processing power of the computer, computer models for generating the required data have to be constructed first. In designing these models, the following three points must be observed:

(1) what is to be made--product model
(2) how many are to be made--scheduling model
(3) how will the product be made--models for production stages and facilities

Computer models designed according to these criteria comprise the basis for the seven services of CIM. Models may be taken advantage of to assess what can be provided; using numerical models, task processes and plant operations can be simulated to identify the best processing conditions and parameters. Also, synchronized with the actual progress of a physical task, a real-time simulation can be conducted running a computer model to achieve dynamic control of plant operation. Such use of computer models can assist management in improving plant operations. The three model criteria and the seven functions are cross examined by relationship in Figure 2.

A company can't introduce CIM into its plant overnight, and expect to be able to take advantage of it. The company has to have a [realistic] vision as to how CIM can help management achieve prosperity for the company. With clear goals in place, a plan for the staged introduction of changes to be implemented at each corporate division must be made, and incorporated into the master plan before construction of CIM can take place. What this means is that a company should take the approach of "planning from the top down, and executing from the bottom up." Later, whenever the introduction of new hardware or software is planned, the company must first make sure such a move will indeed contribute to the upgrading of CIM.

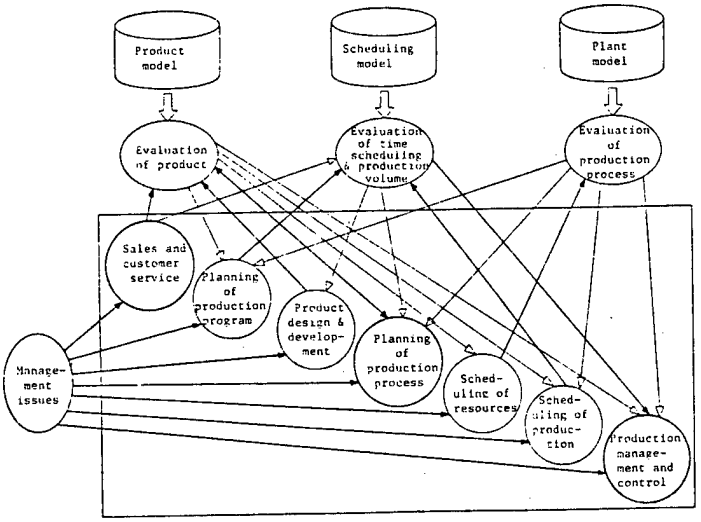


Figure 2. Numerical models for manufacturing automation

The idea of CIM itself was born in the latter half of the 1960's. But back then, system designers failed to clearly recognize the form of technology base that individual systems and devices of CIM had to be built on. Therefore, while advances were made in the applied technology of computer science in the 70's, they did not necessarily lead to computer-based integrated systems. What we did see during that period were [separate] developments in CAD technology for product design, and in CAM for manufacturing needs. Despite the advances made in computer applications, if we look at the results that they had on our current affair with CIM, it is difficult to say that the technology did indeed grow into the truly ideal one we wanted. We have this fear now that we may end up with a revolutionary in concept, but limbo technology because the component systems may turn out to be incompatible for the planned integration of CIM.

Within the context of CIM, the systems of CAD and CAM will play a central role in the vital tasks of manufacturering, i.e., the design of products, and the design of production processes. One of the pressing issues we have at present is to design the right interfaces for consolidating CAD and CAM into the CIM environment, and thus achieve greater productivity in using them. In the following section, I will discuss the problems involved in the interfacing issue as well as recent developments that are taking place in these component technologies.

2. CAD for CIM

It used to be that CAD was regarded more or less as a passive tool for product design. Behind this was the prevalent vision that software developers had which was limited in scope. They wanted programs that could substitute for drafting tables and drafting tools, no more no less. You still had to do the same old drafting tasks on the computer if not on a drafting table. CAD systems that are available off the shelf today share this root. In fact, many of them are merely used to do computerized drafting.

The drafting-type CAD available today obviously cannot fill the needs of CIM. CIM has to do more than just provide a means for preparing drawings for products. It must incorporate the capability for reinterpreting data input during the design stage, and making it available to later stages of plant operations, i.e., production planning, sales and marketing. We need to seriously consider this question: in using CAD under CIM, what is the type of output that the "act of designing" should produce? First of all, the basic design philosophy behind CIM is "to chain link the task processes through data processing (Figure 3)." This indicates that drawings--the goal of conventional design processes--may not be a suitable medium, and that we need to find a better form of design output. Information compiled during the design stage comes in two basic forms: drafting data representing the geometrical construction of the product, and an inventory of materials and components. This information is changed and reused repeatedly during the design processes and thereafter. Therefore, it is essential that this data is stored in such a format that retreival and interpretation at later stages are easily accomplished. Unfortunately, conventional geometric modelling, especially the 2-dimensional lay-out representation made of connecting dots and lines, falls short of delivering the essential.

In response to the abovementioned desired shift in CAD design, a new concept of computer modeling is now emerging, and this is called product modeling. It allows construction of a product model that contains all pertinent information about the product. Such a model may be directly accessed in making evaluations of product durability, designing kinematics simulations, etc. It will be a very effective tool for conducting CAE tasks, i.e., FEM, and solidification simulations. The CAD system will be responsible for constructing such models under CIM. This way, the tasks of preparing information needed for production, and planning the production flow can be greatly streamlined. By entering all available information, vital to later stages of plant operation, into the system at the earliest possible occasion, most subsequent data processing can be automated.

On the other hand, in adopting the idea of product modeling, we need to incorporate information regarding the processing method to be used in production into the modeled product data beforehand, since the same data is supposed to be applied directly at the production stage. However, because the processing method itself is unique and differs from one company to another, the modeling technique has to be unique accordingly. What this means is that having one standard CAD system is not good enough in designing CIM.

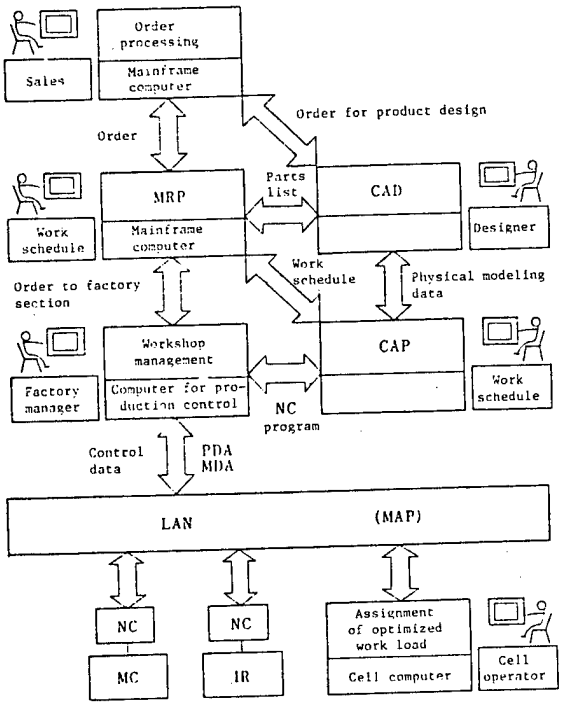


Figure 3. An example of the flow of manufacturing activities under CIM

Because all CAD systems developed previously shared a single purpose, generating output in the form of drawings, they are considered standard systems. After all, drafting is a medium of communication designed for a broad range of people, and there isn't much difference in drafting style and kinds of information represented between drawings by different companies. This general-purpose drafting characterization is precisely the reason that CAD has become popular. But now, under CIM, the CAD system has to assume a very different role, meaning that the previous casting of CAD in a general-purpose mould is no longer valid. Therefore, simply put, developing and distributing an ideal CAD system will not be easy.

3. CAM for CIM

A large segment of the technology called CAD/CAM has been shrouded under a veil. As far as CAD is concerned, since it is designed to produce a general-purpose medium called drawings, it has not been as difficult to create and distribute a CAD system as a CAM system. Since CAM is supposed to digest the data obtained from CAD and generate programs to control the various production flows and manufacturing machines, the whole design of CAM cannot be made according to a standard. On top of this, many previous CAM systems were developed prematurely with the systems engineers having reduced the broad range of manufacturing problems--M in CAM stands for manufacturing after all--to more or less a few simple machining problems. Such a CAM system fails to address the issue of integrating the entire production line since it skips the vital planning stage that follows the design stage and that preceeds the machining process. In fact, it is essential to conduct the following preparatory plans before the transition from product design to part fabrication can be made:

(1) determine what machining tasks to implement, and what machines to use--planning of production processes and manufacturing facilities, and
(2) how to operate the selected machining facilities--planning of production tasks.

With the current-generation CAD/CAM in use, these planning tasks are left to human workers. Recently however, systems developers finally began to recognize the gravity of the above issues--the intermediate steps between CAD and CAM are now given the single broad term of process planning--and are making moves toward designing greater computer automation of these tasks. This newly emerging area of data processing is called CAPP for computer-aided process planning, and is designed to assist in the above tasks, (1) and (2). Task (1) is the narrowly-interpreted planning of production processes, while task (2) is the planning of production tasks, and they are related by the following equation:

Broadly-defined planning of production processes = Narrowly-defined planning of production processes + Planning of production tasks

There is not much commercially available software in the world market, developed with this design philosophy. Of the few that are available, there is the "POPULAR" system developed by Komatsu Software Development Co., Ltd. The processing flow adopted in this program is illustrated in Figure 4. The section for the narrowly-defined production process is semi-automated, meaning that tasks for selecting production processing options, scheduling production processes, and selecting manufacturing machines are done through interactive dialogs between the system operator and the system. On the other hand, the section responsible for task planning is completely automated. As far as the system's processing flow is concerned, the steps involved more or less follow a standard procedure. But, there is room for flexibility in coping with the user's varying needs. The system provides facilities for reading data files and tables so that the user can easily vary the processing parameters to suit his particular needs.

As stressed above, building CAPP into the CAD/CAM system is indispensible for full integration of the latter into the CIM environment. In addition to this requirement, it is necessary to incorporate in the CAD/CAM system another preparatory step in order to realize the ideal CAD/CAM for machine processing. This special step refers to the task of designing and applying jigs and templates. The design of a template cannot start without having already determined the product shape and the materials to be used in it. Also, the details of a production process cannot be determined without having settled on the jigs and templates first. In other words, these tools have to be designed at the very end of the product design, and are needed first in the production of the product. Therefore, unless the task of designing and fabricating the jigs and templates are improved significantly in terms of speed, the design and production of the actual product will not speed up at all. By making special provisions to include the preparation of jigs and templates in the CAD/CAM system, the following advantages may be gained:

(1) Since the information regarding jigs and templates is permanently stored in a type of database, it is easy to utilize previously designed jigs and templates by means of the search function built into the system. The database becomes doubly useful.

(2) It is possible to apply techniques like FEM directly to the product without removing the templates in order to monitor product quality (i.e., durability and resistance to warpage).

(3) It is possible to generate NC programs to control the production process without removing the templates.

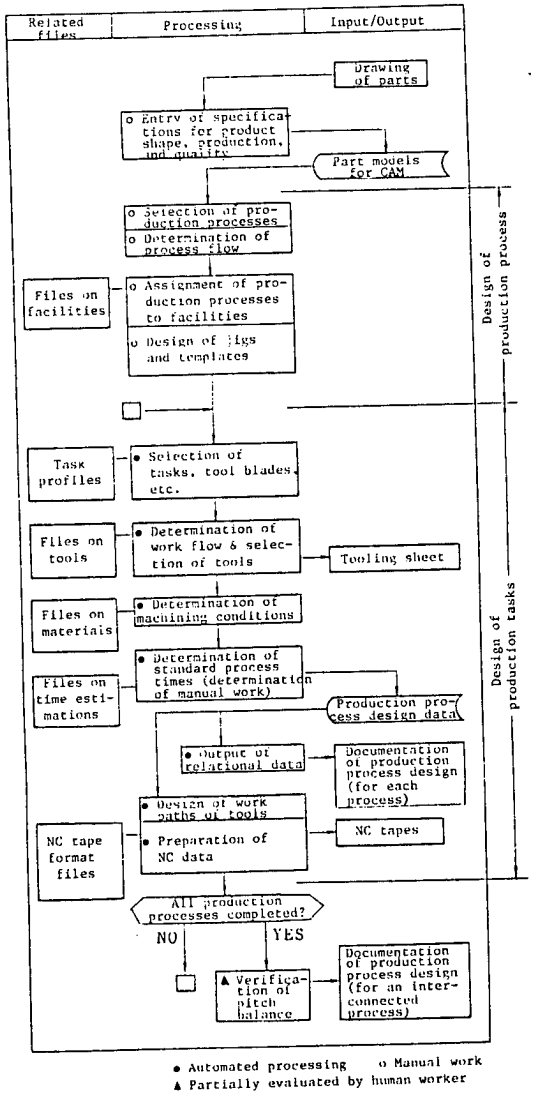


Figure 4. Work flow of a process design system for machining processes

Recent trends in CAM development indicate ongoing active development of online preprocessing techniques, such as production simulation, designed to gain a clearer picture of the production process in advance.

CAM is a program dedicated to the manufacturing aspects [of plant operations] with the aim of maximizing technological and economical gains. But, a quick look at the entire CIM-based monitor/control flow of the total plant floor reveals that the actual planning production processes and manufacturing tasks form only a very small portion of the total, suggesting the tremendously extensive territory of CIM. So, needless to say, CIM is not one of those casual things that can be constructed overnight.

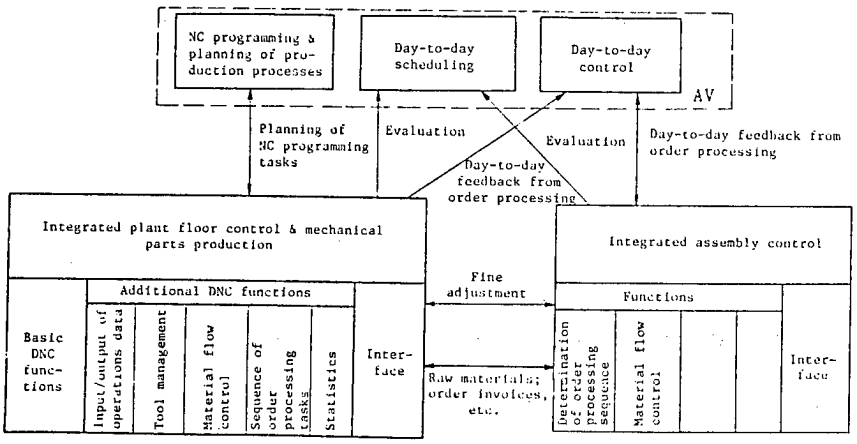


Figure 5. Integration of the control of plant floor and order processing

Once we stop thinking of CAD in the conventional CAD/CAM context, and start designing it as a system for product modeling for the sake of smooth integration into the CIM environment, we will be able to achieve more productive use of information resources by sharing them between different levels of production processes.

As illustrated in Figure 1, the technology of CIM is very broad, consisting of many conventionally available systems and devices. Due to insufficient space, I had to limit my discussion to a few issues and topics of CIM. Those who are interested in the technical aspects should consult many references available on the subject.

When viewed from a different angle, CIM may be considered an agent to "integrate information flow and material flow." The technologies that lend physical function to it are of critical importance. These are the relational database and network technologies. In order to fit a relational database into the production site, the system has to be able to handle autonomously distributed [data] and access it at a sufficiently fast speed. Solutions to these problems still need to be worked out. Also, software systems that use this type of relational data base should be conceived of in an open architecture so that they will be flexible in terms of data base

compatibility. As for networking technology, there is MAP (a communication protocol for manufacturing automation) which has been receiving world-wide attention. According to many experts, however, the MAP protocol is too cumbersome to operate for real-time control of manufacturing equipment on the plant floor. Nowadays, to make up for MAP's big disadvantage, systems engineers are introducing a simpler protocol called mini MAP.

Even without referring to our discussion of CAD/CAM in the previous section, it is obvious that software for the CIM system, for one company or even for a different division of that company, needs to be custom made to reflect the differingt styles and methods of production that the company or the division is pursuing. The development and maintenance of such a system cannot be fully entrusted to an outside developer. The end user company itself must work toward self-reliance in maintaining its own system by investing more in the training of CIM-dedicated engineers.

For production of a product, it is normal to place manufacturing engineers dedicated to the design of production flow and the control of manufacturing equipment beside the assembly line workers who actually manufacture the product. On the other hand, the task of design does not involve such a special staff. This of course is due to the use of primitive tools like rulers and compasses. It is totally unnecessary to explain how they are used. Now with the introduction of CAD, the situation is different. A serious system can take an enormous investment on the order of Y10,000,000 to Y100,000,000. If different users stick to their own individual rules in using such a system, productivity can never be improved, and the money invested will be wasted. It is therefore essential to set up a dedicated support staff, "technical support for product design", which can provide technical information to product designers for making the most use of the CAD system. With CIM moving into plant control scenes today, more and more corporate divisions are directing a greater number of workers to computer terminals, making it increasingly important to launch a similar technical-support program.

When it comes to things related to the "computer", there are still many companies which immediately count on workers in their "data processing divisions" who are only knowledgeable about computers up to the level of preparing management information. Obviously, to make full use of fast-changing computer technology, it is vital for a company to have computer engineers knowledgeable in both the computer itself and the field of application, and to have a provision allowing them to frequently walk through the corporate divisions.

CIM, considered to hold the key to corporate success, must be made to order, and will be different from one company to another. Therefore, the basic issue surrounding the introduction of CIM is whether or not a company can take advantage of standardized tools such as the computer, and construct its own version of CIM for its manufacturing needs. It won't be long before the day when the first corporate division called the "CIM Technical Department" will be launched.

Introduction of Expert System Into Machine Processing

43064080b Tokyo KIKAI TO KOGU in Japanese Nov 88 pp 8-11, 20

[Article by Toru Ihara, College of Engineering, Chuo University]

[Text] Let's suppose that we compare mass production technology to the food service industry. Then, using an analogy, we may liken multi-product low-volume production to family restaurant services. As the standard of living improves, it is quite natural that more and more people become tired of cafeteria food and the set dinners menus offered by local suburban family restaurants. As if to fulfill the yearning of such people for a la carte dishes, manufacturers are now placing greater emphasis on production of made-to-order products.

This shift in purchasing habits has led to the recent debut of the twenty-first century manufacturing system called customized manufacturing, which is geared toward the production of a near-individualized product. This new manufacturing orientation is prompting companies to initiate drastic restructuring measures. One of the issues addressed by these measures is the urgent need to redesign the management of decision making for the core tasks of a manufacturing process, namely, the machining tasks. Machining is essential in a manufacturing process, since it is at this stage that the target product is given its substance--turned, milled, drilled, etc. for precision and quality.

The machining process involves three major aspects: the processing machines, the tool attachments, and the machining procedure. In this article, we will limit our discussion to one of these, namely, the machining procedure, with a special focus on the introduction of the expert system (ES) into the machining environment. First, let's survey the current status and future perspectives regarding machining technology, and see how the external and internal factors as well as human factors are affecting its evolution. Please be aware that this report duplicates parts of some of my previously published articles.

1. The need for the Introduction of ES in Machining Processing

As mentioned earlier, the market environment (external factors) is undergoing a change, and manufacturers are searching for ways to streamline their machining tasks. How do these external factors relate to the manufacturer's decision to introduce an expert system into the machining environment? This is illustrated in Figure 1.

Let me elaborate on this question. The "customized product", referred to in the figure, is considered a semi-individualized product, consisting of parts shared by similar products and parts made to the customer's order. From the production point of view, this means giving up conventional mass production. Since each assembled product may have to be made differently, i.e., different materials and parts, and different shapes, the machining tools have to be quickly readjusted for each production run. In addition, the conventional wisdom of making repeated pre-production tests in advance to determine the

best production conditions cannot be resorted to since the manufacturer is not making mass-produced goods. Publicly accessible databases cannot be relied upon either, since they are not usually designed to provide information on details concerning the machining modes and the new materials that each assembly run requires. Furthermore, the manufacturer needs to consider the possibility of having to cope with fast developing situations where the previously set productivity-boosting technique and task evaluation standard (i.e., the cramping method for multiple parts, and the synchronization of task cycles) for an integrated machining environment no longer holds. These considerations alone indicate that machining processes based on conventional mass production technology are not the way to go. What's needed here is a system which can quickly assess the rapidly changing environment, and perform dynamic management of the machining tasks.

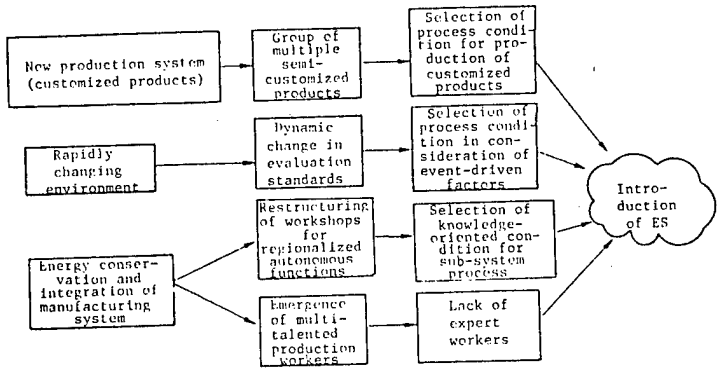


Figure 1. External factors to be considered in introducing an expert system

Manufacturers are introducing computer-based integrated management systems into manufacturing operations. Their primary goal is to improve energy conservation and quality control, and eliminate delivery delays. The introduction of computer-based management has been a blow to some workers, because of the resulting lay offs. The remaining experienced engineers, now fewer in number, have to digest a greater work load and have to attend to many more machines and perform more manufacturing support tasks. Despite these negative side effects, the trend of phasing in computer-based manufacturing management will be here to stay. In fact, it is very likely that manufacturing management technology will even engulf the corporation's financial management, and sales and marketing divisions as well, and grow into a company-wide system.

Let's suppose we have our own company, and it's introducing a computer-based integrated manufacturing system. It is more natural to set up a separate management facility for each level of our corporate operations, i.e., administrative tasks, plant operations, and workshop tasks. Also, we need to organize all the divisions in such a way that each one of them can carry out the assigned tasks on its own. For simplicity's sake, let's limit our discussion to our machining workshop: Each task unit--call it a cell--has to be able to design the machining algorithm, optimized for the available tools and materials, without waiting for commands from cells at higher levels. Our cells will execute independently defined, yet complex machining processes in a parallel fashion.

This is possible only if the system as a whole assumes a hierarchical character with its cell units being regionalized and made independent of each other. In order to achieve such a system, say, for machining needs, we need to develop an automated decision-making system based on artificial intelligence. Such an automating system will prove to be a powerful tool because it can boost productivity significantly by taking over management of complicated but routine jobs like outer-edge lathe turning, where the same results can be accomplished using one of 10 different methods, thus freeing the hands of engineers and technicians for more demanding tasks.

2. Current Status of ES Introduction

The general flow of machining the process is depicted in Figure 2. This does not cover cases where real-time machining tests can be conducted to obtain optimizing conditions. Let's follow the outlined steps: (1) Based on initial estimates, the machining parameters are chosen, and the algorithm determined. The task begins, and (2) the progress is monitored. (3) The collected data is analyzed for ways to improve task output. The processing parameters are readjusted accordingly. (4) If trouble arises, corrective measures are administered, and the processing parameters are again adjusted. Once the task is completed, (5) the operation profile is stored in the database by recording the processing conditioning data, and the profile is reviewed. (6) The results of the machining operation are passed on to upper-stream [divisions of the plant floor] to insure maximum plant-wide coordination.

For each of these six steps, an expert system may be applied. But, because of the constraints imposed by the particular model of computer in use, and because of the limitations in our artificial intelligence-processing technology, we have yet to see actual incorporation of an expert system in all of the abovementioned steps. At present, only some of these tasks, plus a couple of other tasks which used to be considered very difficult to design, such as the "clamping" and "readjustment of the material-specific machining contour", are starting to attract serious consideration for the application of expert systems.

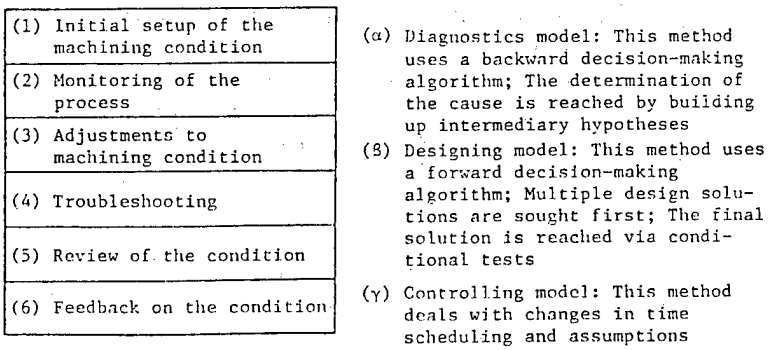


Figure 2. Problems to be considered in determining machining condition

We can see how the above mentioned decision-making steps may be grouped together. Let us follow the classification method used in AI technology. Roughly speaking, task (1) falls under (beta), (2) and (3) under (gamma), and (4) under (alpha) (please refer to the explanations of (alpha), (beta), and (gamma) in Figure 2). In real practice, an AI-based hybrid processing approach is often adopted; Techniques that are more suitable for solving problems, like "structured intelligence" and the "breakup of partial problems", are used in combination with the primary decision-making facility.

Aside from limited development in problem management and fuzzy decision-making techniques, corporations have not shown a strong interest in using an expert system to raise the level of machining processes. In the following section, I will discuss developments in expert systems pursued by the Research Group for Machining Expert Systems, Society of Precision Engineering. There, researchers designed prototype standard-setting expert systems for the management of tooling, monitoring, and troubleshooting processes.

(1) ES for Tooling

In the ES prototype developed for drilling work, a production system was used. Comparison of the decision-making algorithm of this expert system with the decision-making pattern of an experienced engineer indicated that the developed system was very sophisticated indeed. In fact, it even went so far as to review existing decision-making rules and added missing ones to the list. The rules used consisted of the specifications for the contour machining tasks and descriptions of the tools used. The decision-making algorithm was built around a proof-constructing scheme whereby the right combination of rules was sought. The basic rules adopted by the system, which formed the fundamental decision-making criteria that the system followed, were the specifications for the product's material supplies, machining tools, and machining shape profiles. The system possessed a database facility to store the results of analysis. The analysis was made based on a diagram of characterization factors using the critical data generated from the task control conditions, i.e., parameters used to define machining conditions, and data for the finishing allowance.

The prototype expert systems for the lathe turning workshop and the machining center both incorporated the black board technique. The questionnaire survey conducted when the development project began revealed a couple of important generic facts. First, depending on which machining method was chosen for the securing of the target material, the design of the subsequent tooling procedures had to be significantly altered. Secondly, the decision making patterns employed by the systems in breaking up the target task shared common traits.

The algorithm utilized in these expert systems is as follows: Selection of the machining process (or selection of the method for holding down the material) --> Breakup of the process into sub processes (or assignment of the processing sequence) --> and Selection of the tools. Decision making is based on a forward search scheme whereby the rules for the subsequent subprocess are determined according to the characterization that the system has identified [in the previous subprocess]. For example, in defining the processing sequence

for a turning job, e.g., a chucking task, the system first makes note of the unique features of the part, i.e., contour, maximum diameter, and L/D, and then sets the rules for specific task identification. In other words, the system looks for distinguishing features in the target material, searches for applicable rules which match the resultant character, and defines the processing sequence.

Many of the ES-based tooling systems we have today take the abovementioned approach. Let me reiterate. The task setup parameters for subprocess selection, machining specifications, etc., are declared; the decision-making rules are set; machining progress is monitored, and readjustments to the configuration parameters are made to improve performance. To be totally practical, however, the system has to go a step further so that more readily understandable forms of information--graphs and tables--can be directly input, and readjustment of the system can be easily done from the factory floor.

(2) ES for Monitoring and Production Support

These expert systems are designed to provide real-time assistance to ongoing management of a machining task (excluding management of troubleshooting). First, the system collects data from the monitoring of machining progress. It makes a cross reference between the current data and data previously stored in its data bank. If there are deviations, the system generates a prescription for readjustment. Also, it predicts future development of the ongoing task, and generates controlling prescriptions. Both prototype expert systems were designed as production systems. The monitoring ES was specifically designed for diagnosing the operating conditions of the tools. It works by comparing the torque pattern in the working tool and the stored predicted pattern. Please refer to another article in this series for an extensive discussion of this topic. The production support ES on the other hand is only a proposal at this stage, but will be an extensive system incorporating an automated production system and robots. It detects the [target material's] resistance to machining, and readjusts the feeding and machining speeds. As for in-process detection of a tool's wear, the system bases this on examination of the cutting force-component ratio.

The task of monitoring is regarded as the job of recording as "facts" the previously unknown quantities of the ongoing processing operation, which are collected through the system's sensors. This sensing and recording capability represents a very important technology whose development is critical in the design of an autonomous [monitoring] system.

The simplest system may follow "a method whereby the monitored output power of the motor is converted to machining power P, and this value is chronologically recorded as changing 'facts'." Since even under this simple method, comprehension of some sort on the part of the system is still needed, the algorithm required will have to be very complicated. In any sizable program, this always presents a problem when it comes to chronologically synchronizing [a system's subroutines].

As for processing control, techniques along the line of the adaptive control or optimizing control method are being considered. In particular, engineers

are experimenting with an adaptive control technique that relies on heavier use of discrete quantities and vaguely defined conditionals as control parameters. So far, they have not been able to control time-varying conditions. The difficulty in doing so lies in the fact that we still have terribly imcomplete ideas as to how to achieve system stability, throughput, and real-time controllability. At this stage, we are basically limited to research in ecclectical systems based on procedural programs. There is another hopeful technique. This is fuzzy control, which according to Sugano may be considered an expert system. If my understanding of it is correct, as soon as the algorithm for identifying meaningful phenomenon is designed, this system will prove to be very effective when applied to monitoring control.

(3) ES for Troubleshooting

The system developed was designed to locate errors in the lathing operation, and generate measures to eliminate them. This AI system utilized [condition] framing as the basis for determining its decision moves. It identifies the condition of the machining progress by first cross referencing the monitored values to the attribute values stored under the frame defined for the target material, tool, and machining specifications. Then, based on the compiled data that characterizes the machining condition, it performs an analysis to determine if an error has occured. The program jumps to a frame whose attributes match the error profile. It completes the intra-frame processing there, and based on the result, branches again. This pattern is repeated until a solution is found. In this algorithm, the point is to reduce the physically measured error quantities, i.e., resistance to the machining motion, and temperature of the machined material, to generic characterizations so that a solution can be found using sets of universal rules.

Because of this universal characterization scheme--specifically, a scheme for (m,n) evaluation--the system can identify accompanying errors and predict an emminent error. This capability should be contrasted to the limited capabilities of conventional error diagnostic techniques.

As far as error diagnosis is concerned, even a very rudimentary expression of error can be used to conduct a quite effective error diagnosis. For example, a simple rule like {IF (nose radius, R, increase) (rake angle, (alpha), decrease).....--> THEN (breakdown in tool, Pr, decrease)} may even be used.

Of course, this ease in determining a corrective measure may backfire. Because once the system fails to see the fundamental cause that led to the error, no matter what measure is generated, it will be meaningless. Because of this misgiving, when the abovementioned prototype ES was put together, there was an attempt to structure its AI resources. The system's intelligence was divided into material type-based categories (i.e., workshop machines, tools, target materials, and machining lubricants) and task-based categories. It was also necessary to accompany this with another AI-partitioning scheme, a phenomenologically defined method that's based on generic paramters associated with any machining process (i.e., machining power, machining temperature, and chattering [of the machine]). By incorporating the second scheme, the system can easily make readjustments to interfering isolated problem areas, and cross examine different stages of decision making.

3. Developments in Machining Theory

In this section, we will discuss developments that are taking place in machining theory. Remember these advances comprise the internal factors that are prompting ES development. We shall omit a survey of externally driven future developments since these can be readily discerned from preceeding discussions.

Machining theory encompasses the machining mechanism, the machining paramters (i.e., resistance, and temperature), the conditions (i.e., machinability, tool wear, tool breakdown, and chattering vibration), and quality control (i.e., machined precision, quality of the finish, and machining cost). Research in each of these areas has been active, and many techniques for evaluation and simulation of complex configurations have been made available.

This sounds encouraging, but actually, these techniques are not as useful as they could be, because most of the research work focused on subjects with a limited application scope such as narrowly defined tasks, and mass production-based tasks. Also, recent research efforts have been geared toward rapidly developing new techniques and new materials. This situation has led to a rather helter skelter development. This is why we are still plagued with numerous problems when it comes to the simple task of setting the same machining parameters for varying production situations. We simply lack a standard yardstick to quickly assign these data.

The situation is like a bunch of remote islands scattered all over the ocean. There are hodgepodge theories, but they are not systematically connected. So, what has been done is to let the brains of experienced technicians bridge these remote islands.

Figure 3 illustrates how this is done. In the figure, the horizontal space is spanned by the control parameters (the material's machinability/tool's physicality, the [number/types of] tools and attachments, and the machining condition). The scale running perpendicular to this is used to read the degree of machining quality, or the evaluated levels of machined precision, finishing, and machining cost. In setting the state of the machining process as depicted in the figure, I assumed that a successful translation of the quality of the functions of the machined parts into [quantifiable] problems--quantified with respect to a standard precision and so on--was made before they could be evaluated. As you can see from the figure, the machining quality versus control factors forms a multi-shaped relationship.

So, we may say that when we specify a task condition using a theory of machining, we are in effect analyzing a particular portion (static and dynamic) of the abovementioned multi-shaped functional relationship.

As advances in ES are made, more previously unseen portions of the functional relationship will be visible. Through efforts to analyze the function's properties, we may be able to discover discriminating micro quantities. Once a few universal quantities are found, our knowledge of machining engineering will be catapulted to full-fledged machining theory.

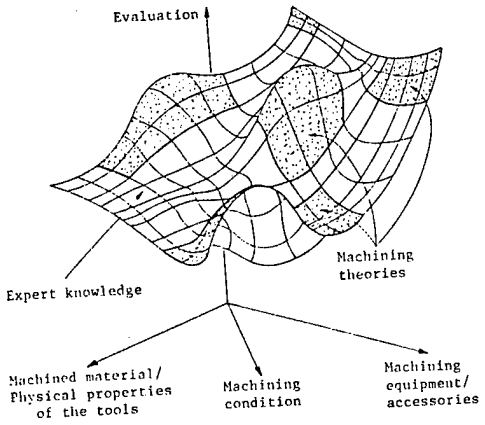


Figure 3. Parameters associated with the setup of a machining work condition

4. Future Perspective and Summary

In this last section, we will survey the issues and topics of R&D of expert systems, which are related to the consideration of human factors. Figure 4 shows a diagram relating the aforementioned factors to successful development of ES.

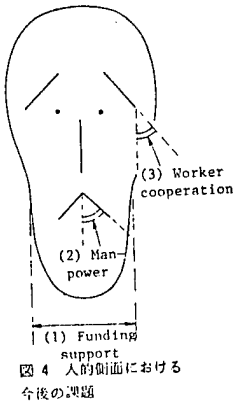


図 4 人的側面における今後の課題

Figure 4. Future issues likened to facial features

First of all, we have the problem of funding. We are on a low budget. Secondly, there aren't enough people working on the development of expert systems. Among many problems that led to this situation, there is the strong adherence on the part of manufacturers to the unyieldingly perfectionistic attitude toward the idea of production planning. It's hard to blame them for this though. After all, they had to do thorough planning before any successful mass production system could be initiated. Also, the level of interest shown by the industry is not sufficiently high. To illustrate the point, aspiring

systems developers are not given enough occasions to publicly present their research results. Outstanding work can end up being wasted. Even if the developed system is a prototype, as long as it shows some creative features, it should receive more attention.

The third problem regards the difficulties often experienced by systems developers in obtaining coorperation from experienced technicians. One of the causes of this concerns the architecture around which a manufacturing system is constructed. Systems we have today were designed on the assumption of one-dimensional processing of information. Workshop technicians become reluctant to lend a hand since, with information arranged one-dimensionally, it is quite easy for it to leak out. Also, this scheme makes it harder for them to make use of the ES-connected machines without running the system. What measures can we take to alleviate this situation? We can break up the AI database functions, and adopt a more regionalized scheme whereby each subsystem is made more autonomous in terms of task territory. Also, we can adopt an object-oriented architecture in the system so that we can keep in a single [data] frame all the pertinent information for the machining procedure, the profiles of the machining equipment, and quality control. With these schemes in place, knowledge coming from outside of one subsystem can be effectively blocked [from reaching other subsystems]. No effective systems have been developed yet. But difficulties are to be expected, since running autonomous task units while ensuring good coordination between them is not one of the easiest things in the world.

As for human factor type issues, as is evident from the figure, the level of positive contribution to any one of them is low. We have to invent ways to rectify this situation, and make the expression on the face change.

Finally, I would like to express my gratitude to the chairman and many members of the Research Group for Machining Expert Systems, Society of Precision Engineering for allowing me to cite their valuable research works and stimulating discussions.

Toward Advanced Monitoring Technology for Machining Processes

43064080c Tokyo KIKAI TO KOGU in Japanese Nov 88 pp 12-15

[Article by Shozo Takada, Mechanical Engineering, Department of Engineering, Toyo University]

[Excerpts] Driven by the recent shift in consumer purchasing patterns toward products that are different, manufacturers today are increasingly pressed to diversify their product lines to appeal to the new generation of consumers. To come up with such high value-added products, companies first need a production technology that enables them to make low-volume products at a competitive cost and still achieve good quality control. No longer, can manufacturers expect to have an assembly line unattended and automated, and safe and sound before any actual production takes place. They can only do this for mass-production products. And they accomplished this by ensuring that everything which had anything to do with the planned production was put under tight control, by

running test after test on the system while it was in the making, making adjustments here and there until the system became flawless. Now a totally different method is being called for. This new technology must be able to keep track of the ongoing process, quickly detect errors, and rapidly administer remedies.

Unfortunately though, our present level of technology does not allow us to design production monitoring systems powerful enough to satisfy these demands. So, we know we have made a big head way. In the following sections, I will describe the R&D direction that research people in manufacturing technology are taking today.

1. Present Status of Machining Process Monitoring Technology and Beyond

1.1 Monitoring Technology

There are as many types of errors that can plague a machining process as there are factors that control the system. To make sure they are caught in time, intensive research and development efforts are underway to develop an equally extensive array of systems and devices. Some of them are being given more serious attention than others. Number one priority is given to systems for monitoring the wear and chipping of tools. Systems for tracking chip disposal and the chattering level of the machine come second.

Numerous types of monitoring schemes have been tried, taking measurements of the cutting strength, vibration intensity, AE, temperature, thermo-electricity, resistance, reflected light intensity, video recorded image, etc. All these techniques fall under one of the following three general categories:

(1) sensor/detector technology,
(2) signal processing technology, and
(3) evaluation technology.

The central problem of the detector technology regards development of a sensor system for effective conversion of a physical event that takes place during the machining session to a quantified signal. Of course, the detected signal may not directly reflect the event, or it may even be contaminated by noise. So, we next need a signal processing technology to track the right information. Then, finally, we need a decision-making system that can evaluate the state of the machining system, and if an error or malfunction is detected, can prescribe what course of action to take based on the processed information. Depending on how the total monitoring system is designed, the same information can lead to different, or even undesirable action.

1.2 Beyond the Present State of Technology

Sadly, the level of technology we have in this area is less than desirable, although we know that plenty of R&D effort has been invested in many different types of techniques. The people who have the say in real-world system development simply aren't using these techniques. Why? There are a couple of reasons:

(1) Lack of sufficiently cheap and reliable sensor devices, which are unobtrusive on the work floor
(2) Lack of highly reliable signal-processing and evaluation techniques that are effective for wide ranges of changing events due to variations in the monitored parameters such as machining devices, tools, and machining conditions

In the meantime, the same systems development people are sending in more requests to their research associates, asking for superior monitoring functionality because of the need to create an unmanned automated system. In the event of a malfunction, it is not as desirable to shut down permanently the whole machining operation as it is to recover from the incident by, say, replacing the failed tool. But this extra service does not come easily. The monitoring system would have to possess the capacity for special decision-making not only to detect the event, but to judge the type and extent of the malfunction and to quickly administer corrective measures. Also, to cope with the wide range of errors that can occur within a given machining process, the monitoring system has to be have an equally extensive monitoring capability.

Another important issue with regard to making improvements in the technology of task watching is early detection. The earlier it is detected the better, if it is detected at all, and if an accident forecast can be made, that's all the better. Imagine the advantage you would have if you have a monitoring system capable of warning you of an eminent, and usually sudden breakdown of a tool.

All these problem points have to be deeply embedded into the minds of the R&D people. Many advances are needed in all areas of sensor/detector technology, signal processing, and task-evaluator systems. Let me introduce R&D developments that are beginning to take place.

(1) Sensor Technology

(1) Research in Advanced Signal Detection Methods

Not so long ago, a popular monitoring scheme was the one based on detection of load current in a spindle motor. In the past few years, however, engineers began to favor AE-based monitoring schemes. Of course, they did not neglect the many other methods, which are based on light detection, ultrasonic signals, and 20k-80kHz vibration intensity (the frequency range sits somewhere between that of the accelerometer and the AE sensor). Most of these techniques have been around for some time. They were simply not put to active use until now. Because of the recent advances in signal processing technology, many engineers are going back to these old inventions and are seeing new potential for breakthroughs. Regardless of the type of sensor device that is adopted, unless the signal is freed of heavy noise and rich in information right from the source of the event, the whole business of monitoring serves no end. The information that was not picked up can never be regenerated, no matter how sophisticated the rest of the system is. More powerful sensing devices with greater sensitivity are strongly needed.

(2) Development of High-Reliability Sensors

As I noted already, one of the biggest bottlenecks hampering development of a practical monitoring technology is the poor sensor technology we have today. One way to improve the sensor system is to design it without any special sensors for tracking the spindle motor's load current, or spindle revolution variation. Anyway, the field is active as far as research is concerned. There have been reports of the development of a torque sensor that is built into a spindle or a motor. Also, advances are being made in cheaper and more compact designs for ITV and lasing components, leading to the active application of these parts into optical sensor devices.

(2) Signal Processing Technology

(1) Application of Digital Signal Processing Technology

Many advances are being made in this area because of the numerous signal processing techniques that have now been made available by recent advances in computer technology for handling digital signals. One of the more famous applications is the tool chipping detection technique based on the use of adaptive filters.

(2) Toward Faster Processing

Recent R&D-related reports indicate that more attention is beginning to be paid to the high frequency spectrum of monitored signal waves, since such close examination should help to extract complete information concerning the signal. The proposed AE-signal frequency analysis is a technique conceived along this line. This extra task does end up complicating the overall mechanics of signal processing for the system. It also slows down the system's processing. And this has led to the search for ways to design a faster machine. Many think that reliance on dedicated DSP's (digital signal processors) will soon be inevitable.

(3) Evaluation Technology

(1) Utilizing Data Prepared in the Production Planning Stage

No monitoring system can determine if an error has occured by basing its decision making process solely on the collected quantities associated with the observable machining parameters. It has to know in advance what is a normal state and what is not. In other words, the system has to be equipped with some form of reference material which it can count on in drawing up a correct evaluation should a suspect event occur. Consider, for example, the task of monitoring the cutting force. In order to judge correctly whether the gathered data represents a normal state or not, you have to know in advance the precise machining specifications like the tools that are being used, the materials being worked on, and the detailed machining conditions. Much of this vital information is generated during the stage of production planning, and depending on how effectively this information can be ported to the monitoring system, the accuracy of the system's evaluation of the task state will be greatly affected. In this respect, there have been proposals for making use of

production simulations [to generate such information]. I will come back to this point later.

(2) Decision Making Based on Models

One may turn to the idea of basing the decision making process on a [numerical] model constructed to represent the manufacturing process in question. For monitoring purposes, techniques based on estimation for on-line parameters may be used. The only problem is how to come up with the right model that fits the profile of the target process. It is not at all easy to prepare a model with a high degree of generic applicability.

(3) Using AI Technology

The idea of using a pattern recognition technique to test the state of an event for abnormality based on observed quantities is not new. But, the concept of pattern recognition is now given a new dimension as people are trying to apply the technology of artificial intelligence to it. AI processing techniques are also favored for designing a technique to maximize the use of the numerous pieces of referece information vital to evaluation of an event/error state. This topic will be discussed again later.

2. AI-Based Monitoring System

2.1 Types of Vital AI Information

As already noted, correct evaluation of an error state requires more than just the raw information gathered from monitoring. It has to be supported by the following types of expert knowledge:

(1) manufacturing information generated during the stage of production planning,
(2) specifications for the mechanics of the target manufacturing process,
(3) specifications for the manufacturing environment including the machining tools, and
(4) expert knowledge of the target process.

This task-specifying information can serve as the criteria on which the monitoring system bases its decision making. First though, this abstract information has to be converted into a concrete quantified format that the system can read. But how? AI researchers say, simulate the manufacturing environment using a computer.

2.2 A Production Simulation System

Conducted in the early stages of machining planning, simulating a manufacturing process on a computer system can assist the manufacturer in two ways: it can help test the planned production settings for integrity, and it can generate default settings of task states for cross reference use, to be accessed by the monitoring system later. Two types of simulations need to be implemented: physical modeling simulation, and processing task simulation

[translator's note: the original text reads "physical simulation"]. The former is used to verify the physical shaping of the product, while the latter is used to predict the state of the processing task such as the change in cutting force, based on the model sculpted by the former.

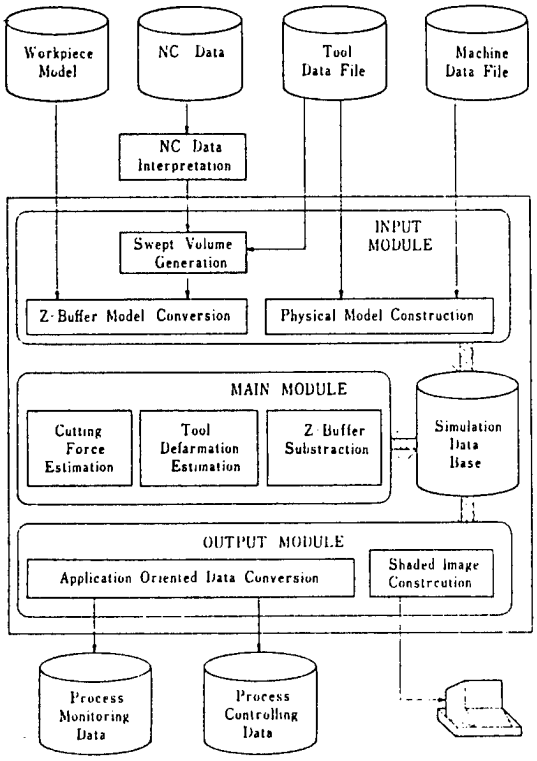
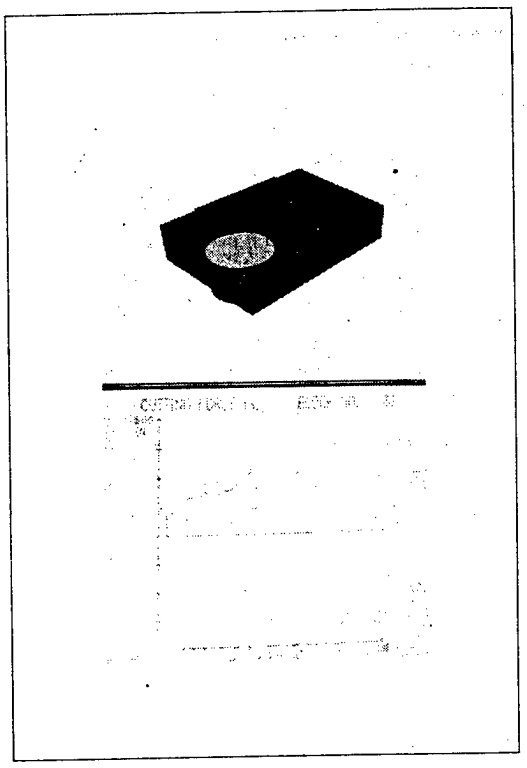


Figure 1. Organization of a machining simulation system



Photograph 1. A display frame from a simulation in progress

Figure 1 depicts an example of how a machining simulation system, capable of predicting the change in cutting force, may be organized. The system is supplied with initial setup information that determines the task configuration, i.e., models for product parts, NC data, specifications for tools and machining processors. Geometrical modeling simulation is carried out with the solid-modeled forms of product parts and tools transformed into Z-buffer data first. This way, the simulation can run at a greater speed. Under the modeling simulation, you can watch the shape of the product change as the modeling progresses on the screen. At the same time, the simulation system monitors the surface being machined by continuously computing the affected surface area. This particular system was equipped with a special facility for estimating the error caused by deformation of the cutting tool. An example of simulated modeling captured on a display is shown in Photograph 1. To the lower side of the display, you can find the plot of the change in cutting strength predicted by the system.

2.3 AI-Based Monitoring System

Running task simulations like the one above, we can predict how machining strength will vary in time. By using the data generated this way and comparing the data obtained through real-time polling, we can determine if the machining processor is operating correctly or not. But, mind you. So there is a deviation in machining behavior, as detected by the monitoring system. But how can the system know that it's indeed a malfunction in the machining processor or not? Before it can actually draw a conclusion, the system has to be capable of estimating the cause that led to the deviation. It can do that provided that strong circumstantial information is made available to it--the behavior during slipage in the monitored value from the predicted value, and the exact condition of the machining process when the error occured--and provided that the system is smart enough to digest all of this input. And this is where an AI system comes into the picture.

A prototype of such a system is illustrated in Figure 2. This system has an extra function, cause estimation, based on AI processing. It is organized with two components: a monitoring unit, and a diagnostics unit.

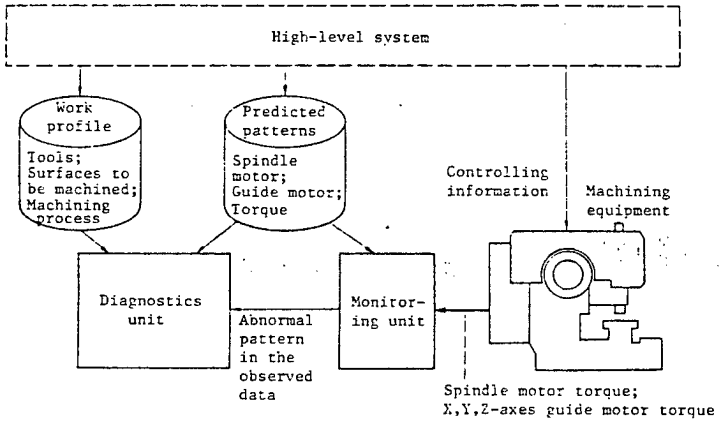


Figure 2. Organization of an expert system for monitoring tasks

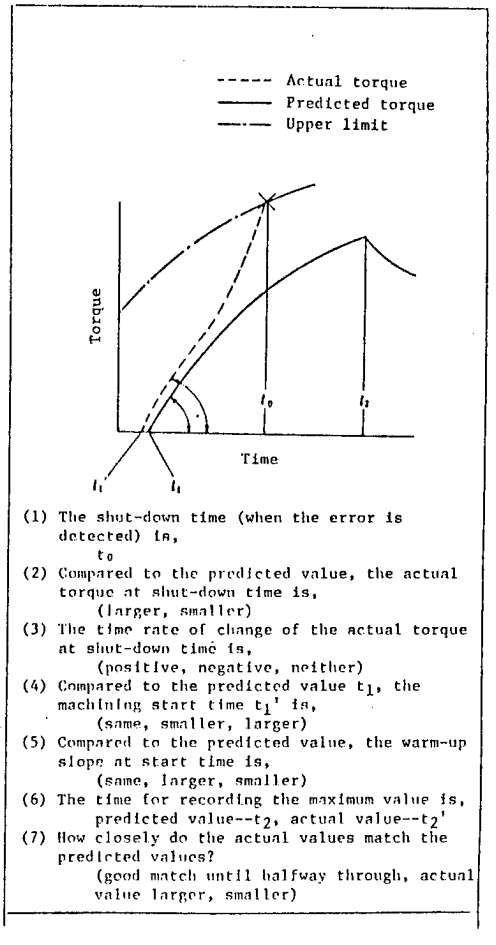


Figure 3. Identification of abnormality of the motor torque pattern

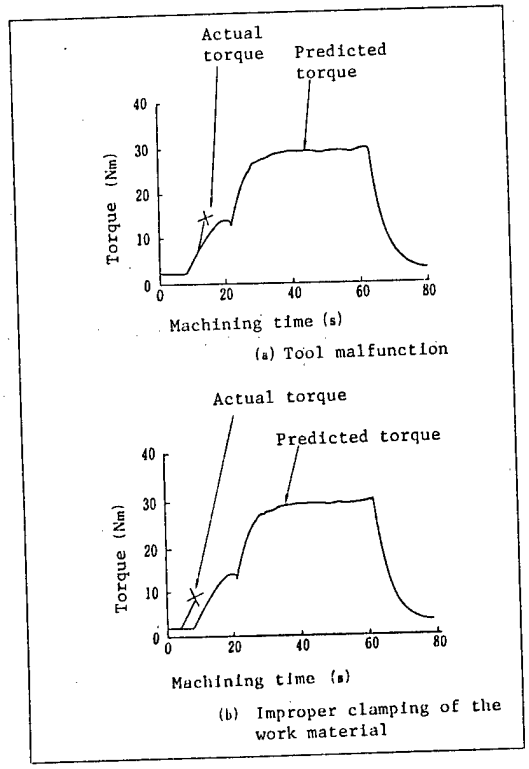


Figure 4. Examples of diagnostics profiles

Table 1. Error Factors That the Prototype System Monitors

Errors in product setup: o Incorrect part assignment o Incorrect positioning o Incomplete clamping o Failure to issue
Errors in machining tool setup: o Incorrect assignment o Incorrect configuration of preset values for tool control o Failure to issue
Defects in tools (chipping, excessive wear)
Errors in machining
Errors in the spindle system
Errors in the guide system: o X-axis guide o Y-axis guide o Z-axis guide

Present Status, Future Perspectives in Abrasive Processing

43064080d Tokyo KIKAI TO KOGU in Japanese Nov 88 pp 16-20

[Article by Toshiji Kurobe, Department of Engineering, Kanazawa University]

[Text] Products of the precision machines and electronics of today contain levels of sophistication in functionality and performance that were unheard of yesterday. From a manufacturing point of view, this advance means a far greater challenge in putting together greater numbers and types of parts and raw materials, many of these being highly resistant to machining. Typical among such materials are the increasing number of super-tough brittle materials or so-called fine ceramics that recent research has been creating. These new breeds of materials have to be ground and resurfaced--lapping and polishing--using special super abrasives.

To achieve high-precision high-efficiency production, suppliers of raw materials, makers of grinding stones, makers of industrial machines, and makers of end products are trying to team up and pool their resources to find the ideal technology to process these troublesome wonder materials. The main issue these people are tackling is how to optimize the size, shape and surface coarseness of the abrasive grains which are determined in accordance to the degree of relative motion between the abrasives and the material being worked on.

In this article, I would like to concentrate on the most recently invented methods of abrasive production.

1. Toward Creation of a Grinding Center

Diamond abrasives are typically used to process tough brittle materials like fine ceramics and superhard alloys. The wheel has to be changed depending on the type of target material to be worked on and the processing conditions prescribed. Professor Nakagawa of Tokyo University's Institute of Industrial Science pointed out specific problems in designing a ceramics machining process for high precision and high efficiency. These are listed in Table 1, and are: (1) how to gain processing efficiency, (2) how to improve the dimensional precision and surface finish, and (3) how to handle complexly-shaped materials.

The term grinding center was coined in a manner similar to machining center. A grinding center houses facilities for heavy-duty machining processes used to reshape ceramics and other hard materials. It seeks to achieve the same level of processing precision and efficiency that is found in a machining center. Creation of such a center requires the development of: (1) suitable processing machines, and (2) extremely tough and long lasting grinding stones.

To qualify for the right grinding stone, the abrasive material must possess the following features: (1) a small amount of wear, comparable to those of cutting tools; (2) a high level of durability, which guarantees a highly efficient grinding operation; (3) long lasting abrasive strength to sustain

prolonged grinding; and (4) a broad profile of effective abrasive strength ranging from low to high grinding speed. Nakagawa et al successfully developed an abrasive material that fulfills these needs. This is the cast-iron-bonded diamond abrasive. Here is how Nakagawa prepared his material:

Diamond abrasive grains are mixed with cast iron powder and (carbonil) iron powder. The mix is given shape, and sintered in anmonium decomposed gases ($3H_2+N_2$) at 1,140°C for 30 minutes (standard). The finished slab is soldered onto a base made of metal to make a grinding stone. Nakagawa's grinding stone has the following advantages: (1) It has an extremely strong abrasive action, and is impervious to mechanical shock. (2) It is highly resistant to blinding, offering good truing and dressing characteristics. (3) It has a low level of wearing and a large grinding[-to-wearing] ratio. (4) The abrasive strength is very good even when used in a slow grinding process.

This wonder grinding stone was fitted not long ago into a machining center for real world testing. According to a recent announcement, the material is performing excellently, as expected. Aside from the grinding stone development scene, there have been proposals for creating hybrid grinding centers by combining an arc-discharger system and an ultrasonic attachment in a grinding center. One way this type of facility might be organized is depicted in Figure 1. Needless to say, to realize such a facility, greater effort on the part of researchers in developing superior abrasive materials and processing machines is needed.

2. Advances in Polishing Technology

Today's sophisticated products require many precision mechanical parts. The level of precision needed to produce these parts must be within the submiron range or less. Then, with some electronic parts, part makers must ensure that the parts are ground and resurfaced to a perfect finish. Not only that, they may have to study crystallography very seriously, and ensure that their products have no surface crystal disorders. These trends indicate that conventional abrasion technology is becoming obsolete, and that there is a need to invent radically different methods, from the ground up, in time to meet the growing demand for high-precision grinding tasks.

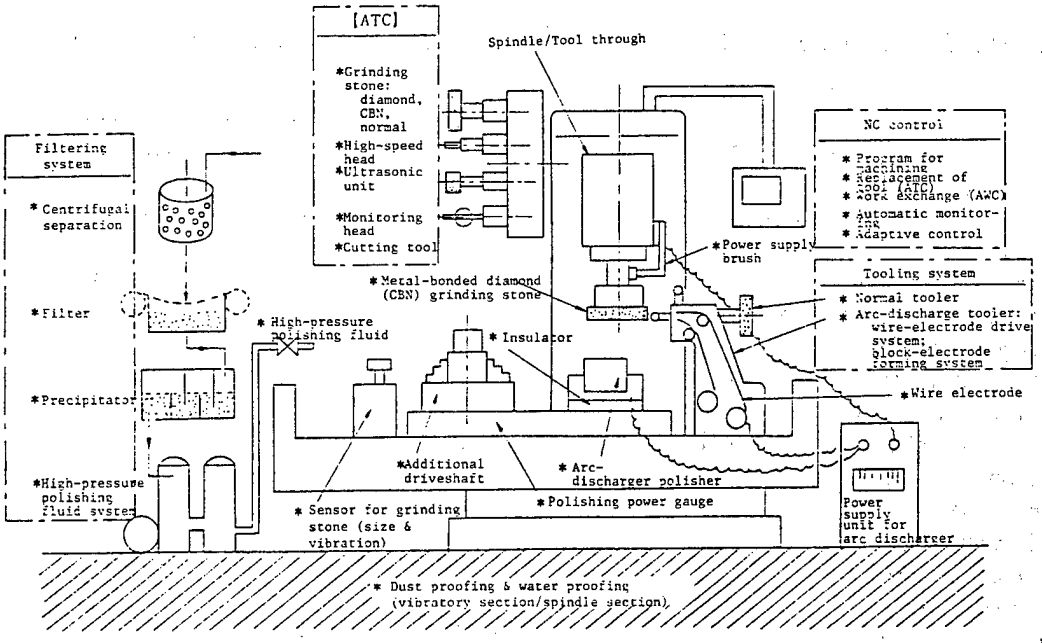


Figure 1. Conceptual design of a multi-function machining center/grinding center (by Suzuki et al.)

Table 2 lists polishing methods developed recently. These techniques are classified into three groups depending on whether or not abrasives are used at all, and on the method of sustaining the abrasive action. Two of the groups defined according to the type of abrasion sustaining method are: (1) processes using spray-dispensed abrasives, and (2) processes using recirculated abrasives [translator's note: the original text reads "free abrasives process"]. The other classification regards some unique processes that do not involve the use of any abrasives. Quite a few of these processing methods have been put to practical use already, although a majority of them are still being worked on in the lab. Research is active, however, and reported developments are encouraging. Next, I would like to discuss recent developments in magnetic polishing technology which many think is coming soon.

Magnetic polishing is a type of pressure controlled process for resurfacing using a special polishing tool. This abrasive tool is highly sensitive to the presence of a magnetic field. It is used on magnetic fluids, magnetic abrasives, and electromagnet materials. At present, three techniques are being researched that incorporate this unique process: (1) robot-controlled magnetic polishing, (2) magnetic polishing using magnetic abrasives, and (3) magnetic polishing using magnetic fluid. The three methods are being developed with different applications in mind.

Table 2. New Polishing Techniques (currently under development)

Type of Method	Type of system	Type of abrasives	Polisher fluid or chamber atmosphere
1) Polishing using magnetic fluid (polishing pressure regulated)	Abrasives are spray-dispensed	o	o
2) Polishing using magnetic fluid (no. of abrasives regulated)		o	Magnetic fluid
3) Polishing using electrophoresis		o	o
4) Chemo-mechanical polishing		o	Chemical solution
5) Mechano-chemical polishing		Soft abrasives	-
6) PMC process		o	Chemical solution
7) Hydration process		-	Heated vapor
8) Non-contamination process		-	Water
9) Hydroplaning process		-	Chemical solution
10) EEM	Recirculated abrasives	o	o
11) Non-solid in-contact polishing		o	o
12) Float polishing		o	o
13) FFF using magnetic fluid		o	Magnetic fluid
14) FFF using electrophoresis		o	o
15) FFF using plasma		o	-
16) Magnetic floata-tion polishing		o	Magnetic fluid
17) Magnetic polishing using magnetic abrasives		Magnetic abrasives	-

<Magnetic Polishing Using a Magnetic Fluid>

S.S. Papel of the United States developed the first magnetic fluid in 1965. It's a colloidal solution that contains a highly magnetic material in a colloidal form (colloid size, 150 angstrom or less). This fluid has the characteristic of being highly sensitive to a magnetic field.

Researchers have taken advantage of this unique feature of the fluid, and so far have developed the following grinding techniques: (1) abrasive suspension magnetic polishing, (2) magnetic polishing with the number of active abrasives regulated, (3) pressure-controlled magnetic polishing, and (4) magnetic floatation magnetic polishing. Newer developments are discussed below.

(1) Super Precision Polishing for Cut Surface of $LiNbO_3$ Single Crystal

Researchers at Mitsubishi Electric Corp.'s Institute of Industrial Technology conducted test grinding of curved surfaces on $LiNbO_3$ single crystal samples using their recently redesigned pressure-controlled magnetic grinding machine. They discarded the old rubber polisher and fitted polyurethane pads (extremely thin polishers made of a foamy material) instead, which are typically used for polishing silicon waffers. The $LiNbO_3$ sample measured (phi) 40 X t 2.3. Before the polish was applied, the crystal was wheel ground on its Y-cut surface using a diamond bit.

Optical micrograms revealed that the new polishing completely removed the traced grooves [left by the initial diamond grinding]. According to the R_{max} measurement, the surface roughness was 0.01 micron or less. Also, there was no unbalanced shaping precision. However, the machine had to work an overly long time. How soon it becomes practically viable will depend on whether or not researchers can come up with a solution for this time problem.

(2) High-Speed Magnetic Polishing for Ceramic Balls

Stimulated by the work of Tani, et al, a group of researchers led by Tohoku University's Engineering Prof. Yasuji Kato recently developed a unique polishing method for ball bearings made of silicon nitride. Kato's method makes use of a magnetic fluid. The polishing processor is schematized in Figure 2. As seen in the figure, the silicon nitride balls are immersed in the magnetic fluid that contains the abrasives. The magnetic field acting from below pushes the balls upward. The balls push against the lower end of the driveshaft which sits on the balls.

Now the driveshaft is turned. This motion is transmitted to the ball bearings. They are forced to revolve about the axis. As they travel through the magnetic fluid, they come into contact with floating abrasives, and get polished. A particularly interesting feature in this method is the incorporation of what is called float. This addition greatly enhanced transmission of the magnetically induced levitation pressure to the balls, improving the abrasion action considerably.

Before the polishing, each ball sample measured 7.7 millimeters in diameter

with a sphericity of 500 microns. Each was made of Si_3N_4. With the driveshaft spun at a high speed of 9,000 rpm, the polishing process produced balls with a diamter of 7.1 millimeters, and a sphericity of 0.14 microns. The level of precision achieved ranks at about the same level as commercially available Si_3N_4 ball bearings. However, it only took roughly 1/40 the of the time that a conventional method does. Because of this extremely fast processing capability, if it becomes practical, the new system will be a very competitive tool.

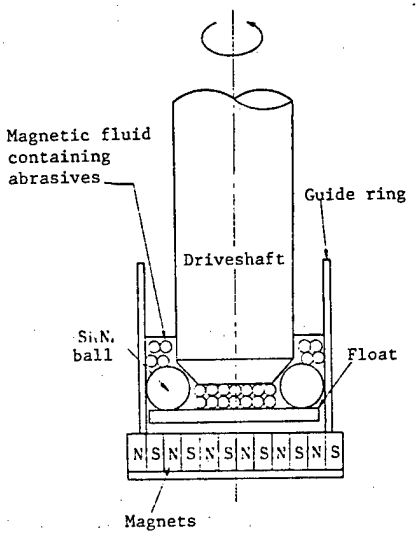


Figure 2. Magnetic polishing of Si_3N_4

(3) Robot-Controlled Magnetic Die Polishing

Under this scheme, the grinding stone is attached to magnets. The combined polishing unit is fitted to the wall of the metal mold, and the polishing action begins. Since the polishing pressure F_n (the force along the normal direction) is given by the magnets, the force F_t which acts along the tangential direction can be generated just by setting the grinding stone into the same tangential motion. Because of the absence of a pressure generating facility, the polishing system itself can be made very simple in construction, and compact in design.

By using electromagnets in place of permanent magnets, the level of pressure [applied to the mold wall] may be controlled through the simple scheme of varying the current supply. Furthermore, by mounting the grinding stone on a robot, the whole operation of die polishing can be precisely manipulated by use of a microcomputer. In fact, this type of die polishing system is being tested now using a cast iron-bonded grinding stone. According to preliminary lab reports, the results are encouraging.

(4) Magnetic Polishing Using Magnetic Abrasives

This polishing method uses a magnetic abrasive material. The powder abrasive used here is a non-magnetic abrasive material like alumina, combined with an iron oxide. The polishing action is accomplished by exploiting the high level of magnetic sensitivity in the abrasives. When a magnetic field is applied, the abrasive particles are start a back-and-forth motion creating a sort of polishing brush. The material to be worked on is positioned between the magenet poles N and S, while the abrasive powder is supplied to the spacings inbetween. As the abrasive brushing action takes place, the material is turned or shaken in such a way as to achieve polishing of the desired surface. This technique is suited to the polishing of complexly-shaped objects.

3. Diamond Polishing Using Thermo-Chemical Reaction

Manufacturers of cutting tools [i.e., blades and drill bits] have to run extra processes to ensure a long life for their products. Typically, tool tips are coated with TiC or TiN to help maintain a sharp cutting edge. Using a diamond coating instead, is suggested by many, and can prolong the tool's life even more significantly. So far, however, researchers have not been successful in growing a flat film diamond. The ones produced exhibited microscopically rough surface features. Flatness of the film surface is a must in order to apply the material to cutting tools.

Recently, the Tokyo Institute of Technology, Prof. Yoshikawa, et al, has successfully developed a new polishing technique for diamond thin films. The method was tested on a diamond film grown on a silicon substrate (8x8 mm, 0.4 mm thick) by micro-plasma CVD. The polishing action was achieved by applying vibrations to the diamond film which was positioned on a heated metal plate polisher. The metal plate adopted by Yoshikawa's group was made of iron, and measured 4 millimeters thick. Other types of plate materials may be used as well such as steel, cast iron, nickle, and cobalt.

To achieve the desired polishing effect, the diamond film was set to self rotate at 0.8-12 rpm and to revolve at 0.2-3 rpm. Also, to speed up the polishing action, hydrogen gas was sprayed onto the polishing surface. Before being sprayed, the gas was run through a dissociation process that consisted of heating it with a filament to 2,000^{o}C or more, and turning it into a radical. The setup parameters were as follows: amount of hydrogen gas flow 100 SCCM, gas pressure 40 Torr, polishing plate temperature of 750^{o}C, polishing pressure of 250 gf/cm^{2}, polishing time of 6 hrs, and rotation speed of approx. 3 rpm.

Evaluation of the polished surface was carried out by using a differential interference optical microscope and a surface probe. Results indicate that the initial roughness was cleanly eradicated from the film surface. To make it a practical system, the process should run faster. Although no definite application plan for this method has been made public yet, once the speed problem is resolved, we can expect diverse applications.

4. Super-Precision Process for X-ray Microscope Optics

Recent advances in state-of-the-art precision manufacturing technology are prompting a fresh look at the design of X-ray microscopes. An X-ray microscope uses X-rays to collect information about atomic-level phenomenon. An X-ray has a shorter wavelength than its visible light cousin. It can penetrate matter far more readily than visible light. Also, the refractive indices of many materials are very close to 1 for the X-ray spectrum. Because of these facts, it was considered extremely difficult to design an optical lens system for X-ray imaging.

This shying away from X-ray optics underwent a radical change when Wolter proposed his X-ray imaging optical system which incorporated the idea of total reflection in 1952. The research community was excited about the possibility, and immediately went to work on many types of X-ray imaging systems. Table 3 lists X-ray microscope designs that have been developed so far. From the table, it is apparent that the microscope's accuracy depends on the level of precision during the machining of the optics, which corresponds to the object lens in a conventional microscope.

Table 3. Major X-ray Microscopic Techniques (by Aoki)

Optical parts	Optical system	Spectrum (nm)	Resolution (nm) Tested wave-length (nm)	Resolution-determining factor
None	Projection magnification	0.1-1	100 (0.1-0.5)	Intensity of X-ray source; degree of diffraction
	Adhesion	0.3-10	10 (approx. 5)	Imaging resolution of photo-emulsion, degree of diffraction
Yes	Zone plate	1-10	50 (4.5)	Minimum line width
	Angled incident mirror	0.1-10	1,000 (0.83)	Contour accuracy and roughness of mirror surface
	Multi-layered mirror	0.1-10	1,000 (0.15)	Contour accuracy and roughness; film growth precision
	Scanner type	0.1-10	150 (3.5)	Size of focused beam

Figure 3 is a schematic of the angled incident X-ray microscope. The accuracy of the microscope depends on the quality of the toroidal mirror. There are two methods available for making a toroidal mirror. In one of them, the inside surface is directly polished. In the other method, a metal mold is first polished, and then a replica made from that is used [as a mirror]. The mirror material made of Mo or WC is resurfaced to the initial finish by sujecting it to a grinding or lapping action. This initial grinding step is followed by the super-precision polishing. Diamond abrasives are used for the grinding job. This reduces the surface roughness to a submicron range. For the final polishing task, a pasty substance made of finer diamond abrasives of a uniform grain size--a hydraulic elutriation is used to achieve the grain homogeneity--mixed with a paste material is used. According to test results, a deviation in shaping accuracy of 1 micron or less can be achieved, with the surface roughness reduced to a few nanometers (maximum roughness).

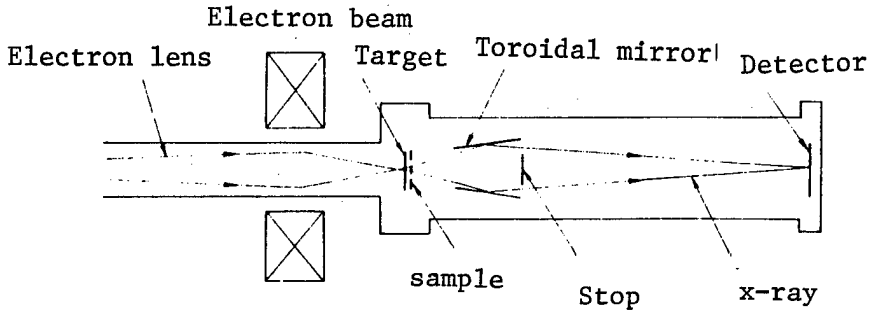


Figure 3. Angled incident X-ray microscope system with a micro focusing X-ray generator unit in use

The demand for high precision polishing is not limited to the machining of toroidal mirrors for X-ray microscopes. Today, there are many more types of special purpose mirrors that require a similar machining process. Polygon mirrors, non-spherical mirrors, and mirrors for synchrotron orbit radiation are some of these. Future developments in mirror design will certainly bring forth even more. Also, besides the abovementioned multi-surfaced mirrors, there will be a need to perfect the process for ultra-flat mirrors as well. By taking advantage of the unique features they possess, the processing methods discussed earlier and listed in Table 2 may be tailored to achieve these specialized processes. As of this writing, the NC EEM process, applied to the polishing process for synchrotron orbit radiation mirrors, is producing outstanding results.

The demand for super-tough brittle materials such as fine ceramics and superhard alloys is expected to grow even more in the future. Industrial products of tommorrow will be required to meet even stricter machining standards. We know that one of the paths leading to high-efficiency high-precision manufacturing lies in the design of super abrasives. Designing next-generation abrasive materials, as well as improving other material aspects--improving the grinder/polisher machines in terms of mechanical strength and precision, bearing quality--is indispensible. At the same time, development of software for better control of the interaction between the abraded material and the abrasives is equally critical to improving productivity.

Super High-Speed Bearings, Guides for Machining Tools

43064080e Tokyo KIKAI TO KOGU in Japanese Nov 88 pp 21-25

[Article by Ichiro Inasaki, Department of Engineering, Keio University]

[Text] One of the new trends in design and production technologies for machining tools is the pursuit of high speed operation. There is nothing new about seeking faster machines, as those have always been desired. The recent rekindling of interest takes on a higher dimension though, as it is being fueled by the rapid advances being made in various fields of industrial and machining control technologies. The automation technology we have today is already dramatically cutting down on the non-machining process time. Now, engineers are contemplating ways to shorten the actual machining time itself. They are also tackling the problem of designing machines with a greater level of flexibility in terms of operation speed. They want the same machines operated at widely different spindle speeds to accommodate diverse machining needs, for instant productivity increases.

The intense research and development activities for faster machining operations are helping to raise the overall quality and level of performance in the machining tools we have today. This, indeed, has a positive impact on the entire society.

In this article, I would like to survey recent developments in the design of bearings and guides, which form the very basic technology for developing faster machining systems.

1. Function of Bearings and Guides in Manufacturing Machines

Before I embark on a discussion of recent developments, let me give you a summary of the functions and capabilities of bearings and guide systems. The bearings and guides constitute the most critical design factors in achieving the desired high level of accuracy for maintenance of the machine's prescribed motion. In general, the spindle and table which are supported by these components are allowed to take a single rotational motion and a single translational motion.

Any given object in a 3-dimensional space has a total of 6 degrees of freedom: 3 translational motions and 3 rotational motions. Therefore, the fundamental function of the bearings and guides is to prevent the machine from entering into 5 of the possible motions, leaving it with a single degree of freedom. Then, these components must ensure that the only allowed motion will be as free of deviations as possible. The quality of rigidity sustaining strength in bearings and guides is evaluated according to the measured degree of deviation from the preset motion.

Paraphrased in a more practial sense, the purpose of bearings and guides may be described as, "fixing the desired motion of the machining tool or the material being worked on, and cushioning the tool and material from motion-induced impacts, and machining impacts."

The quality of bearings and guides is evaluated in terms of the following characteristics:

(1) accuracy in sustaining the motion
(2) static and dynamic rigidity
(3) heating and thermal deformation characteristics
(4) rolling friction, and wearing
(5) maintainability

Improving these characteristics requires research effort into multiple directions for the bearing and guide materials and designs, and lubricants. No matter how fast the spindle turns, unless the above requirements are strictly met, all the effort that goes into the final machine will be wasted.

2. Bearings for Faster Operation

To speed up the spindle operation, one must first redesign the bearing sections. Redesign of the bearings has to be accompanied by a redesign of the table guide mechanism in order to raise the speed so that it matches that of the spindle operation. Generally, by increasing the speed of the machining, a greater level of productivity may be gained. Also, the gain in machining speed can improve the mechanism for removing machined materials, and can lead to better resurfacing accuracy as well as greater reducion in resistance to the machining motion. Raising the spindle speed in the machining center is essential if small diameter drill bits and end mill cutters are to be used in order to gain a sufficiently great increase in the machining speed.

To redesign the antifriction bearings used in ordinary machining spindles, we first have to resolve the problems listed in Figure 1. Different methods to solve these difficulties are now being tried. Some of them are already at a practial application level.

In the next section, I will talk about the oil-air lubrication from the stand point of design of the lubrication method, and the ceramic bearings from the standpoint of mechanical considerations.

2.1 Oil-Air Lubrication

With a normal grease-type lubricant in use, the upper limit of effective lubrication is said to be around a dn value of 800,000. If a higher range of values is sought, a completely different method is needed such as jet lubrication, or oil-air lubrication.

In recent years, the oil-air lubrication technique is gaining significantly large research attention. Under this method, droplets of lubrication oil are intermittantly mixed with a large volume of air. The same flowing air is used for cooling purposes. The lubrication oil is throttled out from the mixing valve at a rate of once every couple of minutes or so. This method has the advantage of making it easy to reduce heat buildup. On the other hand, it may not be sufficiently reliable. It also creates excessive noise as the sprayed air hits the bearing unit.

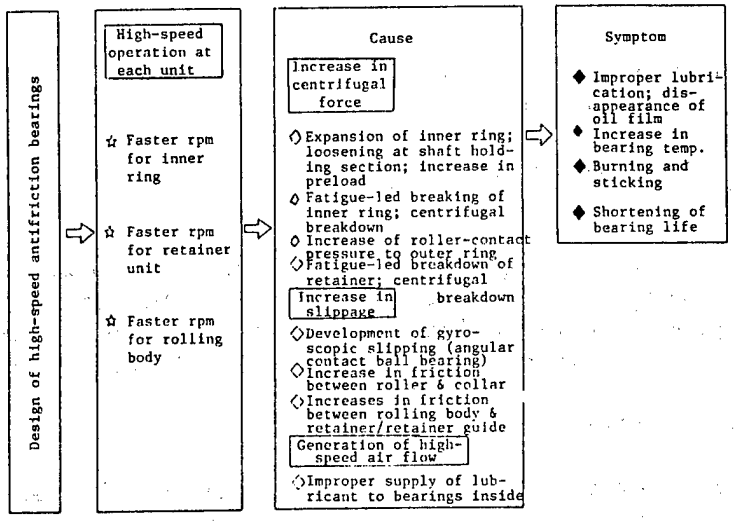


Figure 1. Problems associated with the design of high-speed antifriction bearings

After leaving the mixing valve, the lubricant oil is sent via the pipe to the bearing section. The pipe has to be made long enough to ensure that there will be no major change in the amount of lubricant oil supplied to the bearings. That way, a sufficient amount of oil is always uniformly distributed on the inner wall of the tube, guaranteeing a nonstop spraying of the lubricant from the outlet nozzle.

The most difficult problem to overcome when considering adoption of the oil-air lubrication method is securing of the abovementioned long tubing. There are quite a number of reports of the bearings having burned because of insufficent pipe length, causing an inadequate supply of lubricant to the bearing unit, just after the machine has been cranked up. This type of mishap can be averted by installing a monitoring system which shuts down the system if the supply is not up to the preset level.

At our lab, my associates and I developed an experimental oil-air lubriation system. As shown in Figure 2, it incorporates a nozzle made of an electrostrictive material. We used the electrostricting nozzle device originally made for an ink-jet printer. Our system successfully generated oil droplets whose diametric size reached 100 microns or less. Also, it was designed into a compact construction.

Because of the small size of our system, it is possible to locate it very close to the bearing unit, making it unnecessary to run a very long pipe connecting it to the outlet section. Also, because of the small size of the oil droplets, we could set the constriction frequency of the electrostriction nozzle from a 100 to a couple of 100 Hz. Because of these advantages, we could deliver the lubricant almost continuously to the bearing section. With our system, it is possible to vary the quantity of delivered lubirant by simply adjusting the frequency level.

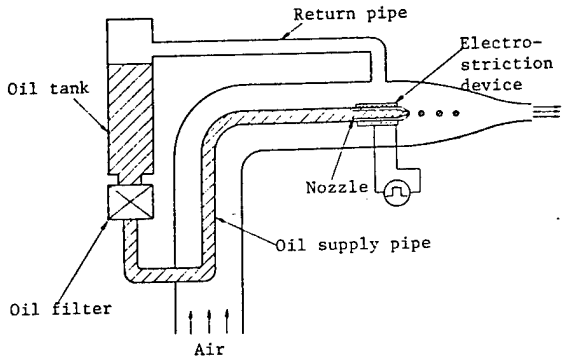


Figure 2. Oil-air lubrication using an electrostriction nozzle

2.2 Ceramic Bearings

There is a steady rise in the use of antifriction bearings made of ceramics. At present, the application is limited to high speed bearing uses, and in most of these instances, ceramics are used in the rolling bodies only. Most of the ceramic materials employed for this purpose are silicon nitrides made by the HP and HIP methods.

The advantages in using ceramics to make the rolling bodies are numerous. It gives very good thermal resistance, resistance to friction and wearing, and excellent rigidity. On top of all of these features, it has a very low density. When subjected to a fast rotational motion, the angular contact bearings which are used in the spindle sections of many machining systems generate a gyroscopic slipping in the rolling bodies, causing a heat buildup.

The greater the mass of the rolling element, the greater the chance of getting a gyroscopic slippage. Therefore, the use of ceramics whose density is comparably low is indeed advantageous. Silicon nitride is a mere half of the density of steel, a typical bearing material.

A plot of the temperature increase in a ceramic bearing unit with oil-air lubrication in use is shown in Figure 3. By using ceramic bearings, we reduced the temperature rise from 35 to 70 percent of the value given by steel bearings. The extent of thermal deformation was smaller too. It was roughly 1/3rd. The upper dn value with ceramic bearings and an oil-air lubrication in use is estimated to reach 1,500,000.

In the near future, more extensive use of ceramic materials is expected. Not only the bearings but the inner and outer rings will also be made of ceramics in order to achieve a dramatic improvement in elastic deformation as well as the static rigidity of the bearing unit. But, from the standpoint of damping, or dynamic rigidity, characteristics, we expect to have some problems. Because of the decrease in contact deformation taking place between the roller elements, and the inner and outer rings, the damping effect is expected to be worse.

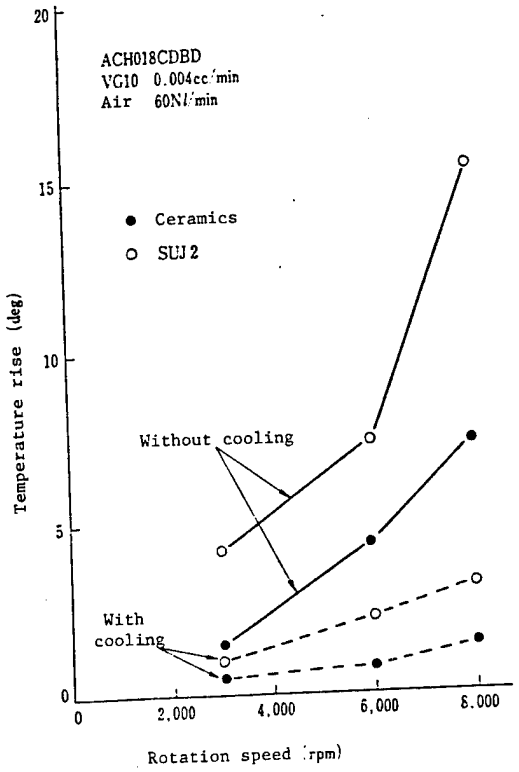


Figure 3. Temperature rise in ceramic bearings (oil-air lubrication in use)

3. Guide for Faster Operation

In order to reduce not only the non-machining time, but also the time spent on actual machining, engineers are designing new methods to achieve a faster table guiding system. Already, with the present level of guiding technology, we can quite easily obtain a guiding speed of up to 20 or so meters per minute.

3.1 Roller Guides and Their Characteristics

One of the techniques frequently used to increase the table guiding speed is the use of a roller guide. The trend of increasing the use of this system is illustrated in Figure 4. It shows the growth trend in the production of manufacturing machines by a certain Japanese tool making company. The demand for NC machines experienced a sharp upward trend in 1983 or so. After that, the same strong performance persisted.

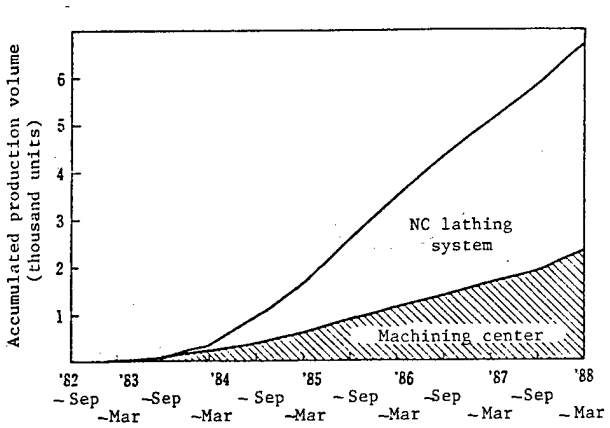


Figure 4. Market demand for linear guide systems

One of the biggest gains in using a roller guide is a significant drop in the value of the friction coeficient. With a smaller friction force acting on the moving parts, the power needed to run these parts does not need to be as large. This fact will be an added advantage in designing a faster operating machine. In addition, since rolling friction is generated as well, the coefficient of static friction will not be so outstandingly large compared to the coefficient of kinetic friction. This means that the stick-slip phenomenon which often creates problems at a slower operation speed can be eliminated, at least theoretically.

Even from the point of view of manufacturing considerations for the guide section, it is generally a goodt idea to adopt the roller guide design too. The system is not problem free, of course. Generally, its damping effect is considered quite bad. At our lab, we asked ourselves, what is the difference in damping effect between the roller guide and a gliding guide? The latter uses a general-purpose resin material as the gliding surface. Let me outline our findings.

3.2 Dynamic Rigidity of the Roller Guide

Excessive damping and excessive dynamic rigidity in a machining system can induce chattering vibration during the machining process. As far as the area of contact made between the guide unit and other machining components is concerned, the gliding guide system should score higher than the roller guide counterpart.

The table system that we used in our test is shown in Figure 5. It weighed about 850 N. Its guide sections, for both X and Y transport requirements, were designed with special consideration being given to the ability to exchange the rolling guide and the gliding guide. The gliding guide had [a sheet of] (tarkite) B glued on. The ball screws were fastened so that the table could not make any movement.

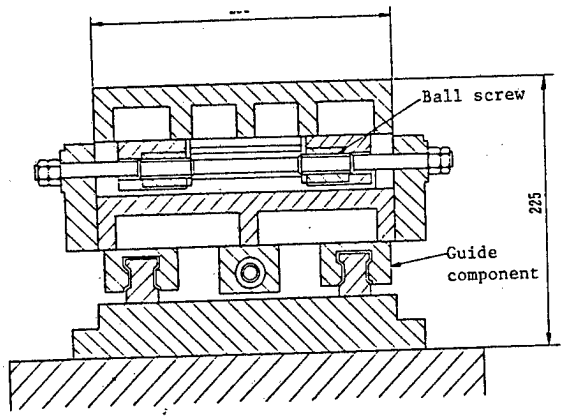


Figure 5. Construction of an X-Y table model

The dynamic compliance was measured for each guide system using an impulse response method. The results are plotted in Figure 6. From this, we can see that the rolling guide has a significantly greater compliance than the gliding one. Also, we can expect an increase in amplitude at resonance. If [the rolling system is] used for an intermittent cutting process such as milling, this effect may lead to problems.

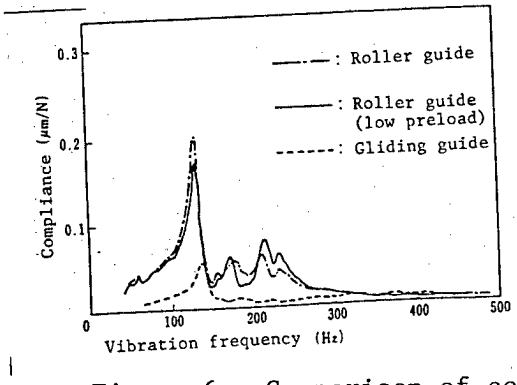


Figure 6. Comparison of compliance

When a smaller diameter roller was selected for the rolling guide unit to make the preload smaller, the amplitude of the fundamental resonance signal near 120 Hz did go down a bit. This indicates that the increase in the rigidity of the contact section caused by the increase in the preload lowered the damping ratio.

There is a strong correlation between the occurence of self-induced chattering vibration, a regenerative effect, and the negative real maximum of dynamic compliance. Figure 7 is a plot of the real part of the compliance as a function of frequency only. This shows that the rolling guide system induced a greater negative real maximum. It has been theoretically shown that for a system having a large absolute value of compliance there is a greater probability that a self-induced chattering vibration will occur.

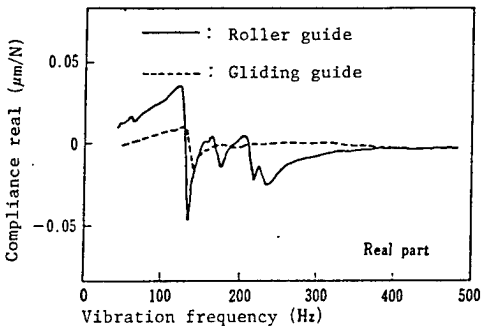


Figure 7. Comparison of real number compliance

We conducted actual machining sessions and measured the intensity of the chattering vibration. The results for the two systems are plotted for comparison in Figure 8. Our setup consisted of mounting the experiment table on the NC milling table, and using the milling table to simulate the motion of a real work table. The material to be milled was a disk and fitted onto the experiment table. An end mill cutter was used to execute circular milling. With accelero-vibration meters fitted to the table along both X and Y directions, the intensity of the vibration generated by the milling action was recorded as the cutter burrowed along the 360^{o} path. In recording the vibration amplitude, we used a 130Hz bandpass filter whose frequency response range matched the guide system's negative maximum compliance region, to isolate the conditions causing the generation of the regenerative self-induced vibration.

The curves indicate that with the gliding guide in use, for a spindle revolution of 700 rpm or less, there is almost no self-induced chattering vibration. On the other hand, when the rolling guide was used, the machining unit became unstable, accompanied by generation of self-induced vibration.

We cannot draw an immdiate conclusion just from the above results, claiming that the rolling guide is disadvantageous to use. The character of the vibration may differ greatly depending on the level of the contact pressure exerted on the guiding unit. After all, the comparison we undertook did not take into account the fact that our experimental setup did not ensure an equally conditioned processing environment. Anyway, even if we wanted, we could not have done so.

Nevertheless, it is true that under certain conditions, there will be a similar difference. In any case, as long as the guide system choosen does not develop a chattering vibration within the range where actual machining takes place, that system can be said to have a sufficiently large dynamic rigidity.

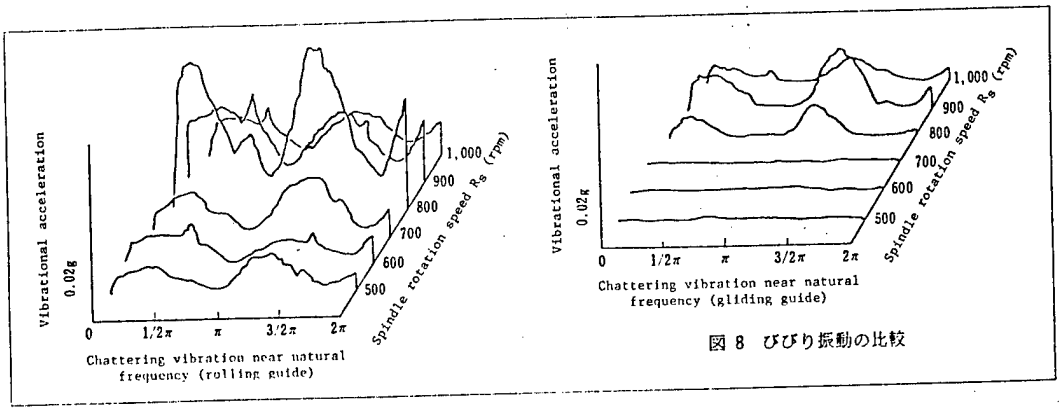


Figure 8. Comparison of chattering vibration

As for research, needless to say, we must continue striving hard to improve the dynamic rigidity characteristics if we want to improve the guide system at all.

In the above sections, we noted that the bearings and guide systems constitute critical parts of machining systems. We discussed the current status and technological issues regarding the direction of R&D toward development of high speed machining systems. We have learned that one of the more practical solutions to development of a fast system is to utilize the roller technology. The technology is promising, albeit, problems do exist. Reliability has to improve. As far as the rolling bearing system is concerned, first, we need to tackle the problem of excessive damping. Let me conclude this article by briefly introducing the new attempts that we are making in our lab to solve these problems.

For the rolling bearings, we are designing a scheme that uses an electrostrictive-device actuator to enable adjustment of the preload that is applied to the bearing section. The electrostriction device is fitted between ball bearing units where spacers are inserted for proper preload adjustment. For a lower-speed revolution, the bearings are given a greater preload while for higher-speed operation, they are give a smaller preload. This way, burning of bearings can be avoided.

As for the rolling guide, we are designing a hybrid system combining both the rolling and gliding guide methods. With it, a different guiding mode may be selected depending on the machining operation mode. For high-speed operation when fast transport is needed, we select the rolling guide mode, while for low-speed transport during actual machining, or for positioning the table into place, we select the gliding mode.

AI-Based 3-D Coordinate Locator

43064080f Tokyo KIKAI TO KOGU in Japanese Nov 88 pp 26-30

[Article by Susumu Oishi, Department of Engineering, Tokyo Metropolitan University]

[Text] The 3-dimensional coordinate locator consists of a pointing device and a computer. The basic function of the pointing device itself is to read the 3-D coordinate position of an object. The probe of the instrument is brought to a certain spot on the object, and the corresponding coordinate information is recorded. Many locator systems use the 3-dimensional rectangular coordinate system to map the collected data. Depending on the way that the probe is handled, the locator system is classified according to several types: the manual-probe model where the probe is moved around by a human operator; the joystick model where the probe is remote controlled via a joystick; and the CNC model where the probe is numerically remote controlled. The CNC model is expected to be most popular in the future. For now though, most of the locator devices we have are of the manual type, and they are usually used as stand-alone machines.

The computer is indispensible in the 3-D locator system. By processing the coordinate data read in from the probe using the built-in computer, the following types of information may be obtained regarding the target object:

(1) position and size,
(2) geometrical deviation, if any, from the expected shape (i.e., sphericity), posture (i.e., parallelism), position (i.e., angular displacement), and course (i.e., deflection from a circular path), and
(3) contour profile.

Based on the above information, the system can perform the following:

(4) a cross reference of the design specifications (i.e., size, dimensional tolerance, displacement tolerance),
(5) preparation of routine statistics, and
(6) preparation of NC tapes.

As you can see, [with a computer built in], numerous types of measurement become possible, making it a highly general-purpose system.

To illustrate the diverse capabilities of the system, let us consider the typical measurements of items (1) and (2). In this case, the system performs the following:

- o correction to the angular displacement of the object placed on the locator's measuring table,
- o identification of the coordinate origin, and
- o identification of the coordinate axes.

After these measurements are taken, the system performs many more measurements such as locating specific points, lines and circles. With all the identification algorithms and probe settings counted, the number of measurement-related jobs that the system performs reaches 100.

One of the factors that led to the rapid spread of the 3-D coordinate locator is the great improvement in the measuring precision and reliability of the system. It's one thing for the system to perform many different types of data collection. But, the fact that it takes an expert, well experienced in the use of the measuring system, to interpret the data and drawings is a totally different thing. The pointing device gives outstanding accuracy, but the system as a whole may fall short of being used efficiently. This sets the background for initiating heavy research effort into ways to ease the data interpretation part. Already many proposals have been made. Among these is the incorporation of an AI system.

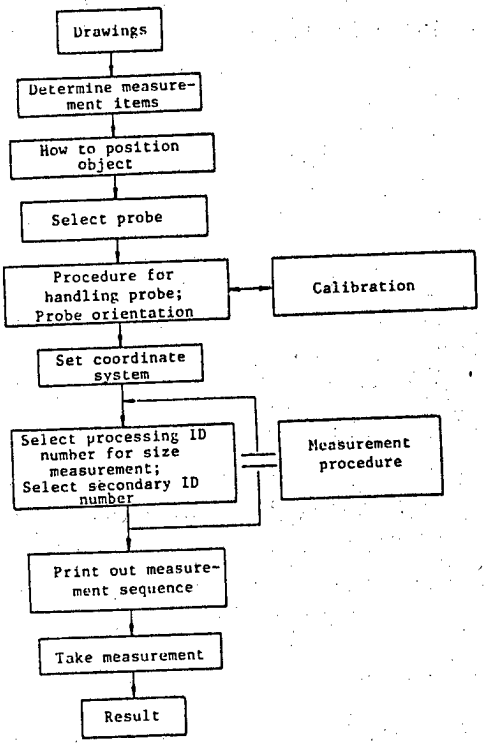


Figure 1. Typical procedure for a measurement process

1. Sequence of Steps in Generic Measurement

Figure 1 illustrates the sequence of steps taken to complete a typical measurement task. Usually, before you can start taking measurements, you need to know where, how, and in what sequence the measurements need to be taken. These problems are resolved by considering the following points:

(1) In many instances, when design drawings are made available for locating the spot at which the measurement is taken, the reference point may be selected from the drawings. Sometimes, the drawings may not be as helpful as they could be. If the former is your case, let the reference point equal the predefined reference used throughout the manufacturing process. But if you cannot rely on the drawings, determine it at the site by considering the designed capabilities and functions of the target product.

(2) In order to minimize the number of probe settings, carefully consider the way to position the pointing device on the site.

(3) Select a probe which does not need to be replaced frequently. This consideration does depend on what is to be measured however.

(4) Position the probe in such a way as to facilitate the identification of the coordinate system.

(5) In setting the coordinate system, first, the coordinate references must be determined. There are two ways to do this. One is based on a system of referece planes, and the other on reference axes. If there is no indication as to how the referece is to be made, stretch the reference planes or axes as far out as possible.

(6) As for the assignment of the processing ID number, follow the basic rule, one measurement item = one ID number. In the case where the measured item does not fall on a reference plane, you need to give it a secondary ID number. Later, measured items should be re-categorized according to the same [principle] ID number.

(7) Consider the preparation of the measurement procedure as preparation of a [formal] handbook.

So, before the actual taking of measurements can begin, a handbook of measurements or a similar guide needs to be prepared. This documentation apparently consists of expert knowledge of the instrument operator.

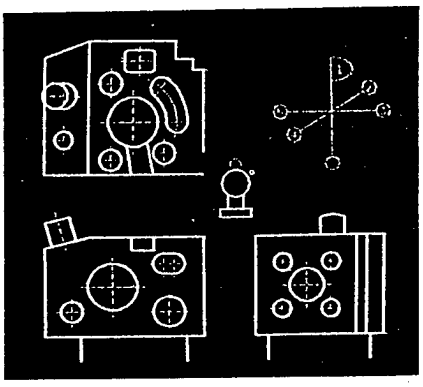


Figure 2. A CRT display frame from a probe-positioning tutorial; a set of rough constructions of the target object

2. Ways to Facilitate Use of the 3-D Locator System

Because of the complexity involved in using the system, learning to use it is a major, time consuming task for many beginning operators. Basically, the training consists of the following:

(1) orientation lasting for a few days after the purchase of the instrument, and

(2) purchase of tutorial training kits consisting of demo programs and a training manual.

Aside from the general training of personnel, the following considerations may be made:

(3) A tutorial session is given to personnel using computer-generated CRT displays of graphical figures and maps for the measured object and the measuring probe, which is prepared according to the measurement handbook. The purpose is to have people learn on their own where to place the probe and in what sequence the measurement is to be carried out. Figure 2 shows a typical screen display.

(4) A "Teaching playback" session is conducted using a CNC system as illustrated in Figure 3. The actual locator system is given instructions as to where and how the probe needs to be directed for the specific target point. When its time for an actual measurement session, the system carries out the instructions automatically.

(5) "Off-line teaching" session is conducted using a CAD database facility. The purpose is to maximize the time-wise productivity of the locator system, and to integrate the stages of product design, manufacturing and quality control.

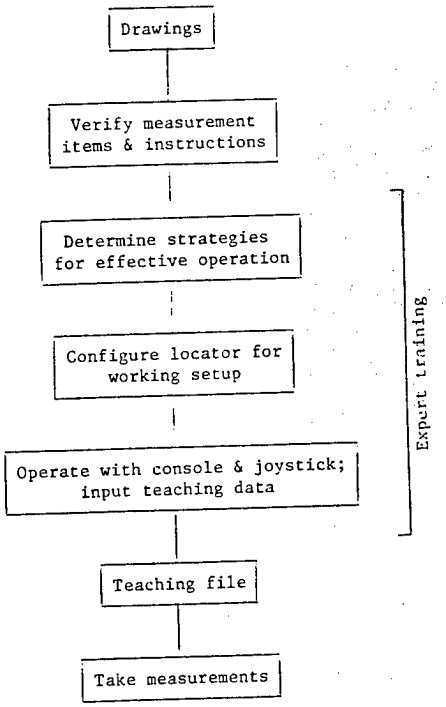


Figure 3. CNC-based teaching procedure for locator/probe operation

The training aspects (1) and (2) are geared toward beginning personnel. The program of (3) assumes that well experienced operators are present already, and the graphic presentation program has been prepared by them. As for techniques (4) and (5), here too, expert operators must be present before these sessions can be scheduled, because the specific questions of where, how, and in what sequence need to be answered in advance.

3. What Is an AI-Based 3-D Locator System?

AI stands for artificial intelligence, It represents a field of computer science dedicated to the design of intelligence processors using machines that mimic the human brain. It is hard to give the precise definition of "intelligence". Generally, it means scholarly knowlege, knowledge from experience, expertise in certain tasks, etc.

One of the AI applications is the expert system. The computer provides consulting services using expert knowledge stored in a data bank. Figure 4 schematizes the basic organization of an expert system. The knowledge base is a database facility that stores the knowledge of experts in a particular task. The decision-making engine is where the computer executes an algorithm to solve the given problem by making reference to the knowledge base.

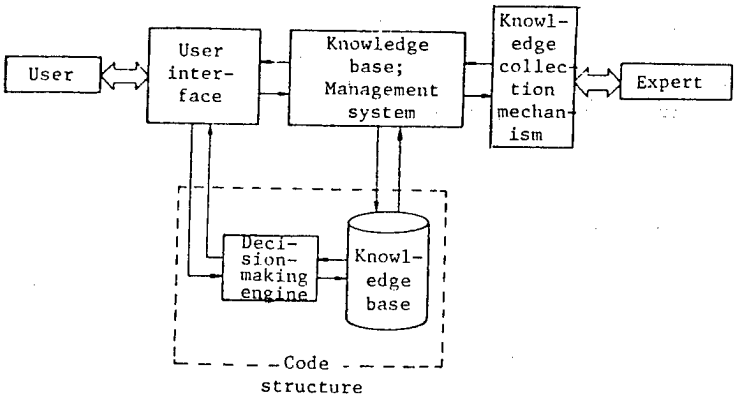


Figure 4. Organization of an expert system

Because of the complexity involved and the requisite expert knowledge, the undertaking of the operation of a 3-D locator system is not an overnight task. By incorporating an AI system into it, we can remove the unnecessarily tedius routine task loads from the operator's shoulders so as to make it sufficiently painless for beginners to learn the operation quickly.

4. Design of an AI-Based System

A majority of the operators we have now do not know the purpose of the objects that they are pointing at. They can read the drawings, and so, they can identify the pointing references and where to point at. On the other hand, with beginners, in an extreme case, they may not even be able to read the drawings. Also, there may be no drawings available for some assembly parts. In such cases, first of all, the operator needs to determine what to measure on his own. The pointing locations may be then deduced from the shape and usage of the measured object.

This task of determination may also be done by an expert system. In designing such an expert system, we first input the data for the object's shape and usage into the system. For instance, let's see how computer pattern recognition for an object's shape may be done by citing research work on pattern recognition using angular and round shapes. In a dialog mode with the host computer, the operator enters the features of the shape in question into the system. The system makes a cross reference of the knowledge (expressed as "frames") that it has in its database regarding the shape of various objects. The indicated shape is thus identified. Next, the system looks to the information expressed in terms of production rules regarding pointable locations of the object. After a cross examination of the queried data, and the determined features and attributes of the shape, the system deduces the pointing locations and coordinate references. At the same time, it generates a measurement process ID code, and completes the investigation of the shape in question.

Because the computer system has to work with a limited database to generate any type of display output, the image that the human operator sees on the screen may not accurately represent what he has to actually place a probe on. This may cause confusion among operators. Why not squeeze more information into the system so that it can draw a finer image? There are an infinite number of possible shapes that the measured object may take. Obviously, it is practically impossible to prepare a database for all of them. In any case, even if the specifics of the measurement task--pointing locations and reference frame--are known in advance, it is still beneficial to consult a type of expert system and go over CRT simulation images of the object's shape and the locations for probe setting that the expert system generates no matter how primitive they look.

What would the ultimate 3-D locator system be? With such a system, all we have to do with the target object is to put it down on the measuring table. All the probe settting and reading will be automatically handled by the system. But no matter how advanced the system may be, it will still need shape modeling information and the capability to recognize patterns, like the not-so sophisticated siblings we now have. Aside from the abovementioned manual method for inputting shape-related data, we have two other methods available. One uses a camera and an image processor. The object to be measured is photographed. The picture image is converted into numerical video data, and entered into the expert system. The system generates a probe path from the input data (See Figure 5). In reality however this technique involves many unsurmountable technical difficulties at present. Because of blurs in the captured image, or because of the sheer complexity of the shape, the system cannot correctly comprehend it. Let's discuss the other method. This utilizes CAD. Figure 6 illustrates how CAD may be used to accomplish image entry. In this specific scheme, consideration was taken to avoid taxing the locator system, and so the CNC-based "teaching" session was conducted off-line. A CAD system was instead used to supply the required shape-related data. It was an extensively automated locator system, and was designed to run under an integrated CAD/CAM system. Similar design attempts have been made elsewhere.

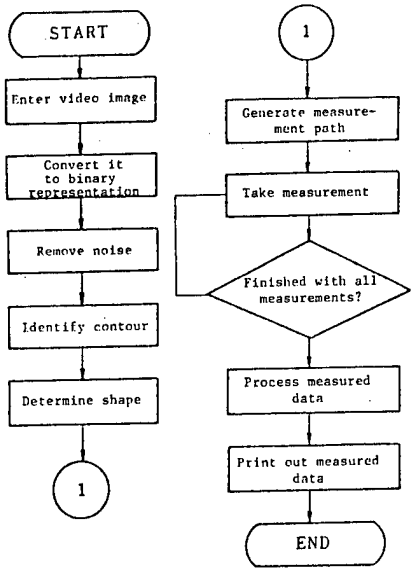


Figure 5. Input of shape-modeling data of the object to be measured using a video imaging device, and generation of the measurement path

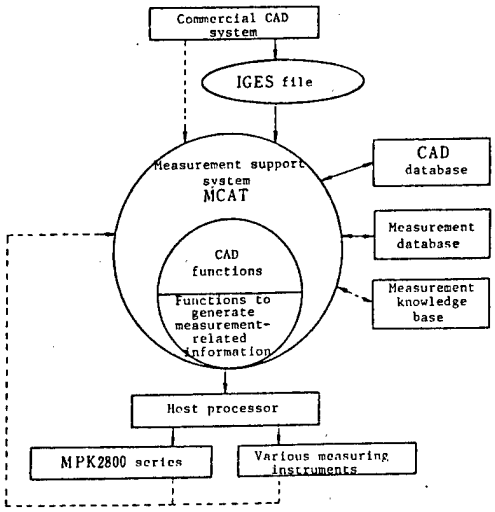


Figure 6. Input of shape-modeling data using CAD as a measurement support system

Following the determination of the pointing locations of the object to be probed, we need to determine how and where to position it; which probe to use, how to use it, and where to set the reference frame. All of these aspects have to be considered with an eye for greater productivity and measurement quality. In fact, this is the most complicated stage of all in the entire measurement process, and knowledge of expert operators is essential. Even when the locator system is considered part of an integrated CAD/CAM process, the same thing can be said.

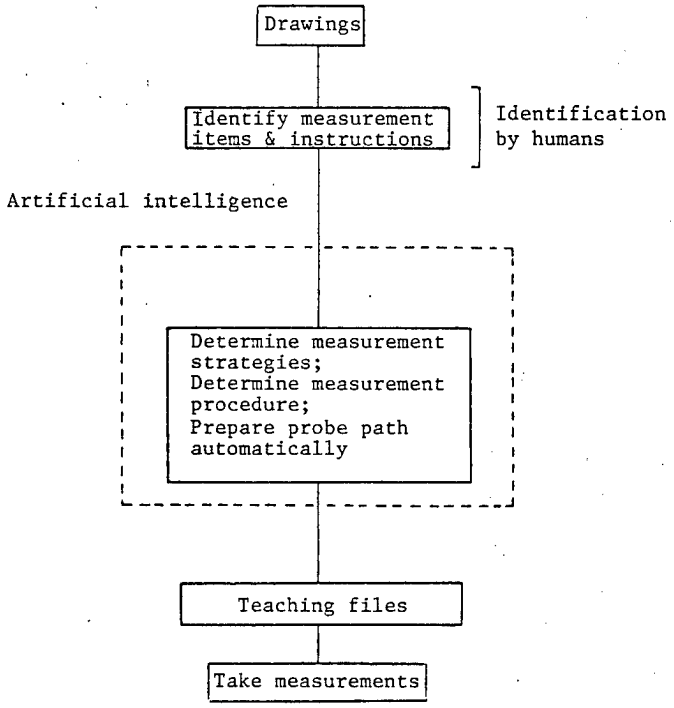


Figure 7. Teaching task using an expert system

The procedural sequence for an expert system-based off-line teaching task is illustrated in Figure 7. In this setup, the operator performs the following:

(1) selection of the measurement items,
(2) determination of the order of measurement,
(3) determination of the positioning for the object, and determination of the reference frame, and
(4) selection of the probe.

Following the completion of the above tasks, the expert system determines the following:

(5) measurement point-specifying factors,
(6) projection planes,

(7) number of measurement points, and measurement pattern,
(8) probe angles,
(9) pointing locations, and
(10) point-to-point path.

The AI data used to conduct these tasks is expressed in terms of frames and production rules. The system uses a foreward-referencing decision-making algorithm.

As far as the system illustrated in Figure 7 is concerned, the tasks (1) to (4) will also be conducted by the expert system, according to the developer. As for the previously cited research work on pattern recognition, an AI system was incorporated to handle even tasks (1) and (3). Refer back to the eariler discussion for the actual processing of task (1).

Ideally, the only thing that the operator has to do is to place the target object down on the table. The machine will take care of all the measurements. This might not be a far fetched idea under an integrated CAD/CAM scheme. But while we still use stand-alone manual models quite often, the need for an effective method to acquaint freshman operators with the 3-D locator system is definitely there. For even an untrained worker to be able to operate it, the operation task must be reduced to simply following what the expert system indicates on the screen. The task for researchers is to create this expert system based on the manual instructions and the experienced worker's expert knowledge.

And as we have noted, this task is being implemented from many different angles.

Laser Microprocessing

Status, Trends in Laser Microprocessing

43064082a Tokyo OPTRONICS in Japanese Oct 88 pp 88-90

[Article by Mitsugu Hanabusa, Toyohashi University of Technology]

[Text] 1. Microprocessing

Today, industrial needs for microprocessing are extensive, and still growing rapidly, prompting greater efforts in research and development for innovative processes. In this article, I would like to discuss the present status and future trends in microprocessing technology with the focus on laser applications. In particular, I will talk about three technologies: laser CVD, excimer laser lithography, and laser plasma X-ray lithography.

It is in the semiconductor industry that microprocessing technology is most widely employed, and the industry's performance for the last few years indicates that the accelerating trend of heavier circuit integration in semiconductor devices will not reach a ceiling any time soon. As far as the history of MOS memory is concerned, we are in the age of mass production of 1-mega bit memory, with the era of the next-generation 4-mega bit storage technology just around the corner. As if to say that the pace is not fast enough, semiconductor makers are already releasing prototype designs for 16-mega bit MOS specifications. The smallest logic element in the circuit pattern for a 4-mega bit integrated circuit measures 0.7 to 0.8 microns. With a 16-mega bit chip, the required dimensions will have to be as small as 0.5 microns.

By the way, the semiconductor industry is not the only one benefitting from microprocessing technology. In fact, many other fields of science and technology are finding needs for the new technology. For example, recently, engineers have become more and more interested in devices that take advantage of quantum effects. But to achieve the desired quantum effect, which appears as electrons are confined in an extremely small region of space, the circuit has to be patterned with a line resolution greater than that for a conventional semiconductor integrated circuit. Then, there are micro machines which can also take advantage of the new technology. These may be mini robots made with tiny toothed wheels and motors. The logic elements in an IC controller for such a machine may be made to a size of roughly 100 microns. Also, many manufacturers are now turning to the use of light, instead of

electricity, in designing various devices. Here too, the technology of microprocessing is expected to play a vital role.

2. Beam Generation Technology

With new technological innovations constantly taking place, the technology of microprocessing is undergoing rapid change to suit the diverse needs of tomorrow's industry. Basic to this technology is the technology of beam generation. There are three types of beams available: light, electron, and ion. Among these, the ion beam offers the best resolution, but there are unsolved problems such as the achievement of a stable supply of ions and the designing of high-precision optics. Therefore, prospects for extensive applications in the near future are slim. With the electron beam, it is possible to draw patterns with a line resolution of 0.1 micron or less. In fact, with present commercially available systems, line drawing widths ranging from 0.2 to 0.3 microns may be readily obtained. But, because of the poor throughput (processed volume of work), an electron beam system is not suitable for mass production. It is, however, used in R&D-type jobs where [waffer] masks and prototype IC chip sets are made.

However, most industrial lithographic methods rely on the use of light beams. Under a photolithographic technique, by exposing the photoresist material coated on a waffer to a beam of light through a mask, the circuit pattern is transferred onto the waffer. This photo-exposure technology has continued to evolve, incorporating innovation after innovation, to match the need for a greater pattern resolution of the time. Since the maximum degree of pattern resolution is related to the diffraction limit, a beam of light with a shorter wavelength is preferred. This fact first led to a scheme whereby a component of light with a shorter wavelength was screened for use from the light spectra emitted by a mercury lamp. Today, a new breed of light spectra with an even shorter wavelength--excimer laser--is used in what is known as laser lithography. At present, a KrF laser with an oscillation frequency of 248 nanometers is usually employed for the job. This laser system is capable of a working resolution for pattern imaging on the order of 0.5 microns. These developments are encouraging chip designers to continue using light beams well into the age of the 16-mega bit specification.

A light beam is also used in photo-CVD which is utilized for vapor deposition of thin films. With laser in use in place of ordinary light, a CVD technique can allow simultaneous growth and direct patterning of thin film. There are two laser CVD methods: a direct laser drawing method, and a pattern transfer method. The latter process may be more properly thought of as laser lithography that involves a resist exposure process. But in terms of a historical perspective, laser CVD technology appeared first before its lithographic counterpart, and was put to use as a process for direct laser drawing by taking advantage of its good spatial selectivity for thin film growth.

The continued quest for a shorter wavelength spectra makes it unavoidable for IC designers to ultimately turn to X-ray's. It is hoped that X-ray lithography will set a new standard for photo-exposure technology. Laser technology is applied to the task of generating X-ray radiation -- X-ray's

are generated by laser-induced plasma -- and this is where conventional X-ray tubes fail. The industry is enthusiastic about the possibilities that the new laser plasma X-ray lithography can offer.

3. Laser Microprocessing

For detailed technical aspects, the readers are advised to refer to references by experts on the subject. I shall limit my discussion to a survey of the methods available and a cross reference of them.

3.1 Laser Lithography

For production of large scale integrated circuits (LSI), the image-reducing projection photo-exposure technique is adopted at present. The technique is highly reliable, and is widely employed for mass production. The light source used is a high-pressure mercury lamp. Out of the mercury lamp's emission spectra, the shorter-wavelength g-line (wavelength 436 nm) and i-line (365 nm) are selected for use. The degree of pattern resolution depends also on the numerical aperture (NA) of the projector's optics. To achieve a higher degree of NA, the lens has to be made to a higher precision. Also, the depth of focus becomes smaller, making it difficult to achieve a sharp focus at the time of photo-sensitization. Due to these disadvantages, it has been agreed upon that the use of short-wave light is more advantageous, leading to the introduction of the excimer laser. As a result, the photo-exposure technology--considered to be limited in its pattern-drawing capability before, with a maximum line resolution of roughly 0.6 microns--is now being adapted to the production of next-generation integrated circuit chip sets.

Current R&D efforts are concentrated on the KrF laser, which plays a central role in first-generation laser lithography. In choosing a shorter wavelength, engineers will have to modify the optics. Ordinary excimer lasers are known for a rather large lasing spectral bandwidth. The KrF laser has a bandwidth of 0.3 nanometers or so. This can create a problem as chromatic abbrasion enters the projector lens. Even if this is compensated for by switching to a different lens material, a suitable material usually cannot be found, since the selection is limited to the following: quartz, CaF_2, and MgF_2. So, some engineers are trying a scheme of narrowing the laser emission spectra--by inserting an etalon in the lasing cavity, or incorporating a special mechanism for injection locking. For a working all-quartz projector optical system, the spectral bandwidth must be reduced to 0.003 nanometers or so. Also, the central wavelength must be kept fixed.

To be used as a light source for laser lithography, an excimer laser system has to meet still other technical requirements. One of these regards the lifetime of the laser. Originally, the excimer laser was concieved of as a tool for research scientists, and so, the ability to generate 10^9 laser pulses in a lifetime was considered quite satisfactory. But now that the laser is to be used for industrial semiconductor processing, it is a totally different ball game. Operated at a frequency of 200Hz for 8 hours a day, five days a week, the laser would last a mere 8 months. So, it is no wonder that extending the laser's lifetime has become a big issue in the industry.

In a move to prepare for the coming age of the 64-mega bit specification (smallest pattern dimension 0.25 microns), the industry is initating R&D to improve the excimer laser lithography technology. Already, with the use of an all-quartz optics system, a technique easily yielding a line-drawing resolution of up to 0.35 microns has been developed. To further enhance resolution, an ArF laser may be incorporated which can lase at an even shorter wavelength (193 nm). Also, better resist materials, ones more sensitive to laser light, may have to be developed.

3.2 Laser CVD

An emission of short wavelength light is characteristic of the excimer laser, and at the same time, this laser features a high level of output. Compared to a power density of 1 W/cm^2--measured on the surface of a waffer--given out by a high pressure mercury lamp, if the excimer laser is used instead, the density increases approximately 10^7 fold. With this capability for concentrating power, not only can the laser system expose the resist coated on a mask to the laser light, but it can also be used to directly remove [the etched portions from the emulsion] through abrasive action. Since average power is large as well, the laser system can simultaneously grow and pattern thin films directly on a waffer placed in a suitable gas atmosphere. This type of laser CVD may be applied to etching and doping processes as well.

As mentioned above, it was the laser CVD that appeared first, and its capability for outstanding spatial selectivity has been utilized in direct laser-drawing techniques. Using this method, circuit designers have grown silicon layers patterned with a line width of 0.2 microns. Also, with the aid of this technique, thin metal wirings of Al and W as well as wirings of dielectric materials such as SiO_2, and SiN_x have been successfully laid down. Recently, there have been reports of attempts to selectively grow layers of chemical compound semiconductors. Because of its low throughput and because of its limited line-drawing capabilities (a line width of 1 micron or more), the direct laser drawing technique has not enjoyed wide spread use in the semiconductor field. But, as it evolves into a mature technology, it should play a vital role in other areas like optronics in the future.

Together with laser lithography, the pattern transfer method based on laser CVD has recently caught industry's attention. In fact, there is already a report of a successful growth of a patterned thin film of aluminum. In order to raise the resolution level using this method, it is necessary to come up with a scheme for taking advantage of the photo-dissociation of the molecules absorbed into the waffer surface instead of the photo-dissociation achieved in a vapor under normal a CVD process. It is therefore strongly desirable for the advancement of basic research in the mechanics of this type of surface reaction.

3.3 Laser Plasma X-ray Lithography

As mentioned before, X-ray lithography is now heralded as a promising technology that will break the current limitations of photo-exposure technology. The soft X-ray used for this purpose ranges from 0.5 to 1.5 nanometers in wavelength, which means that the X-ray is shorter in wavelength

than photo-sensitizing ultraviolet light by a ratio factor of 1 to 100. So, theoretically, a line resolution of such quality can be offered, but in reality, since X-ray optics are yet to be developed, this cannot be realized. So, the present approach is to place the mask and the waffer side by side in close proximity in order to achieve the right exposure. Because of this roundabout setup, the exposure quality suffers more from the effects of geometric distortion (blurred images, etc., caused by the inherent physical dimensions associated with the X-ray wave) than from the Fresnel's diffraction effect.

The problem with this technology lies in the designing of an X-ray beam source. Usually, a metal target is excited by an electron beam, but the efficiency of X-ray generation achieved this way is low, ranging from 0.01 to 0.1 percent. The spot diameter of the electron beam, which, if large enough, can cause a blurring of the image, ranges from 2 to 4 millimeters, and is not small at all. To improve performance, research is underway with a setup based on the use of a hot gas plasma, whereby an efficiency rate of 0.5 to 2 percent with an X-ray beam spot diameter of 0.5 to 1 millimeter has been obtained.

On the other hand, by concentrating a pulsed light beam from a high power laser onto a metal target, plasma may be generated with an accompanying generation of an X-ray beam. The laser used has a power density of 10^{12} to 10^{15} W/cm^2, and the achieved conversion efficiency is 10 percent or more. Also, the diameter of the X-ray beam is small, ranging from a few tens of microns to several microns. Because of these superior capabilities, laser plasma X-ray lithography is attracting industry's attention.

The efficiency of generating soft X-ray's goes up as the lasing wavelength of the laser is made shorter. At the same time, the efficiency also depends on the laser pulse width: several hundred picoseconds is most suitable. Because of this, it is best to use a solid-state laser with a flashlamp excitation mechanism. Especially notable here is the invention of the zigzag slab laser. With it, a laser operation with a large output at a high operation frequency is made possible. The lasing materials found in such laser systems are Nd:YAG, Nd glass, and Nd:GGG.

Laser plasma X-ray lithography is a brand new technology, and so, there are many technical obstacles to overcome. For example, there is the problem of beam-shot particles flying apart from the target. Also, there is the problem of finding the right resist material, which by the way is a problem said to be common of any type of X-ray lithography. In order to improve the sensitivity of the emulsion, it is indispensible to develop materials with properties of inner shell electrons readily undergoing excitation.

4. Conclusion

We discussed the present status of laser microprocessing focusing on applications of this technology to semiconductor technology. Heralded as a technology which will set a new standard for the aging photo-exposure technology that has long been used in the production of integrated circuits, laser lithography is now ready to be put to industrial use. The use of X-ray

beams for emulsion sensitization is believed to greatly enhance the potential of the photo-exposure technology in general, and thus, the laser plasma X-ray lithography appears especially promising. On the other hand, laser CVD shines for its ability to make direct pattern transfers as well as its direct drawing capability.

The laser equipment which plays the central role in all of these techniques varies in specification and design depending on the type of microprocessing it is used for. Here the tasks are to identify the diverse needs that will arise in conjunction with different laser configurations, and to meet the requirements by advancing the state of laser design through innovations and modifications.

Not only will laser microprocessing work for the production of semiconductor devices, but it will also be adopted in many more areas including the fabrication of optical devices and micro machines. I shall close this article by looking forward to many innovations in this area in the near future.

High Peak Laser Plasma X-Ray Source

43064082b Tokyo OPTRONICS in Japanese Oct 88 pp 91-96

[Article by Kazuo Tanaka, Osaka University: "High Peak Laser Plasma X-Ray Source and Its Applications in X-Ray Lithography"]

[Text] 1. Introduction

The idea of making plasma using a laser is not widely known among the general public. This is not so surprising because after all, major research endeavors in this area have been limited to nulear fusion based on laser inertial confinement. But as a spin-off, these mammoth science projects spread valuable information regarding the characteristics of the X-ray radiation emitted from laser-generated plasma to many, diverse fields of science and engineering, prompting the design of applications for X-ray lasers, X-ray lithography, light sources for soft X-ray microscopes, XFAFS (extended X-ray absorption fine structure), etc.

The above-mentioned laser facility for nuclear fusion refers to a giant laser system with the capability of generating extremely strong multiple laser beams, which is housed in a huge gymnasium-like clean room. Obviously, it is not a machine that anyone can readily use. However, recent developments in university labs indicate table-top compact lasers are on their way into commercial production. These experimental lasers remain highly efficient even after very frequent use, and are still powerful enough to generate X-ray radiation. Incidentally, in the United States, there are a number of companies eyeing the commercial production of such an easy-to-use table-top laser system. One company actually embarked on R&D ten years ago, and since then, it has been working exclusively on the new X-ray laser.

It thus appears that industry is increasingly interested in laser plasma X-ray (shortened to LAPLAX). But what actually is this LAPLAX? How is this different from other X-ray sources (i.e., synchrotron orbit radiation)? We

shall discuss these points in detail below.

First of all, let us talk about X-ray lithography which is one of the major applications of LAPLAX.

There have been reports of photo-resist materials entering the commercial market which are designed for the excimer laser (having a laser line of around 250 nanometers), that are going to make it possible to start mass production of chip sets with a patterning resolution of 0.5 microns (5000 angstrom). But, the number that designers of X-ray lithography are aiming at is actually still smaller than that. The technology is not, however, mature enough to achieve this as there are many problems left unsolved. In any case, it's not too early to consider possible applications of LAPLAX in X-ray lithography, and indeed, various proposals and designs have been made. So, now I would like to discuss the properties of LAPLAX, and its potential benefits to X-ray lithography.

2. LAPLAX (Laser Plasma X-ray)

2.1 X-ray Generated by Plasma

In reality, how does an X-ray appear from a plasma produced by a laser? Laser light is electromagnetic radiation. X-rays are also electromagnetic radiation, but they have a shorter wavelength. In other words, you might say that with plasma as somewhat of a catalyst, the visible light gets transformed into X-rays through the working of the right setup, and that the plasma acts like a wavelength converter.

One important fact that should not be missed is the fact that laser light is different from ordinary light (such as the light from the sun), and that it is in fact a coherent beam of extremely monochromatic light (the typical sprectral spread in a laser beam is 1 angstrom or 10^{-8} centimeters or less). On the other hand, radiated X-rays usually have a spectral spread of 100 angstrom or more, and are not phase-coherent. But, with the right combination of plasma parameters and the laser, a special kind of X-ray radiation called an X-ray laser may be generated. In this case, the radiated X-ray itself assumes the profile of a laser having a matched phase and a very small spectral spread. In either one of the above cases, the plasma is used as an intermediary agent to help convert laser light into X-rays through frequency shifting.

A laser beam with a thickness of 0.1 cm to 80 cm is emitted from a laser amplifier. The size of the beam becomes larger as the power of the laser is increased. The cited figure of 80 cm is a rather extreme example: it refers to the beam size of the world's largest laser facility for inertial confinement of nuclear fusion, known as "NOVA", in the U.S.'s Lawrence Livermore National Laboratory. This laser system is capable of generating a laser beam of several kilo-joules with a pulse width of 1 nanosecond (10^{-9} seconds). By focusing this type of laser light on a spot using a lens, one can obtain a beam with all component waves nearly perfectly aligned in parallel, and the beam can be made to focus on a spot as small as the lens' diffraction limit, which is actually determined by the lens quality. For

example, with a 10-joule laser light having a 1-nanosecond pulse width, it is an easy job to focus the beam on a 100-micron wide spot. With this concentration of power, the intensity of the electric field per unit area becomes extremely large:

$$I = \frac{10\ [J]}{1\times10^{-9}[sec] \times pi\ (50\times10^{-4})^2} = 1.3\times10^{14}\ [W/cm^2]$$

When subjected to an electrical field as large as $10^{10}W/cm^2$, electrons are torn apart from the atoms that form the target material, generating plasma. The electrons transfer the energy absorbed from the laser to the atoms that make up the matter through collision with ionized atoms. This transmission of energy is performed at a speed of 1/10th or less of the speed of sound travelling in matter. We have to skip the details of the mechanism of energy transfer since it's too involved. Anyway, X-ray and shock [waves] (the laser plasma pressure is extremely high: a very high pressure of 10Mbar (140,000 atmospheres) or more is generated within the spot) are also contributing factors in this energy transfer mechanism.

In a plasma state with ions and electrons forming separate plasma clouds, X-rays are radiated for various reasons. A free electron that travels freely inside the plasma, when attracted by an ion, changes its course of orbital flight generating a special type of X-ray called braking radiation. Because of the nature of the emission mechanism--invoving many electrons moving at different speeds and in different directions inside the plasma--the generated X-ray spectra is continuous. X-ray radiation with this property is called free-to-free transition X-ray.

In another mechanism, free electrons are captured by atoms and bound to atomic orbitals. The electron's leftover kinetic and thermal energy is converted to X-ray radiation. This process is called free-to-bound transition. Although this radiation is essentially a continuous spectrum, because of the presense of numerous electron orbitals (the figure depends on the primary quantum number), and because the electrons in these orbitals carry different energies, the X-ray spectrum assumes a non-continuous sawtooth-like profile.

There is one more type of X-ray emission. This process is called bound-to-bound transition whereby the X-ray emission is achieved via the up and down motion of the remaining electrons between the abovementioned orbitals inside the atoms. Since the wavelength of the X-ray is predetermined by the value of the energy gap between the transition orbitals, the emitted spectrum is narrow, and is called normal-line spectrum. For instance, if an aluminum target is made into a plasma using a 10^{14}-W/cm^2 or so of laser power, 10 percent of the generated X-rays are due to the free-to-free transition; 60 percent to the free-to-bound transition; and 20 percent to the bound-to-bound transition, according to a computer calculation. Figure 1 is a plot of the measured spectrums for an X-ray energy (h*nu) range of 0 to 1.5 keV, obtained from an experiment with a gold target. All of the curves represent plasma-generated X-ray radiation produced by a laser with a power of $7\times10^{13}W/cm^2$. The thick curve represents the experimental data for a laser

line wavelength (lamda) of 260 nanometers, the dotted line represents lamda=530nm, and the dashed line represents lamda=1050nm. The figure shows that the laser line with the shortest wavelength, 260 nanometers, which is in the vacuum ultraviolet region has the highest soft-X-ray conversion efficiency. This particular experiment was sensational in that, in fact, 80 percent of the laser input energy was transformed into soft X-ray radiation. In other words, 8J of the 10J of laser energy was radiated out at a solid angle of 2*(pi) from the focused spot of roughly 100 microns wide. Actually, the intensity of the X-ray emission was extremely short having a radiation duration on the order of one nanosecond. This fact led to the very large energy value. In Figure 2, comparison is made of the emission intensity for various X-ray generation methods. According to this figure, LAPLAX gives an incomparably high peak value compared to the synchrotron orbit radiation method. Obviously the two techniques are different: while the latter method produces a nearly continuous spectrum of radiation, the total energy of the LAPLAX radiation is predetermined by the laser pulse frequency. At this writing, research is still active on the development of a 30- to 60-watt compact laser. Also, at the frontline of research and development, working table-top laser systems for 1J, 1ps, and 0.1Hz are already here.

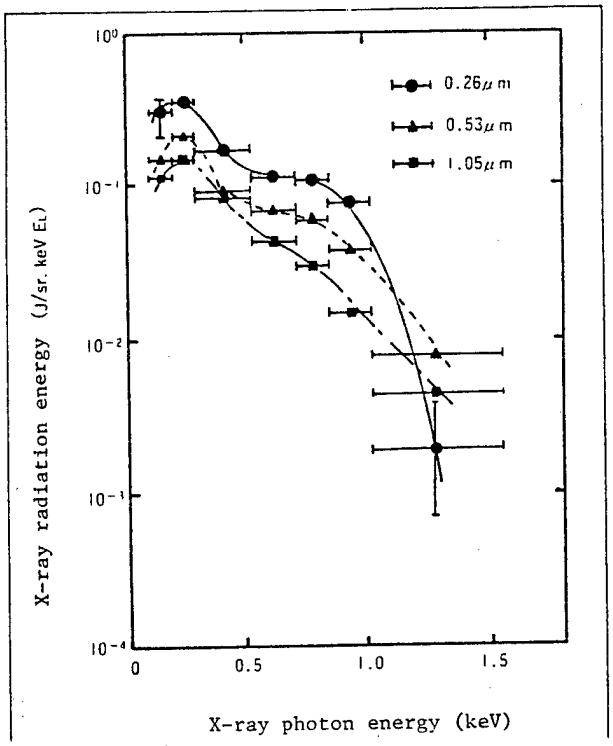


Figure 1. Soft X-ray spectra (h*(nu)=0-1.5keV) from an Au target. The solid curve indicates the spectra for a laser line of 260 nm, the dotted line indicates one of 530 nm, and the dashed line indicates 1054 nm. The pulse width is 400 psec, and the laser beam intensity $7X10^{13}$ w/cm^2.

The vertical scale is expressed as the amount of energy (J) per unit solid angle (Sr) X unit photon energy (keV) X unit laser energy (E_L).

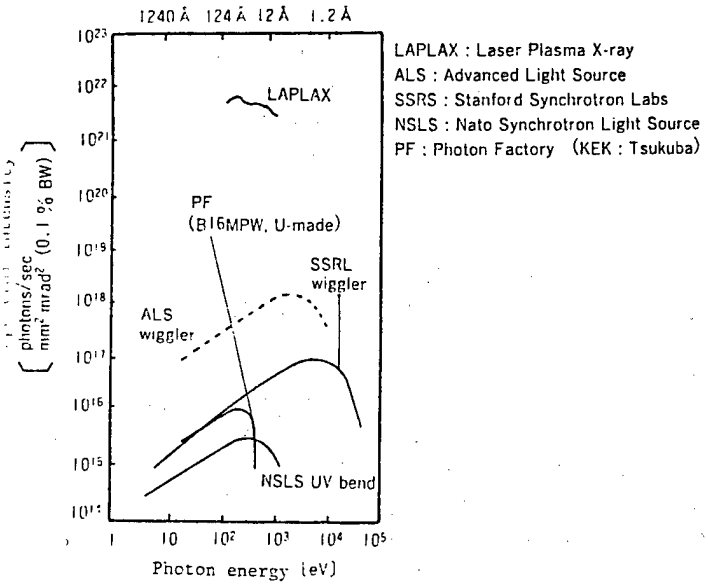


Figure 2. Intensity of the X-ray emission spectra for different X-ray sources $[\frac{\text{Number of photons/sec}}{\text{mm}^2\text{mrad}^2\ (0.1\%\text{B.W.})}]$

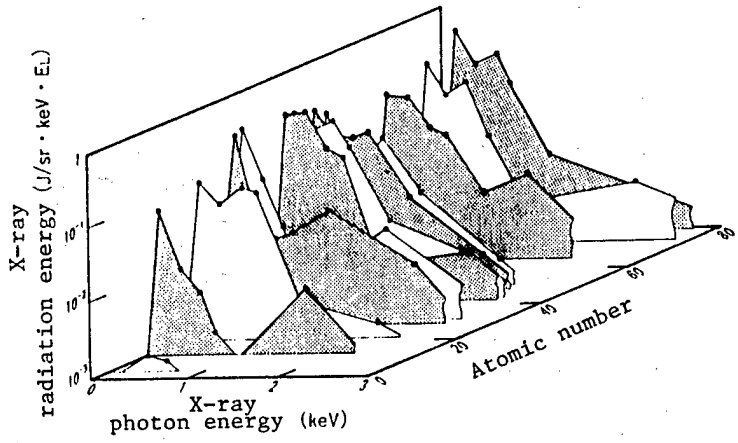


Figure 3. Dependence of the X-ray spectral intensity on the atomic number; laser line 527 nm, laser beam intensity 10^{14} W/cm^2, laser pulse width 1 nsec.

2.2 Features of LAPLAX

The conversion efficiency and spectrum of the LAPLAX's laser energy may be controlled by varying the parameters (i.e., the kind of target material, the material quality, the pulse width, and the wavelength and intensity of the laser line). Let me elaborate on this by citing a few specific examples.

2.2.1 Dependence on Atomic Number

Figure 3 is a plot of X-ray conversion efficiency versus photon energy (keV) for various target materials ranging from a plastic (CH) with Z=3.5 to gold (Au) with Z=79 under the same conditions of laser irradiation (laser intensity 10^{14} W/cm^2, pulse width 1 nsec, and laser line 530 nm). The photon energy ranges from 0 to 3 keV indicating that the region considered represents the soft X-ray spectrum. The areas under the plotted spectral distribution curves are alternately shaded and left blank to avoid confusion in locating the specific spectral profile for the atom to be investigated. According to the plot, the trend is that as the atomic number increases so does the conversion efficiency. Some spectral profiles have 2 or even 3 peaks. This occurs because the abovementioned main transition process changes due to an interaction with photon energy. For these spectral peaks, the full width at half maximum was read and traced onto Figure 4. According to the figure, for photon energy, say, of 2 keV, and for atomic number 40 on the vertical scale, electronic transitions primarily responsible for the spectral peak in question involve the L shell. From this graph, one can select the kind of atom, and hence, the kind of material that's most suitable for efficient laser-to-X-ray conversion to obtain the desired X-ray.

2.2.2 Dependence on Plasma Volume

As the irradiated volume of LAPLAX is increased, the amount of X-rays generated increases. To verify this fact, an experiment was conducted where the target was irradiated with laser pulses twice in succession. The initial pulse was weak (usually 1/10th or less of the intensity of the main pulse that followed), and formed a plasma over the target material. After sufficient thermal expansion of the plasma (approximately 1 nanosecond) was obtained, the main pulse was delivered. Through this process, the amount of the generated soft X-ray with an energy level greater than 3 keV can be increased roughly 3 fold, at the most, over that of a single pulse process. This information is invaluable for later design of applications for lithographic technology.

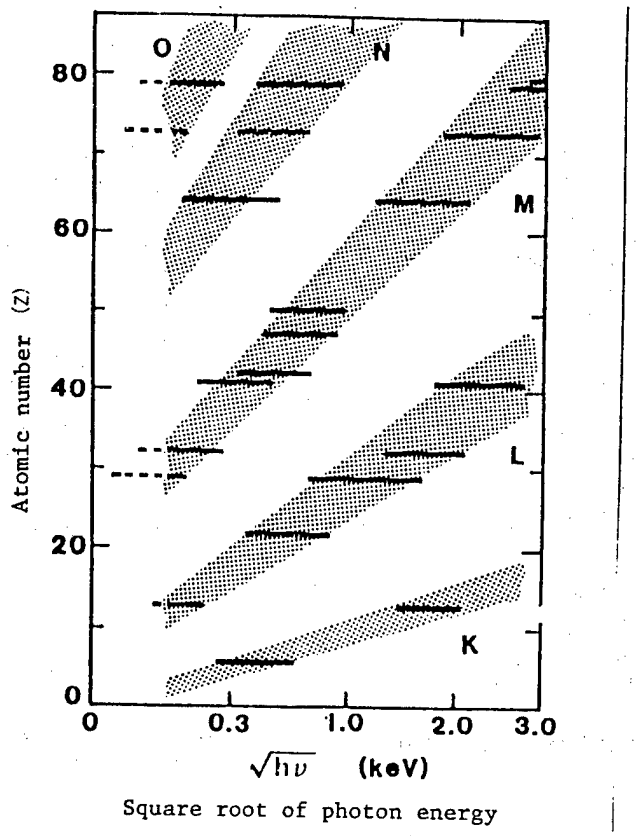


Figure 4. Relationship between the position of X-ray spectral peak (FMHW) and the atomic number. The horizontal solid bars indicate the position of the spectral peaks for different elements. The shaded areas indicate the corresponding electronic shells O, N, M, L, K.

The experimental setup is the same as that in Figure 3.

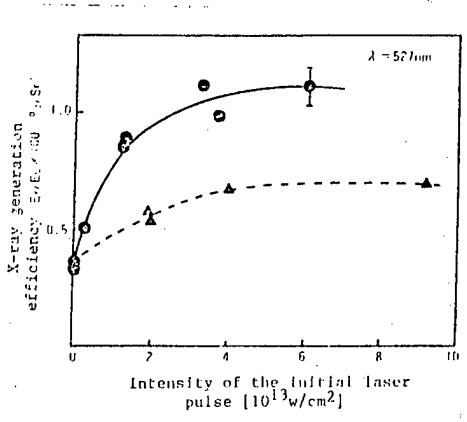


Figure 5. Efficiency of X-ray generation using a copper target (photon energy, h*(nu)=1 to 3keV)

The double pulse setup (the initial pulse creates a large plasma which is later irradiated by the main pulse) produced an X-ray generation efficiency roughly 3 times higher than the single pulse one (corresponding to the position indicated by 0 on the horizontal scale). E_X is the X-ray energy, and E_L the laser energy.

2.2.3 Dependence on Laser Line Wavelength

Dependence on laser line wavelength is as shown in Figure 1. It may be observed from the figure that by using a shorter wave laser, conversion efficiency can be improved. The reason is as follows:

A laser beam incident on a plasma is reflected if the plasma density exceeds a certain level, or if the plasma density is exactly at the threshold level--known as the critical density--the beam is strongly absorbed by the plasma. The critical density is directly proportional to the square reciprocal of the laser line wavelength. In other words, a small increase in the wavelength of the laser line causes a great increase in critical density. For example, the density value for a 260-nanometer laser line is 16 times greater than that of a 1054-nanomemter one, and there is a corresponding difference in absorption of laser energy. But, given the same absorption energy, [the plasma] having the higher density contains a greater number of atoms which need to be heated by means of the absorption of laser energy, and this results in a lowering of the plasma temeperature. Hence, there will be an increase in the relatively low temperature (100eV to 1keV) emission of soft X-rays.

3. Application of LAPLAX in X-ray Lithography

3.1 Introduction to X-ray Lithography

At present, the X-ray lithography system based on the use of LAPLAX is still at the developmental stage. This is advantageous from the point of view of system development since a lot of unrestrained innovative designs can be poured into it. Many researchers are already involved in photo-exposure experiments, producing satisfactory results with the expected sharp patterning resolution (delta-X approximately 1000 angstrom) (refer to Figure 6[8]). It is believed that development of a suitable photo-resist, more sensitive to LAPLAX, will be underway shortly. The experiment listed under reference 8 used FBM-120. Therefore, researchers also concluded that the spectral component responsible for photo-sensitization is mainly an X-ray with 1 keV or less. As for the development of responsive X-ray source materials, it is best to stick to target materials with a high atomic number such as gold since these materials have a high soft-X-ray conversion efficiency and are capable of stronger emissions of X-ray spectra with 1 keV or less energy.

3.2 Proposed Designs

Studies to evaluate the technique of direct irradiation using a system setup with the mask and resist placed closely side by side have been conducted to a certain degree. FBM-120 based resist materials rated for a $10mJ/cm^2$ sensitivity are successfully sensitized in a setup where the laser exposure lasts for 10 seconds at 50W with the mask and the LAPLAX separated by 20 centimeters. The exposure time is expressed as:

$$T = \frac{2(pi)L^2S}{(eta)_x W_L} \quad (msec),$$

where L is the distance between the LAPLAX and the mask, S the sensitivity (mJ/cm^2) of the resist material, $(eta)_x$ the laser-LAPLAX conversion efficiency, and W_L the laser's output power.

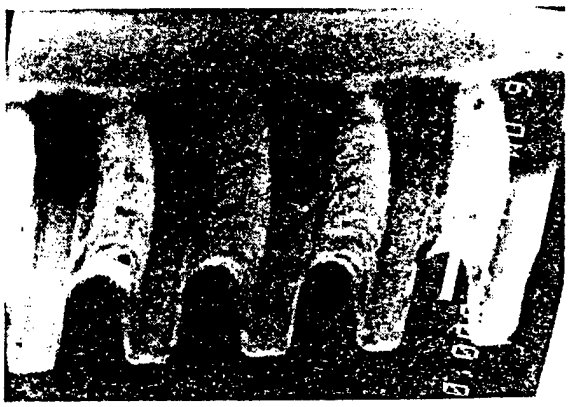


Figure 6. A surface processed by a LAPLAX-based X-ray lithography; resist material FBM-120, X-ray mask 8 microns thick, Au lattice structure, line width 5000 angstrom, periodicity 1 micron.

The distance between the mask and the resist, and the distance between the LAPLAX and the mask directly affects the patterning resolution and the extent of blurring of the image. In order to hold the blurring effect to less than 0.2 microns, the mask-resist distance has to be reduced to less than 50 microns.

With this type of setup, where the mask and resist materials are supposedly placed inside the same vacuum chamber as the target, there is the problem of the mask getting damaged by plasma and neutrons flying apart from the target at the time of laser irradiation. The use of a buffer gas to fill the vacuum chamber and act as a partition has been proposed. But I feel that there can be more improvements, especially in the generation of plasma resulting from the reaction of the gas to the laser, and in the design of a better scheme for loading and unloading the mask and resist materials.

The proximity setup discussed above may follow the scheme shown in Figure 7.

The point of this scheme is to make the handling task easier, and to prevent free flying particles from coming out of the plasma and damaging the mask and resist materials: the vacuum chamber where the target plasma is kept and where LAPLAX is generated is made separately from the mask/resist sensitization unit. In this setup, the desired X-rays are screened using an X-ray filter, and the exposure of the resist to X-rays is performed in an environment of air, and helium (He) or another light gas. With the attenuation at the X-ray level taken into account, there shouldn't be any problem in computing the proper distance for separating the X-ray filter and

the mask. For instance, by using He gas (held at 1 atmosphere), an X-ray with a photon energy h*(nu)=1keV has its energy level reduced to half, which means that the mask-resist cassette should be positioned 60 centimeters away from the X-ray filter--the actual positioning is done via adjustment of the cassette rotor assembly.

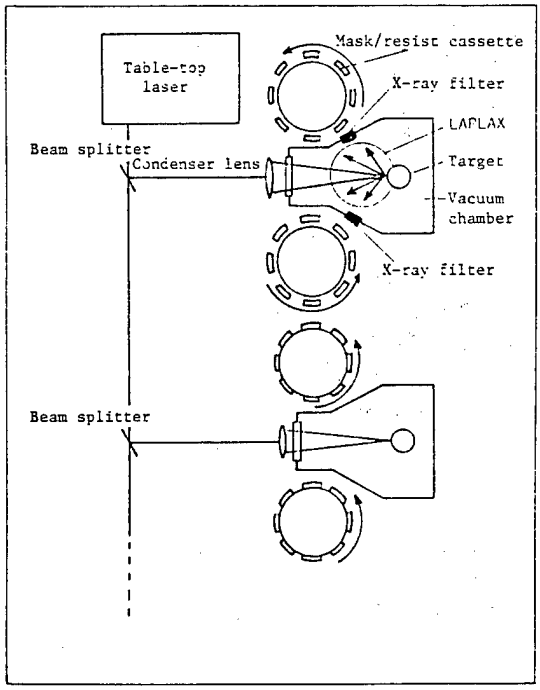


Figure 7. Proximity-type lithography system based on LAPLAX

In this system, the setting of the unsensitized mask/resist is automatic: an exposed waffer appears at each swing of the rotor. This is good news for mass production. Also, it's possible to install multiple rotor units for one laser beam: one for each side of the chamber and two more for the upper and lower sides. Furthermore, by using a commonly used filter material such as K-edge, the unwanted component of the LAPLAX is automatically cut off resulting in an improvement in the resolution.

To increase the number of design options, methods other than the one discussed above need to be studied. For example, by incorporating the (Schwartz-Schulz) soft X-ray microscope that's undergoing a lot of research, the [emission] can be achieved at a greater solid angle.

Also, by using the abovementioned microscope, reduced-size projections can be performed: the precision pattern on a carefully prepared enlarged mask is traced with a spatial resolution limited only by that of the microscope's own. According to calculations, this microscope can have a resolution as powerful as 500 angstrom. No wonder new developments are being carefully watched.

Devices like this microscope, which can be used to collimate LAPLAX, have great practical value considering the magnitude of possible applications of X-ray spectra radiated at a solid angle of nearly 2*(pi).

4. Conclusion

Studies of LAPLAX have been carried out as part of the research for inertial confinement nuclear fusion, and detailed characteristics of this phenomenon are unravelling, albeit this may not be apparent to ordinary people. As our knowledge of its outstanding properties, especially the highly peaked spectra discussed in the above sections, has deepened, we are now discovering diverse possibilities for industrial applications of LAPLAX. Our discussion in this article focused on applications in X-ray lithography, and considered possible setups to achieve a working system.

Of course, in order to create a commercially viable system, the development of an easy-to-handle laser which is capable of high frequency operation is a prerequisite. But, recent research developments in this area indicate that we are indeed very close. Finally, I would like to conclude my discussion by saying that LAPLAX is going to make a simple yet very powerful X-ray source for the coming decades, full of potential for diverse applications.

Excimer Laser Lithography

43064082c Tokyo OPTRONICS in Japanese Oct 88 pp 97-102

[Article by Kazufumi Ogawa, Matsushita Electric Industrial Co., Ltd.]

[Text] 1. Introduction

With the recent breakthroughs in excimer laser technology, the technology of photolithography based on the excimer laser is suddenly becoming very attractive as a means of achieving the ultra fine line resolution that no other source of light beam can offer. In fact, with the excimer laser-based lithography, one can go beyond photolithography's previously accepted resolution limit of 0.5 microns.

Even before the advent of the excimer laser, there was already a need to break this resolution limit, and for this purpose, the electron-beam (EB) exposure and X-ray exposure techniques were considered. But problems abounded: The EB method can make the micro-patterning job easier, but has poor throughput. As for the X-ray method, a source that can produce the needed high power spectra is lacking. Also, the sensitivities of available resist materials are not up to the desired level. Furthermore, the production of masks involves very complex steps.

Then, along came the excimer laser-based exposure technique. Because of the use of an extremely short wavelength spectra, the new technique can offer an incredible resolution compared to the conventional stepper which uses the

g-line or i-line of the spectrum from a super high-pressure mercury lamp. For example, the currently available options for selectable wavelengths in an excimer laser are 193 nanometers (ArF), 248 nm (KrF), and 308 nm (XeCl). With an ArF laser, the maximum resolution can be brought up to about 0.25 microns. Also, the image-reducing projected photo-exposure technique using an optical lens can be used. Furthermore, it offers the advantage of ease of mask production and high throughput. These factors make the excimer laser-based lithography a very promising technology for submicron-level microprocessing needs.

At present, there are basically two methods available for realization of the excimer laser photo-exposure system:

One of them is a technique carried over from the conventional stepper with the same chromatic abrasion-corrected image-reducing lens, but its source of exposure light is not a super high-pressure mercury lamp, but an excimer laser. There is already a working model based on this method. Developed by G.M. Dubroeuco, it incorporates a KrF laser and object lens. This system does offer a greater degree of freedom in the lasing spectral width, but it makes it difficult to design the lens system for large-area exposure needs. It thus appears very difficult, at this time, to create a practical system which can handle 15 X 15 mm^2, the integration size required for a 16-megabit DRAM memory chip set.

The other method consists of an image-reducing lens designed for monochromatic spectra, and an excimer laser as the exposing light source which generates a narrow band near the monochromatic lasing spectra. A working system developed by V. Pol et al incorporates a narrow-band lasing spectra KrF laser and monochromatic lenses made of quartz. Although there is the possibility of the generation of coherent noise (called speckles), use of this setup can greatly simplify the design of the lens system, and a resolution on the order of half a micron for large exposure areas can be achieved. Table 1 lists the specifications for the KrF excimer laser steppers which have been developed so far.

The photo-exposure system alone cannot create the lithography; it has to be accompanied by equally sensitive resist materials--precision resist processing technology. With the advent of a new concept exposure system, a new type of resist is needed.

Although there are very outstanding emulsions--high performance positive resists of a (Nobolak type)--for use in conventional steppers, when it comes to the resist material for the excimer laser system, there is no practical emulsion material at this time. Intensified research efforts are underway, however.

What is being called for in the resist material for 0.5-micron line patterns is the capacity for a high level of throughput in forming the micro pattern on waffers. In other words, the emulsion has to score high marks in photo-sensitivity, image resolution, resistance to dry etching, and adherence to the substrate.

Table 1. Specifications for Excimer Laser Stepper by Different Manufacturers

Item / Research Institutions	NA reducing ratio	Side of exposed square area (mm)	Lens	FWHM of laser line (nm)	Exposure illuminance (mJ/cm^2-pulse)	Illuminator system
AT&T Bell Labs	0.38 5:1	10	No color correction	0.004	0.4	Injection synchronization & scan sys.
Toshiba	0.37 10:1	5	Color corrected	-	20-40	-
Matsushita	0.36 5:1	15	No color correction	0.007	8	Etalons used
Nikon	0.37 10:1	5	Color corrected	-	-	-
Canon	0.35 5:1	4	No color correction	0.002 or less	-	Injection synchronization & fly's eye lens

Actually, there are few photoemulsions available in the market which may be used with a KrF laser (248.4nm)--an excimer laser. These are actually some of the positive resists used in conventional deep-UV systems. But, considering their compatibility with the alignment light and exposure specifications, they still need to have their absorption wavelength characteristics modified. For instance, popular (naphthoquinone-diazide) photoemulsions have a broad absorption spectrum. Therefore, without any modifications to the material, the alignment wavelength has to be shifted away from the photo-exposure wavelength by a significantly large margin, which means that the alignment precision can worsen. Also, with existing resists, one cannot gain good line resolution when using a KrF excimer laser. This is what's prompting development efforts in positive-type and negative-type emulsions.

In the following section, I would like to discuss the current status of the excimer laser lithography with a special focus on the workings of our in-house developed KrF excimer laser stepper which is capable of photosensitizing a $15X15mm^2$ area. Also, we will study photosensitization experiments using positive resists recently developed specifically for excimer laser lithography.

2. KrF Excimer Laser Stepper

In general-purpose KrF lasers, the normal lasing spectra's FWHM (full width at half maximum) is roughly plus or mimus 0.2 to 0.3 nanometers. As already mentioned, the laser spectral bandwidth has to be made narrow enough to make it a monochromatic spectra. What is also important when using a light source with good coherency is the means of preventing generation of speckles while achieving the needed monochromatic property without causing a drop in number of the lasing mode.

There are excimer lasers which incorporate an injection locking system--a method for monochromatic lasing--and wavelength selection devices (etalons). Although it's not perfect, having the disadvantage of an extensive lowering of the power density coupled with a drop in the number of modes, the injection locking model does offer better laser beam coherency. On the other hand, by using etalons, the drop in power density and the number of modes can be controlled, but the etalons can contribute to a reduction in laser coherency.

Having weighed these facts, we adopted the etalon method in our lab to achieve a narrow bandwidth in the excimer laser spectra. We successfully obtained a narrow spectra without the speckle generation.

Figure 1 is a photograph of an overall view of the experimental KrF excimer laser stepper we developed. The installation space, including the space for the laser equipment, took two to three times the space needed for a conventional stepper system. The laser system is separated from the stepper assembly (consisting of an image-reducing projector lenses, alignment optics, and XYZ stages) since each rests on separate vibration-free tables. As for the etalons, one is placed inside the laser cavity, and another in front of the cavity. After the laser beam is tuned to a narrow bandwidth, it is bent by 2 reflector mirrors and directed to the integrator where it is amplified and homogenized. After that, the beam is reflected off a large total-reflection mirror to a quartz condensor lens, and is focused on the quartz mask (reticle). Through this process, the reticle pattern is then projected onto the waffer fixed on the stage through the image-reducing projector's quartz lens.

The laser used was a compact model specially designed for stepper use, based on the electric-discharge excitation method. The laser head is made of ceramics, and is separated from the gas blower to achieve ease of replacement. It has roughly 1/4th the size of a standard model. The power is rated for 100Hz, 8W, and 40mJ. As for the life, with a power degradation of 50 percent regarded as the lifetime limit, our laser has approximately 10^9 pulses. Also, with the two etalons in use, the FMHW in the spectra is reduced to 0.005nm or less; taking the fluctuations into account, the value becomes 0.007nm. The observed spectra of the laser is shown in Figure 2. The plot indicates a good possibility of using projection lenses for designing a monochromatic system. Table 2 lists the specifications applicable to this laser system.

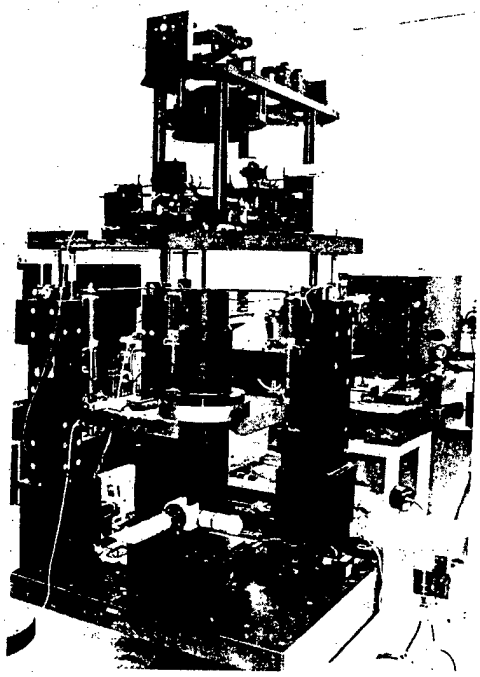


Figure 1. Snapshot of a KrF excimer laser

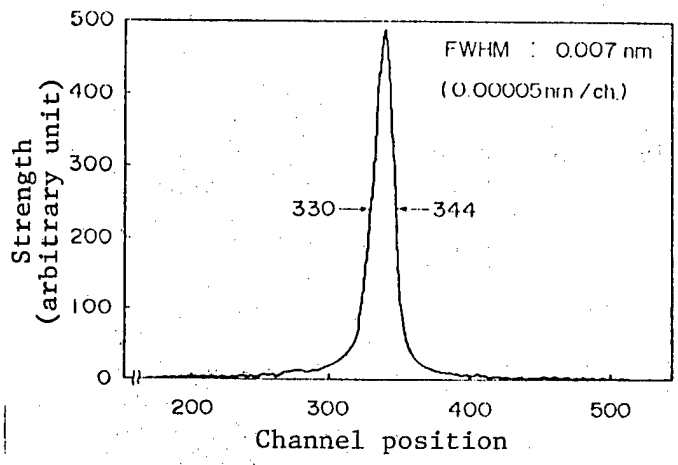


Figure 2. Spectra from a KrF excimer laser for a photo-exposure system

Table 2. Specifications for a Photo-Exposure KrF Excimer Laser

Wavelength (KrF)	248.4 nm
FWHM	0.007 nm
FWHM stability	±0.001 nm
Lasing frequency	max. 200 pps
Pulse energy	40 mJ
Pulse energy stability	±5 percent
Output power	max. 8 W
Output power stability	±2 percent
Lifetime of gas	$3X10^9$ pulses
Dimensions of laser head	300 X 500 X 1000 mm
Gas replacement	Continuous replacement possible

The present level of technology available cannot provide us with the means to control the stability of the laser output power; typically, the power fluctuates by plus or minus 5 percent for each pulse delivered. This fact led us to believe that it would be impossible to regulate the amount of the sensitization light under the scheme of a single-pulse exposure. So, what we ended up doing was to readjust the power output in such a way that the sufficient amount of exposure light was secured using 10 pulses per chip. This way, we succeeded in holding the exposure fluctuation to within 0.5 to 1 percent. By the way, under these conditions of usage, the lifetime of the gas is sufficiently long, and therefore, the advantages of the excimer laser are not lost.

The lens used for image-reducing projection is made of pure quartz without any choromatic adjustment. Its specifications are shown in Table 3. The numerical aperture (NA) is 0.36. The image reduction ratio 5:1. The exposure area $15X15mm^2$. The NA value of 0.36 was used to take into account the FWHM limit of 0.007nm within which our test laser was stable. We ran simulations varying FWHM from 0, 0.0035, 0.007, to 0.32 nanometers to find the respective OTF (optical transfer function; contrast in spatial imaging) characteristics. The results are plotted in Figure 3. Also, Figure 4 illustrates the results for the spatial imaging contrast obtained under the simulated conditions. According to these results, the narrow bandwidth laser (FWHM=0.007nm) gives a better resolution when compared to the normal laser (FWHM=0.32nm).

With regard to the lens with an NA=0.36, as shown in Figure 3, the OTF becomes particularly bad for a FWHM=0.32nm (standard laser), but if FWHM is increased to 0.0035nm, it comes very close to the value for FWHM=0nm (theoretical monochromatic laser).

Table 3. Specifications for the Image-Reducing Projector Lens in a KrF Excimer Laser Stepper

Method	Image-reducing projection
Magnification	5 X
Exposed area	15 X 15 mm^2 (dia. 21.2 mm)
NA	0.36
Wavelength	248.4 nm (KrF)
FWHM	0.007 nm
Lens material	Quartz
Length of lens	675.5 mm

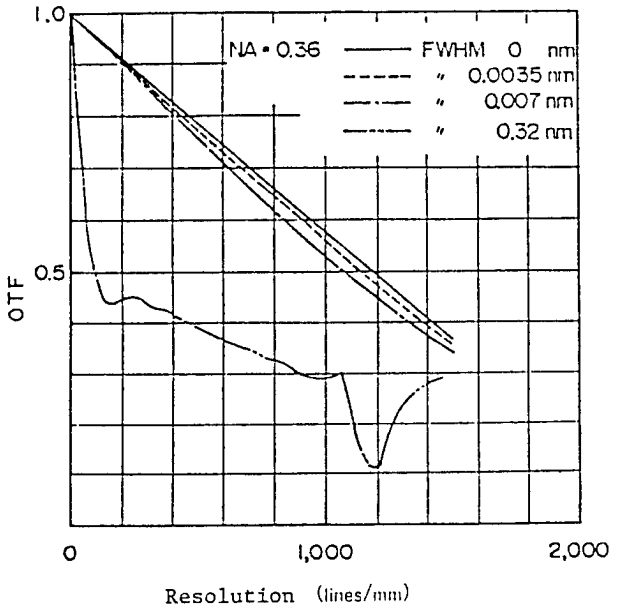


Figure 3. Plot of OTF from simulations with an image-reducing projector lens (NA=0.36)

3. Exposure Experiment

To evaluate our new lithographic system, we conducted photosensitization tests using our custom-designed resist (STAR-P2) and the commerically available positive resist (MP2400; Shipley Co., Ltd.) which is for deep UV exposure. Table 4 lists the processing conditions used in the experiment. The amount of exposure light chosen for the STAR-P2 and the MP2400 are roughly, 30 pulses and 10 pulses (respectively) with the exposure illumination strength fixed at $8mJ/cm^2$ per pulse. What this means is that the exposure time per chip runs 0.1 to 0.5 seconds. In other words, with the right peripheral systems coming in the future, we can expect a throughput of 30 to 40 waffers per hour for the job of masking 6-inch waffers. Incidentally, there were no speckles observed under the exposure conditions we employed. Refer to the pictures shown in Figures 5 and 6 for comparison of the pattern resolutions obtained with STAR-P2 and MP2400. As far as the level of pattern resolution is concerned, both resist materials gave a line-and-space (L&S) pattern resolution of 0.5 microns--measured at the center of the chip

surface--throughout the chip (15 by 15 mm^2, 0.4 microns thick). But from the pictures, it is obvious that STAR-P2 outperformed the other in terms of [the spatial contrast of] patterned shapes. Also, on the gamma characteristics (Figure 7), STAR-P2 scored better than MP2400. Figure 8 is a plot of the ultraviolet spectral characteristics before and after the KrF excimer laser irradiation for STAR-P2 and MP2400. With the resist coating of 1.0 microns thick, our excimer laser beam achieved sufficient exposure of the resist. As you can see from Figure 8, the post-exposure transmittance for the near-248 nanometer spectra has a better value for STAR-P2 than for MP2400. These results indicate that STAR-P2 is a quite satisfactory performer. The badly patterned shapes of the MP2400 were caused by the fact that the illuminating beam did not penetrate deep enough into the resist layer since most of it ended up being absorbed in the top surface. In other words, it's very likely that the unwanted absorption by the base polymer of the resist is causing the MP2400's poor spatial contrast.

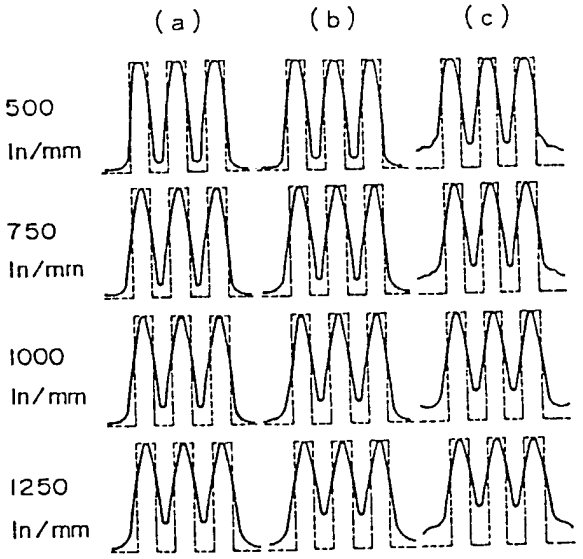


Figure 4. Plot of the spatial contrast from simulations
(a) FWHM 0.000 nm (monochrome)
(b) FWHM 0.007 nm (narrow band KrF laser)
(c) FWHM 0.32 nm (normal band KrF laser)

Table 4. Parameters for Single-Layer Resist and Double-Layer Resist Processes

Type of process	Resist coating	Prebake	KrF excimer laser exposure	Development time	O_2 RIE
Mono-layer resist process	STAR-P2 1.0 μm	90°C 120 sec	30 pulses	60 sec	-
Mono-layer resist process	MP2400 0.6 μm	90°C 120 sec	10 pulses	60 sec	-
Double-layer resist process	(1) BL RG3900B 1.8 μm (2) TL SNR 0.2 μm	(1) BL 240°C 20 min (2) TL 80°C 20 min	900 pulses	20 sec DIBK 90% ECH 10%	O_2 Gas: 20×10^{-3}Torr Power density: 80mW/cm^2 Bias: 100V Gap: 40mm Etching time: 17+Over 2min Etching rate: 1250ang./min

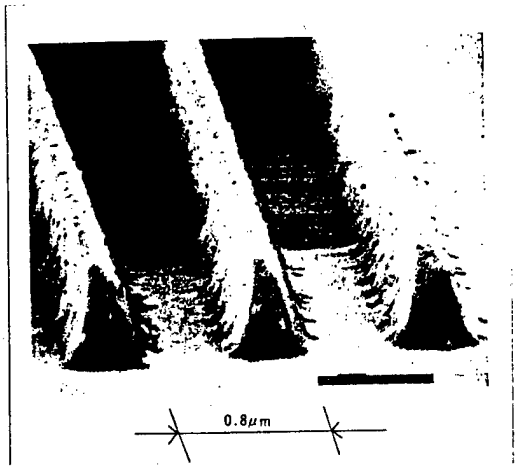


Figure 5. SEM photograph of a photo-sensitized pattern using a KrF excimer laser stepper (resist: STAR-P2, coating thickness: 1.0 micron)

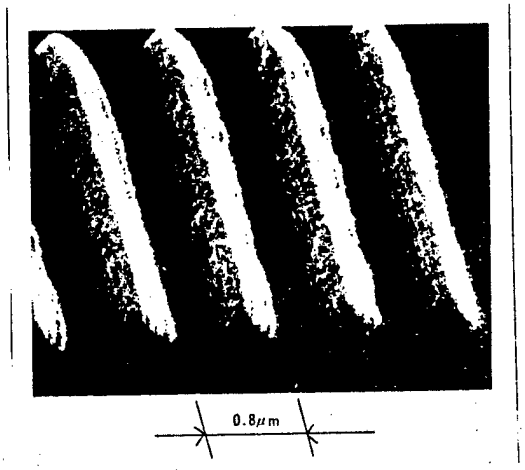


Figure 6. SEM photograph of a photo-sensitized pattern using a KrF excimer laser stepper (resist: MP2400, coating thickness: 0.6 microns)

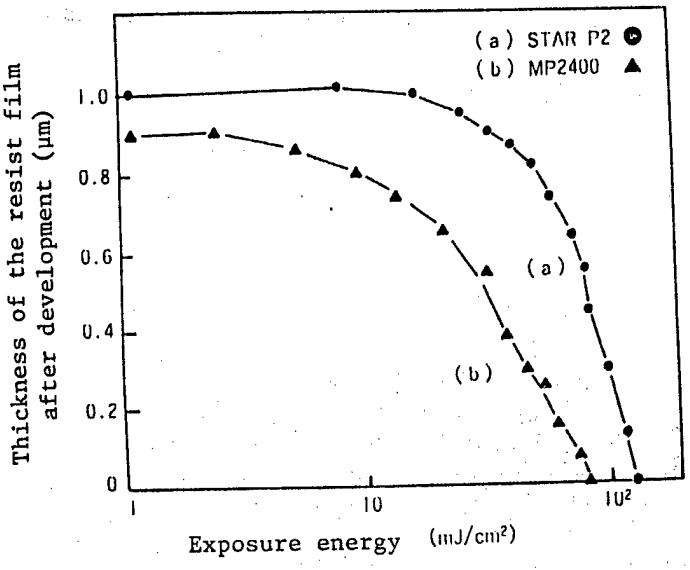


Figure 7. (Gamma) characteristics for STAR-P2 and MP2400, exposed with a KrF excimer laser

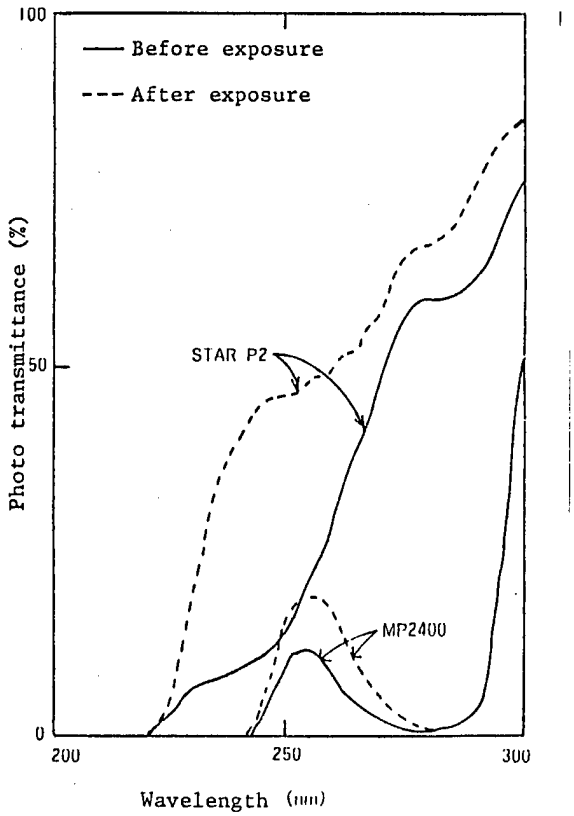


Figure 8. Changes in the spectral transmittance characteristics of STAR-P2 and MP2400 before and after the exposure using a KrF excimer laser

Besides the modifications to photo-resists, there is an alternative way to improve the L&S resolution--through the use of improved resist processing technology, e.g., a technique to grow thinner resist films. Making thinner films naturally improves resolution, but can cause a lowering of strength of resistance to etching. There is a way to circumvent this problem; this is the 3-layer resist process proposed by J. Moran et al. The technical basis for this technique has already been well established.

However, the 3-layer resist process consists of too many steps to be used in the design of a practical semiconductor device production system. What we studied instead is a bi-layer resist (BLR) process with a resist material containing silicon in its use.

Figure 9 illustrates the steps used in the BLR process, and Table 4 lists the processing specifications. The resists used are RG-3900B for the lower layer, and SNR (Torey Silicon Co., Ltd.) for the upper one. With this process, the upper and lower layers correspond, in terms of capability, to the upper and middle layers for the 3-layer process. With no need to grow a middle layer, the BPL system can do away with the steps for the middle-layer coating and patterning, which are needed in the 3-layer process.

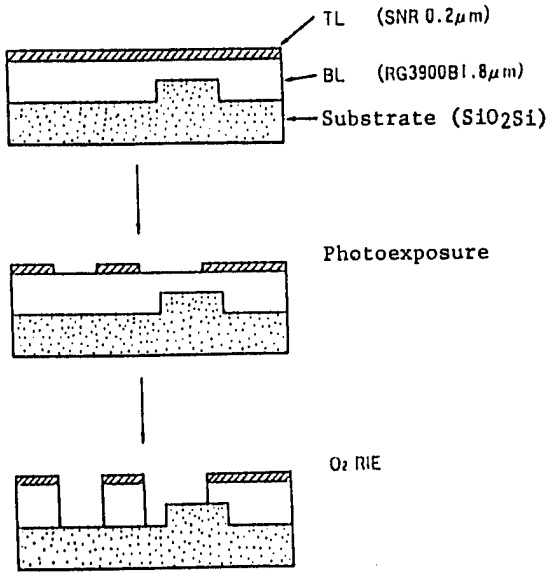


Figure 9. Steps in the 2-layer resist process

Of course, compared to single-layer resist processes, there are still a lot of steps. It does, however, produce a negative pattern with a high aspect ratio. Figure 10 shows an SEM photograph of the completed 0.5-micron L&S pattern. As shown in Figure 11, although the shift for the UV spectra at 248.4 nm before and after the exposure on the SNR layer is quite large, the SNR's photosensitivity is low. In other words, if it is too good, the transmittance in the resist can cause problems, and a certain level of absorption will be a great help. As for the gamma characteristics, shown in Figure 12, they are much better for SNR than for MP2400 because of SNR's better transmittance for the exposure spectra wavelength. Lastly, little thinning of the film was observed during the development of the emulsion.

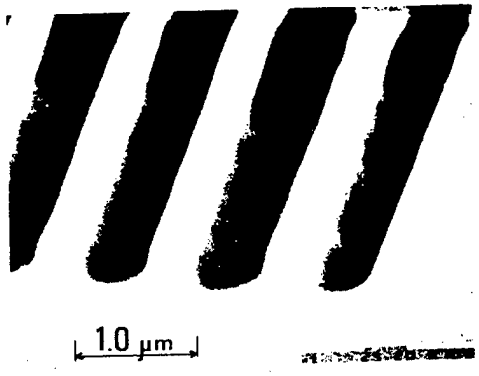


Figure 10. SEM picture of a resist pattern prepared by the 2-layer resist process

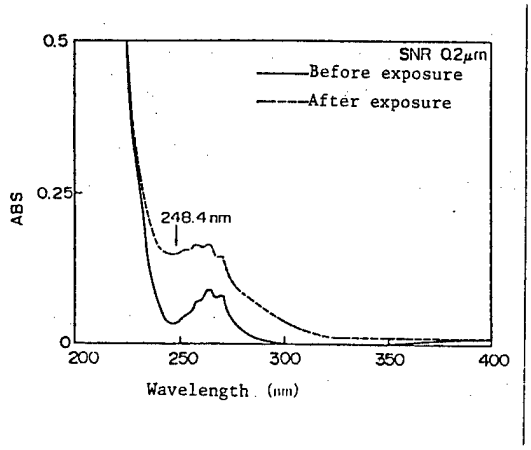


Figure 11. Changes in the absorption spectra of SNR before and after the exposure utilizing a KrF excimer laser

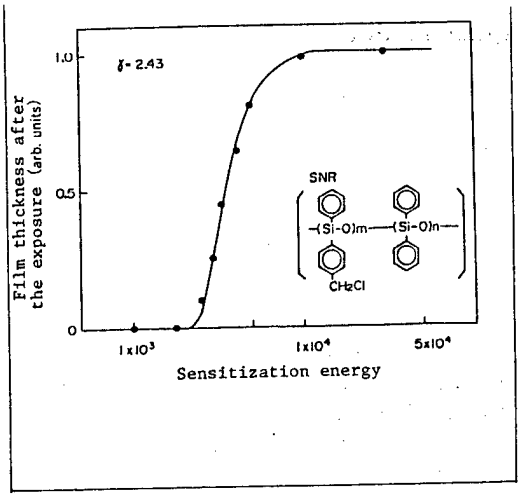


Figure 12. (Gamma) characteristics of SNR, photo-exposed by a KrF excimer laser

4. Summary

In the above sections, we discussed the current status and technical issues surrounding the excimer laser lithography which is now favored as a replacement for the present lithographic systems.

As far as the photolithographic process is concerned, there is no difference between the excimer laser stepper and conventional steppers that use the i- or g-line. Also, compared to the direct-drawing EB photo-exposure, and X-ray exposure methods (i.e., SOR), the excimer laser system presents fewer technological obstacles. With the use of a KrF excimer laser stepper, the throughput can be increased to as high as the currently maintained level

using a conventional stepper. Also, it gives a sufficiently good resolution (0.4 microns). Furthermore, the technology of the KrF excimer laser stepper basically is part of the same developmental extension as the technology of optics. With suitable modifications to the conventional resist processing technology, it will be immediately compatible with the accumulated resources we now have, and so, its introduction to production plants brings about many practical benefits. The technology is sufficiently mature, and will very soon be ready to start churning out the 16-megabit DRAM as soon as the peripheral technology for the chip is ready.

The remaining problem is how to improve alignment precision. Production of next generation super LSI's is expected to require a solid working resolution of 0.5 microns at least, and an alignment precision of plus or minus 0.1 micron. Therefore, it is essential to step up the efforts for improvements in alignment precision in order to materialize the introduction of the excimer laser lithography system into production lines.

Thin Film Production With Laser CVD

43064082d Tokyo OPTRONICS in Japanese Oct 88 pp 104-110

[Article by Susumu Hoshinouchi, Noriko Morita, Hiroshi Onishi; Mitshbishi Electric Corp.]

[Text] 1. Introduction

Prompted by the growing industrial need for greater semiconductor circuit integration, urgent technological issues are now surfacing regarding ways to run thin-film growth processes at lower temperatures and at the same time to reduce the number of defects. Even outside the field of semiconductor devices, e.g., optical components, where improvements in product performance and capabilites, and compact packaging have been made, there is a strong need for thin-film growth processes that make it possible to precisely control the growth of particular films at particular spots and layers according to the blue print through the controlling of physical phenomenon at atomic and molecular levels.

One of the more promising technologies being developed to meet these needs is the laser CVD method. Under this method, laser light is shined on the target--the raw material gas, and the substrate--to excite molecules. The excited molecules undergo chemical reactions, forming a film on the substrate.

In 1978, Christensen et al succeeded in growing silicon films by way of thermal decomposition in SiH_4 achieved by irradiating the substrate with a condensed CO_2 laser beam. Then, in 1979, Hanabusa et al succeeded in the same

task, but did it using the photodissociation of SiH_4 induced by a multiple photon absorption of a CO_2 laser.

Since then, intensive research efforts have continued producing extensive results; as shown in Table 1, numerous reports of thin film growth for semiconductors, metals and insulators have been made during the last 4 to 5 years. It's a young technology having just left its infancy, on its way to full maturation. Although some of the effects are not yet fully understood, the prospect for a breakthrough is there since there have been many reports of fruitful attempts to incorporate laser technology in industrial applications.

In this article, we will present a brief summary of recent research activities that studied the role of light in film growth processes. Later, we will introduce a few applications that are being tried.

Table 1. Laser CVD-Grown Thin-Films (II): 2nd harmonics

Laser	CO_2	Nd : YAG	Ar	Ar (II)	Cu (II)	XeCl	KrF	ArF
Thin films	Si, Ni Cu, Au TiO_2 TiC	Si	Al, Ni Zn, Cd	Si, Al Cd, Sn Zn, W Cr, Fe	W Mo Cr	Si Ge	Si Ge W Cr ZnO Al_2O_3 Al	W, Si Al, Zn Fe, Ge Cr, Cd ZnO InP BN SiO_2 Al_2O_3

2. Equipment Organization

Basically, there are two laser CVD systems depending on the type of laser illumination chosen. In one of them, the laser beam is directed horizontally in parallel to the broad side of the substrate. Also, the use of light is limited to excitation and dissociation of the raw material gas. In the other systems, the laser beam directly illuminates the substrate in order to cause desired reactions at the substrate surface.

The choice of illumination method depends on what is to be achieved through the processing. With the horizontal illumination, although there is the advantage of being able to spatially specify the excitation of the raw material gas, the actual film growth reactions that take place on the substrate are limited to thermal energy-induced reactions. With the substrate illumination method, a film deposition process, a photoexcitation process occuring at the surface, can be controlled, but since the gas too undergoes a similar excitation process, the mechanism of the total process gets to be very complex. With regard to the latter deposition process, which will be introduced later, there is a related method; this technique utilizes one of

the laser's features--coherency--to enable selective localization of the deposition.

As for the laser system, the excimer laser is mostly used for its large photon energy level. Lasers in visible and infrared spectrum are in use, too.

As you can see, there are numerous system configurations reported for laser CVD equipment. They differ in the type of laser in use and the type of illumination method. In Figure 1, we have illustrated the organization of the excimer laser CVD system that we are using in our W-film research.

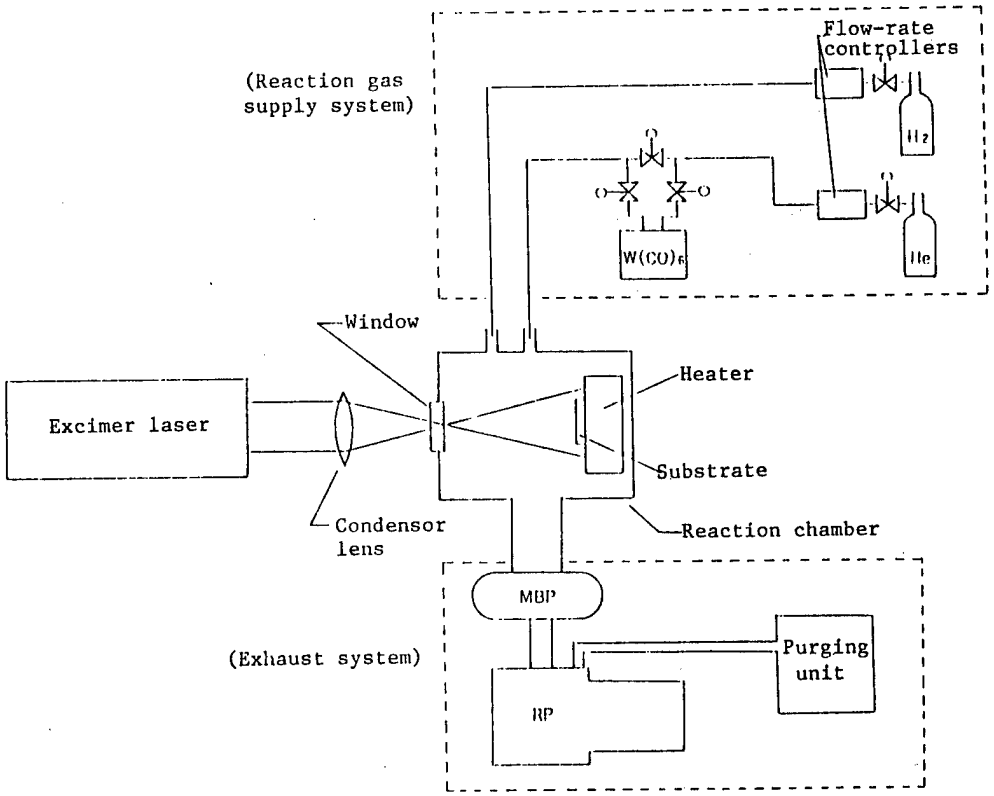


Figure 1. Configuration for the excimer laser CVD system

3. Features of Film-Growth Process

CVD methods may be classified into three categories according to the method of activation chosen for the raw material gas mixture: thermal CVD, plasma CVD, and photo CVD. Laser CVD is included in the photo CVD.

With regard to film growth processing, the following features are common to all CVD methods:

(1) A wide range of film materials, i.e., metals, non-metals, and chemical compounds, can be used.
(2) Numerous types of film, i.e., mono-layer films, multi-layer films, metastable-phase films, polycrystal films, and single crystal films, can be grown easily through selection of the raw material gas, and control of reaction conditions.
(3) Films grown have excellent adherence strength because of the use of vapor reactions.
(4) Film growth is fast.

In addition to these, laser CVD offers the following:

(1) By choosing the wavelength of the laser line, a specific decomposition recombinant can be selectively deposited, making it possible to control the quality of the film.
(2) Through adjustment to laser beam concentration/diffusion, reaction space can be specified, thereby enabling control of film quality and localization of film growth.
(3) Since a photodissociation process is used, film growth can be performed at low temperatures.

These advantages have many practical implications.

4. Role of Light in CVD

Depending on the laser illumination method elected, the role of light in the film growth process differs. In this article, we limit our discussion to cases of direct substrate illumination where the temperature rise in the laser-an irradiated substrate may be ignored. We will thus discuss the effects of light illumination on the processes of vapor reactions and surface reactions.

4.1 Vapor Reaction Process

There are three stages involved in vapor reaction: photon absorption by the raw material gas; decomposition of the gas; and transport of chemical species created during the decomposition process.

During the decomposition process of the gas which has absorbed photon energy, two types of dissiciation procesess take place: the direct dissociation whereby molecules may get excited and reach their dissociation potentials; and the predissociation whereby molecules are initially brought to a stable bonding potential, and are later shifted to the dissociation potential. Certainly, if the excitation energy is not sufficiently high, the molecules cannot move to the dissociation potentials, preventing the photo-dissociation. What this means is that once we know the wavelength for the excitation, the absorption characteristics of the gas, and its bonding

dissociation energies, we can roughly predict the probability of having a desired photo-dissociation. Refer to Tables 2, 3 and 4 for listing of, respectively, the excimer laser's lasing wavelengths with corresponding photon energy levels; the absorption characteristics of often-used gases; and the levels of bonding dissociation energy.

Table 2. Lasing Lines and the Corresponding Photon Energies for an Excimer Laser

	波　長	光子エネルギー	
	(nm)	(eV)	(kJ/mol)
ArF	193	6.4	618
KrF	249	5.0	479
XeCl	308	4.0	387
XeF	350	3.5	341

Table 3. Photon Absorption Characteristics of Often-Used CVD Raw Material Gases

化　学　種	吸収端 (nm)	最大吸収波長(nm)
SiH_4	160	120
Si_2H_6	210	
O_2	242	
O_3	300	250
CO_2	228	
NO_2	398	
N_2O	240	180
NH_3	210	190
PH_3	220	<200
P_2H_4	260	<220
B_2H_6	200	180
H_2S	270	200
HCl	280	
Cl_2	400	340
Br_2	510	400
$CFCl_3$	265	
CF_2Cl_2	240	
CCl_4	235	
CH_3Br	250	200
CH_3Cl	200	<180
BrCCH	280	210
$Al(CH_3)_3$	260	
$W(CO)_6$	300	
$Mo(CO)_6$	300	

Table 4. Bonding/Dissociation Energy of Chemical Species Often Used in CVD

結　　合	結合解離エネルギー D (kJ/mol)	
H_2As-H	310.6	
H_3B-BH_3	146	±17
Cl_3C-Cl	297	
CH_3-Cl	345	± 5
CH_3-H	434	± 6
Cl-Cl	242.58	± 0.01
Cl-H	431.6	± 0.2
H_3Ge-H	364	±21
H_2N-H	432	±13
NN-O	167	± 1
H_2P-H	284	±59
H_3Si-H	377.8	
SiH_3-SiH_3	310	

化　学　種	平均結合解離エネルギー D (kJ/mol)	
$Al(CH_3)_3$	279	± 5
$As(CH_3)_3$	239	± 6
$Ga(CH_3)_3$	247	± 4
$Mo(CO)_6$	152	± 1
$W(CO)_6$	178	± 1
$Y(C_5H_3)_3$	374	± 4

Chemical species produced by the photo-dissociation undergo diffusion and convection, and eventually reach the substrate surface where they become a part of the growing film. Since the concentration of chemical species in the reaction space near the substrate surface is determined by the competition between the secondary reaction speed and the transport speed, in order to understand the process of the vapor reaction, it is essential to know the type and speed of reaction that's taking place amongst the species that are responsible for the vapor reaction, as well as the diffusion speed. Furthermore, the spatial energy distribution of the illuminating light--variations are caused by the vapor reaction-produced convections and irregular concentrations--which need to be carefully watched.

As you can see from the above passages, when compared to plasma CVD, the laser CVD is much simpler in terms of the complexity involved in the vapor reaction process. Nevertheless, there is no changing the fact that the process of laser CVD is still very complicated. In any case, since the vapor reaction process greatly affects the film composition and deposition speed, understanding its details has practical implications for our ability to control the reaction and to ultimately design practical laser CVD systems.

With the present technology, it is very difficult to directly monitor the reaction-formed species generated during the photo-dissociation, and their quantities. We conducted reaction simulations for the SiH_4 reaction that takes place under mercury-sensitizing photo CVD, a technique with a

relatively simple reaction environment. The simulation results are shown in Figure 2. Our success with the simulation indicates that it is very possible to obtain detailed data regarding the variation in the concentration of and time variation in chemical species. Actually, It will not take too much effort to extend our method to laser CVD's. In fact, we believe that such study will prove to be very productive in selecting the suitable parameters for controlling the vapor reaction process and in designing the best equipment organization for an optimum laser CVD system.

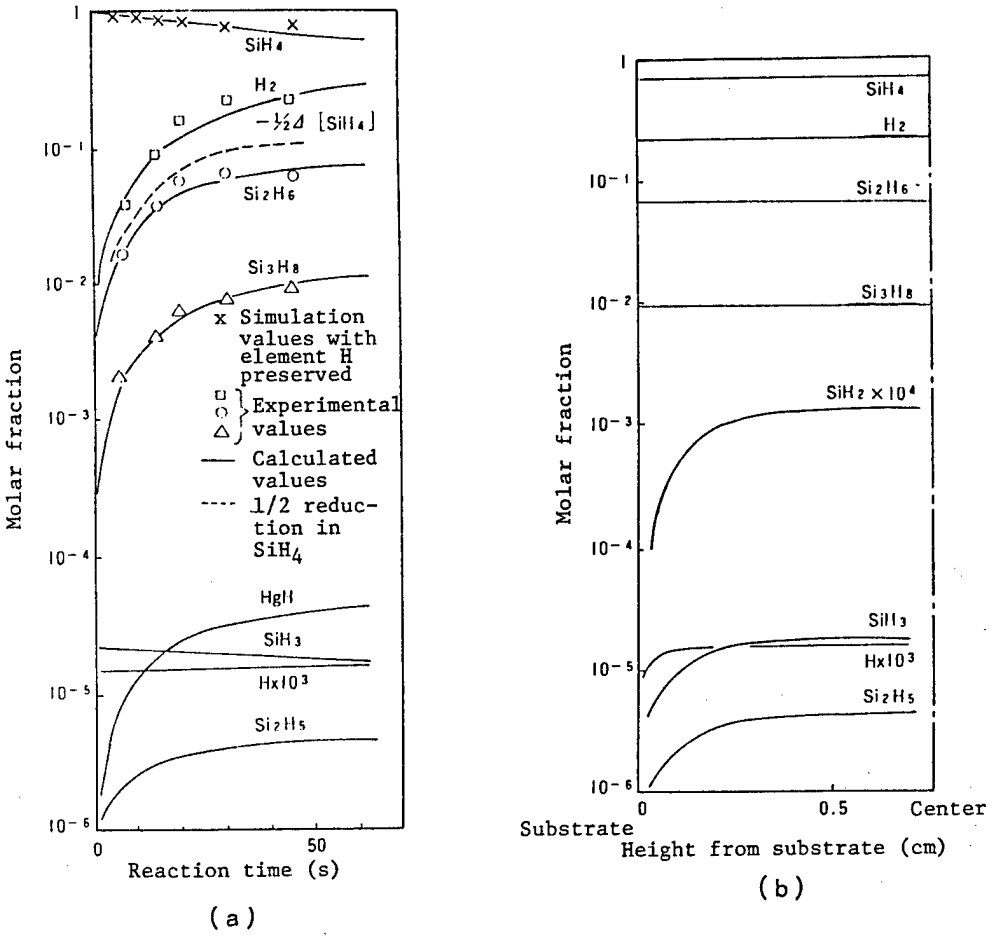


Figure 2. Plot of the data from reaction simulations for the Hg-sensitized CVD with SiH_4; (a) Time variations of the chemical species concentrations, (b) Variations of the chemical species concentrations

4.2 Surface Reaction Process

Chemical species created through vapor reactions fall onto and adhere to the substrate surface becoming a part of the deposition film. This phenomenon may follow the Rideal or Langmuir-Hinshelwood mechanism.

In general, with the partial pressure of a chemical specie near the substrate denoted by P_i, the deposition volume per unit time, unit area V_{Wi} is expressed by the following formula called the Hertz-Knudsen formula:

$$V_{Wi} = (beta)_i * P_i \; / \; \sqrt{2(pi)M_i RT}$$

where $(beta)_i$, and M_i are chemical specie i's adherence probability and molecular mass, respectively; R the gas constant; and T the gas temperature.

$(beta)_i$ varies depending on the surface state, temperature and pressure as well as the chemical specie's internal energy. These dependencies however are hotly debated among many researchers. When the substrate is illuminated with light, some of the already adhered chemical species may become excited as well. This fact can forceably alter the determination of $(beta)_i$, and therefore, should be carefully noted.

In order to obtain a high-performance high-quality thin film, it is essential that the growth process is equipped with the capability to: (1) clean the surface, (2) form a high density nuclei, and (3) grow a film. The effect that these factors can cause upon the illumination-induced reaction deserve detailed research as they need to be precisely controlled to bring out the superior capabilities that laser CVD, and no other conventional film-growth method, possesses. Already, there have been reports on the use of UV light to clean the surface; improvements in the adhered specie's surface migration; and formations of selector nuclei. Further development in this area is eagerly awaited.

5. Proposed Applications

In this section, we discuss some of the proposed applications in crystal growth, growth of functional thin films, and localized selective deposition.

5.1 Applications in Crystal Growth

As already mentioned in the preceeding section, the extent of the effect of light illumination on the reaction process at the surface appears to influence the laser CVD's capabilities. Therefore, applications in epitaxial growth processes, where reactions at the surface play an important role, deserve special attention.

Aoyagi et al discovered a phenomenon where the growth speed is significantly improved for growth of GaAs film; in their setup, an Ar ion laser lasing in the visible spectrum irradiates the reaction gas, which does not absorb visible light, in order to form the film MOVPE. They then applied what they learned into the design of an atomic layer epitaxy (ALE) system, and successfully performed atomic layer growth of GaAs films. The process consists of alternately supplying TEG and AsH_3 which are premixed with H_2--carrier--gas at 1 second intervals, while completing one growth cycle in 4 seconds; and delivering laser beams at a pace of 1 second per cycle--synchronized to the supply rate for the group-III gases.

The results for the growth speed as a function of laser irradiation intensity are plotted in Figure 3. For the cases of UV- or visible-spectra laser irradiation, atomic laser deposition occurs at a region where the laser intensity is high. This effect is thought of as one whereby, under an irradiation of UV or visible laser, the growth of GaAs is accelerated by reacting gas [species] adhered to the surface. The results are very interesting because they successfully brought to light a significant relationship between growth and the type of irradiating beam.

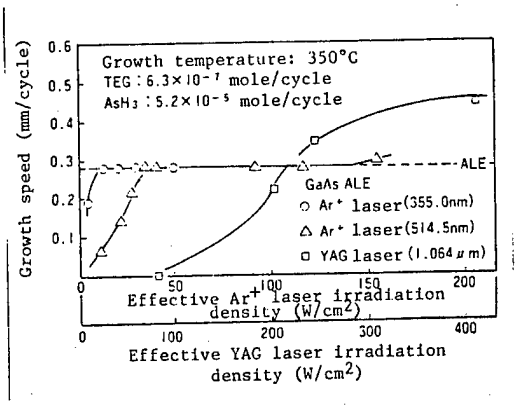


Figure 3. Dependence of the growth speed on the laser irradiation density in the laser MO-ALE method

This technique may be applied not only to chemical compound semiconductors, but also to new types of advanced materials, e.g., high-temperature superconductor thin-films. Development of the technology is sure to bring great benefits.

5.2 Applications in Growth of Functional Thin-Films

The field in question presents many possibilites for the application of laser CVD's. We have to however limit our discussion to diamond thin films, a recently invented wonder material, and W thin film, a high-fusion metal wiring material that's geared for future needs.

(1) Diamond Thin-Film

A diamond possesses not only a super hardness but also outstanding thermal conductivity and insulation strength, and these properties make it stand out among the crowded array of materials. Because of these remarkable properties, ever since the discovery that even under a depressurized condition a diamond in non-equillibrium phase can be synthesized, active research has been underway to synthesis such a diamond. In particular, strong efforts have been directed to developing low temperature processes whereby obtaining thin films with good crystallinity is easier.

Yagi et al succeeded in growing diamond films with a relatively good crystallinity using a low substrate temperature ranging from 300 to 600°C. They used an ArF excimer laser to cause a reaction in the gas mixture of CCl_4 and H_2. Table 5 lists the interplanar distances measured with RHEED. According to the researchers, the 6 diffraction patterns matched the ASTM values with an uncertainty of 5 percent.

Table 5. Interplanar Distance in Thin Film Samples Prepared by the Excimer Laser CVD

Diamond (ASTM)		Diamond thin-film
hkl	d (Å)	d (Å)
111	2.06	2.13
220	1.26	1.24
311	1.08	1.04
400	0.89	—
331	0.81	0.80
422	0.73	0.70
440	0.63	0.62

With the added capability in the near future for better control of crystallinity and with more substrate materials available, we can apply [the diamond thin-film technology] to tribology materials, audiophile components, and advanced semiconductors.

(2) W Thin-Film

Common to all CVD processes that use the gases of organic metals as a reaction gas is the issue of how to lower the contamination of the gas by carbon.

We sought to control the intrusion of contaminants C and O for the laser CVD. In our setup, the gas used was $W(CO)_6$ mixed with H_2. The substrate was irradiated with an ArF excimer laser. We found that the C and O in-film concentrations were reduced by a greater extent for the laser CVD than for the thermal CVD. Figure 4 compares the degrees of the C and O concentrations for ArF excimer laser CVD with H_2 gas mixture in use, and for thermal CVD. We observed a reduction in the amounts of both O and C. Also, we obtained a resistivity of 30 micro-ohms/cm, and with this value, it can be said that the grown substance qualifies well as a wiring material.

At this time, it is not clear what changes occured in the reaction process and how these changes brought about the drops in the impurity concentrations. It is however believed that since the dependency of the growth speed on the substrate temperature (see Figure 5) varies significantly from when the ArF excimer laser irradiation is on to when it's turned off, certain changes in the surface reaction process may have caused it.

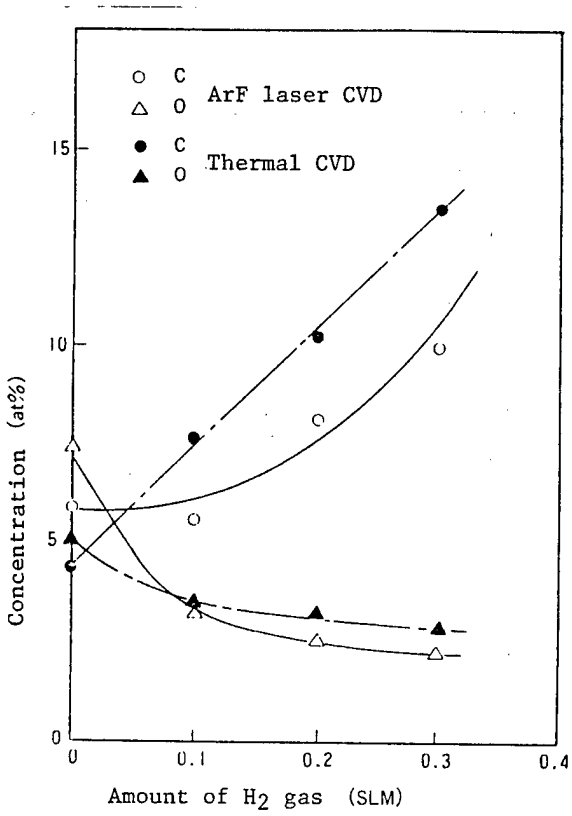


Figure 4. Effect of adding H_2 gas to the C and O concentrations in W films prepared by different CVD methods

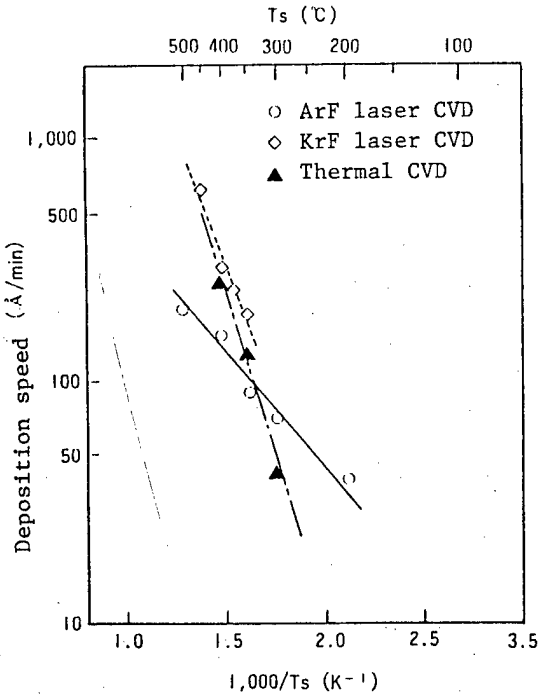


Figure 5. Substrate temperature dependence of the W-film deposition speed

In order to use the $W(CO)_6$-type laser CVD in designing practical devices, an even higher degree of purity will be desirable. At this point, we are thinking of incorporating new methods, e.g., surface excitation in a high vacuum, and through this, we are hoping to achieve a higher level of purity.

5.3 Applications in Localized Selective Deposition

Many researchers are increasingly attracted to the technique of localized selective deposition--direct drawing process--which takes advantage of the laser's excellent spatial controllability and selectivity. This outstanding microprocessing technique makes possible tasks which were considered impossible using conventional techniques, i.e., simplifying device production processes based on the batch transfer method; modifying mask patterns and micro circuits; and fabricating micro optics through simultaneous processing of deposition and shaping with advanced materials.

Ehlich et al reported a discovery of the phenomenon of photo-initiated deposition; with the natural nuclei generation held at a negligible level, the nuclei selectively grown by photo reaction induce successive growth of film.

This process consists of supplying an organic metal such as $Zn(CH_3)_2$ to the reaction chamber; evacuating the chamber using a turbo molecular drag pump to achieve a high vacuum state; and irradiating [the gas] with a condensed beam of second harmonics from an ArF excimer laser or an Ar ion laser. The ultraviolet light in the condensed beam induces a photodissociation in the organic metal adhering to the substrate surface, which is followed by the formation of a nuclei. Following these steps, the reaction gas is introduced into the reaction chamber, and defocused UV light is evenly illuminated on the substrate. This results in localized growth of film where nuclei were selectively formed earlier.

This method is promising not only as a selective growth technique but also as a low temperature growth process.

Finally, we introduce the research of Sugiyama et al. They applied a laser CVD technique to production of a micro lens. This process took advantage of the selective surface reaction initiated by laser-induced localized selective substrate heating. In order to control the selective reaction, it is essential not to induce a vapor reaction. For this reason, a CO_2 laser and an Ar ion laser were employed as the light source. Figure 6 shows the variation in film tickness--due to the temperature distribution on the substrate surface--in the silicon nitride thin film deposited by this method. Near the beam center, the film density is high with the distribution curve more or less following the path of the inscribed semicircle indicated by the dashed line. This sphere-resembling portion may be used as a lens, and its focal distance can be controlled by manipulating the dependencies of the diameter and thickness of the deposited film on the laser beam specifications, etc. So far, the smallest focal distance achieved is 0.6mm.

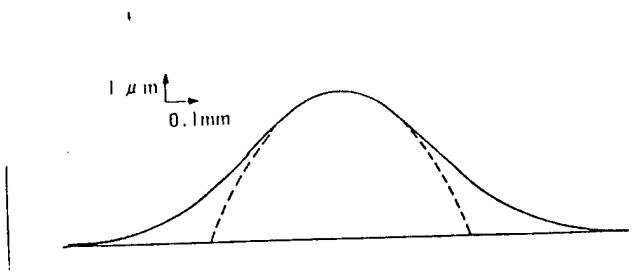


Figure 6. Variation in film thickness of the silicon nitride thin film grown by the localized selective laser CVD method (the dotted line is the inscribed circle)

As the need for micro mechanism components, as well as optical components, is expected to grow, this technique may grow into a promising microprocessing tool.

6. Conclusion

We have discussed the reaction processes of laser CVD, and recent attempts to apply the technology to industrial needs. As mentioned in a preceeding section, in regard to the role of light in film growth processes, many new effects, which cannot be achieved using other methods, have been reported. The technology of light as a means to design advanced materials is indeed very promising.

Laser CVD is a young technology; research took off only 10 years ago. Therefore, there are still many questions regarding the details of the film growth mechanism. In order to raise the technical status of the laser CVD among the other film growth methods, unraveling the photo-induced mechanism behind the film growth is indispensible. Also, researchers should not stop at the mere proposition of possibilities. They must strive for an understanding of the basic phenomena and systematize the acquired knowledge.

As for the equipment side, advances have to be made in the industrial technology for producing laser systems, especially, for those systems operating in the visible to ultraviolet region. Since anticipated industrial tasks are for high value-added devices, the development of the technology is doubly critical.

Nuclear Safety, Radiation Research Examined

LWR Fuel Safety Research

43062019a Tokyo PROMETEUSU in Japanese Dec 88 pp 30-38

[Article by Rikuo Ichikawa, manager of Fuel Safety Engineering Department, Tokai Research Institute of Japan Atomic Energy Research Institute: "Light-Water Fuel Safety Research"]

[Excerpts] [passage omitted] 2. History of Light-Water Reactor Fuel Safety Research in Japan Atomic Energy Research Institute

Uranium dioxide fuel is used for a light-water reactor. Japan Atomic Energy Research Institute (JAERI) has conducted the irradiation tests of this fuel in JAERI, starting with the latter half of 1960 in the JRRA-2 research reactor and started the full-fledged irradiation tests accompanied by the development of high-power fuel required by the power doubling project for the JPDR or Japan Power Demonstration Reactor, which is the first light-water power reactor in Japan, in an attempt to produce domestic fuel in Japan. In 1965, JAERI carried out the safety test of the fuel assembly installed on the JPDR for testing purposes. However, the power of this reactor was too low to be adopted for the power doubling project. JAERI joined the OECD or Organization for Economic Cooperation Development's Halden project and conducted the irradiation tests of the domestic fuel assemblies, using a Halden reactor in Norway. In these tests, JAERI made effective use of the fuel instrumentation, such as fuel rod central temperature, fuel rod elongation, fuel rod internal pressure and fuel output developed by the Halden project.

The measurement of fuel elongation in particular out of the subjects stated above triggered fuel deformation research and formed the major research theme about the high-power fuel irradiation tests continuously carried out by JAERI in the Halden reactor.

Photo 1 [not reproduced] shows the fuel instrumentation assembly.

In the latter half of 1970, worldwide attention was focused on pellet/coating interaction. More specifically, much international attention is being given to the problems of the tensile stress of a coated fuel

pipe generated by fuel deformation during a sudden power increase of the fuel rod and the stress corrosion cracking of a coated pipe produced by repeated actions of iodine compound (part of fission products) atmosphere, which are considered responsible for the generation of fuel troubles.

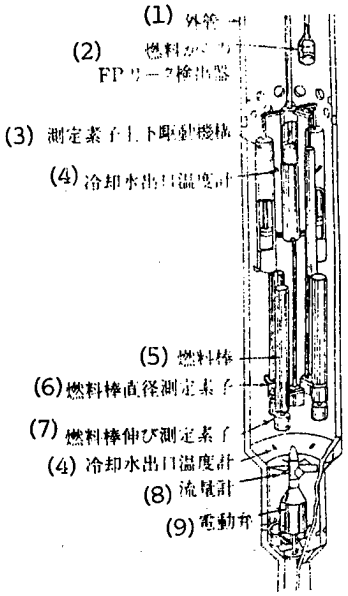


Figure 1. Concept Diagram of Irradiation Testing Unit

Key:
1. Sleeve
2. FP peak detector from fuel
3. Measurement component vertical drive mechanism
4. Cooling water outlet thermometer
5. Fuel rod
6. Fuel rod diameter measurement component
7. Fuel rod elongation measurement component
8. Flow meter
9. Motor-driven valve

JAERI has conducted a series of irradiation tests in the Halden reactor, engaged in research on pellet/coating interreaction as its major theme about fuel safety based on the research foundation created by fuel deformation research, conforming to the safety research annual project established by the state。 Figure 1 shows an example of an irradiation testing unit. In addition, JAERI participated in the power up test project conducted by Sweden's Stassvik Research Institute as its international program, and tried to expand the data base associated with the aforesaid research. Furthermore, JAERI has conducted a power up test in the Halden reactor, using a boiling water type capsule so as to study the characteristics of domestic fuel pellet/coated pipe interactions, thereby having supplemented the irradiation tests in the Halden reactor.

To understand the interaction of the pellet and coating, it is very important to grasp the entire behavior of fuel irradiation. Figure 2 shows the outlined relation between outside factors and events involved in fuel behavior. Since the feedback performance illustrated in the figure exists among these events, it is necessary to develop a calculation code in order to grasp fuel behavior comprehensively. As a result, JAERI has developed a code called FEMAXI. Concerning a series of irradiation research about the Halden reactor, the institute received a special prize from the Atomic Energy Society of Japan (JAES) in 1987 and a technological prize from the society for the development of the FEMAX 1-3 code in 1982.

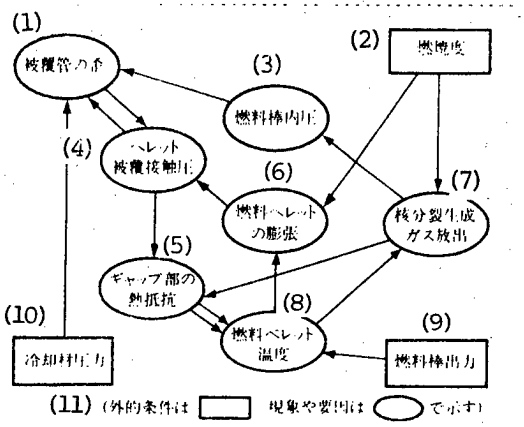


Figure 2. Main Outside Conditions, Phenomena and Factors Associated with the Behavior of LWR Fuel

Key:
1. Strain of coated pipe
2. Burnup
3. Internal pressure of fuel rod
4. Pellet coating contact pressure
5. Thermal resistance at the Gap section
6. Expansion of fuel pellet
7. Discharge of fission generation gas
8. Temperature of fuel pellet
9. Output of fuel rod
10. Pressure of coolant
11. (The outside conditions stand for ▭ phenomena while the factors are expressed in ⬭ .)

3. Research about Fuel Irradiation Behavior and Development of FEMAX Code

The temperature of fuel plays a key role in the irradiation behavior of fuel. Including the irradiation tests in the Halden reactor in the Halden project, JAERI carries out direct measurement of central temperature of the fuel pellets under the irradiation tests, using several hundreds of fuel rods. Figure 3 shows the measurement results of the central measurement of fuels as a typical example. The thermal output of the fuel rods used in the tests were at a low level similar to that of the power reactor fuel. As a result, only a small amount of fission gas

was discharged while the temperature up to the highest level of combustion was relatively low. However, when the fission-generation gas is discharged in the fuel rod, the thermal resistance at the gap section becomes greater, thereby increasing the temperature of the pellet. Large numbers of experimental results about the aforesaid subject were used for the development and demonstration of the FEMAXI code, allowing the prediction of fuel temperature with greater accuracy. The most important factor in an attempt to evaluate the safety of fuel rods is the subject of stress and strain of a coated pipe. The basic data about these items are obtained from the measurement of incore deformation. Figure 4 shows an example of output dependency in terms of the elongation of fuel rod measured in the Halden reactor. The thermal expansion of uranium dioxide pellets is greater than that of the zircaloy of a coated pipe, which increases the output. Therefore, when the pellet and the coated pipe are brought into contact with each other, the pellet affects the coated pipe, producing the stress and strain. Since the elongation of a fuel rod is generated by the dynamical interactions stated above, the elongation of a fuel rod can be used as a standard to denote the magnitude of interaction. As illustrated in the figure, the pellet whose length is short indicates a smaller interaction. The results of the research have served to improve the design.

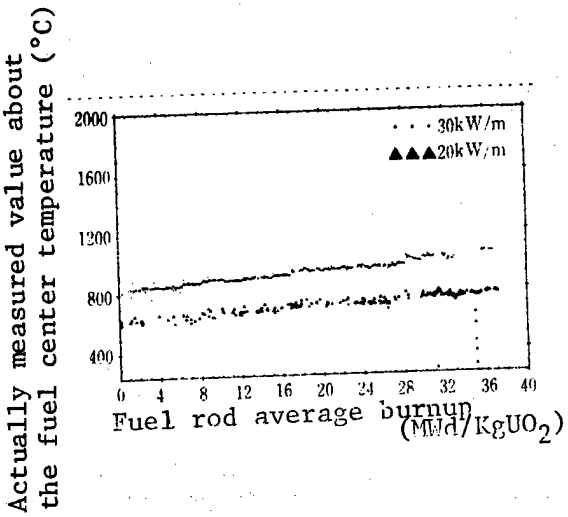


Figure 3. Fuel Center Temperature Measured by High Temperature Thermoelectric couple up to burnup of about 40,000MWd/t
A rise in temperature accompanied by combustion is low. The gap of this fuel rod produced during manufacture is 75 microns and smaller than that of ordinary ones。

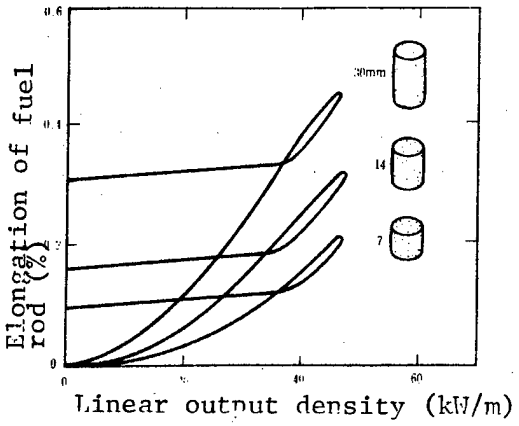


Figure 4. Example of Effects of Pellet Dimensions Affecting the Elongation of Fuel Rod

The elongation of a fuel rod is very convenient to describe the tendency of dynamic interaction with accuracy. However, it only denotes the sum of elongation over the entire length. Therefore, it is more effective to measure the diameter of a fuel rod in order to evaluate local stress and strain. JAERI has given specific attention to the measurement of the diameter of a fuel rod in the core in the Halden reactor. Figure 5 shows an example of measuring the diameter of an incore fuel rod. When dynamic interaction is produced, the fuel rod indicated a bamboo-joint-shaped deformation. Specific attention must be given to the fact that an axial deformation is actually generated in terms of micron dimension and presented as a markedly expanded form. It is virtually impossible to recognize such a level of deformation with the naked eye. The measurement data was very useful to demonstrate and evaluate the sections required to analyze the dynamic behavior which forms the basis for the FEMAXI code.

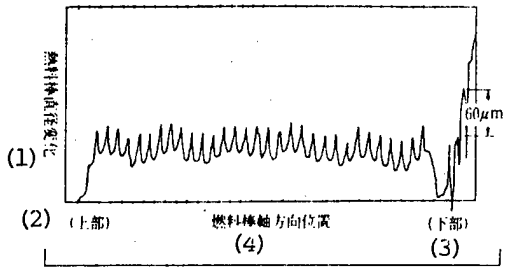


Figure 5. Measurement Example of Incore Fuel Rod Diameter in the Halden Reactor

Key:
1. Variations in the diameter of fuel rod
2. Upper
3. Axial direction position of fuel rod
4. Lower

The FEMAXI code adopts a finite-element method for the sections required to analyze dynamic behavior and it is characterized in that local dynamic interaction can be described properly. Figure 6 shows the conceptional diagram which models the new version (FEMAXI-4) of the FEMAXI. The old version is intended for handling the steady state while the new version is characteristic in that it can handle transient time, such as sudden power up time.

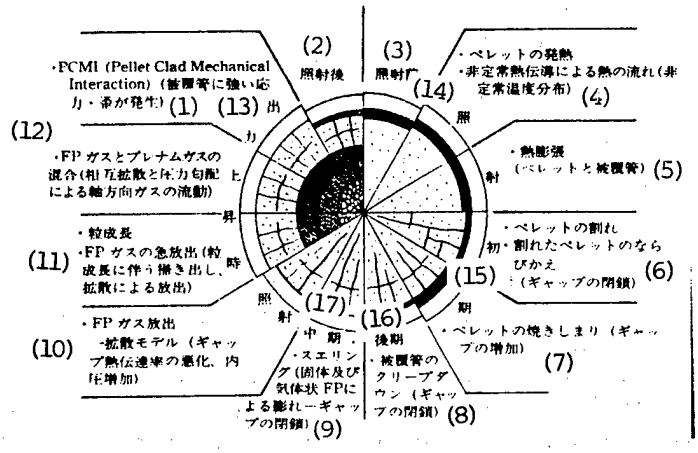


Figure 6。 Various Phenomena Considered in FEMAX-IV

Key:
1. (Strong stress and strain generated on a coated pipe)
2。 After irradiation
3. Before irradiation
4. Heat generation of pellet Heat flow produced by non-steady heat conduction (non-steady temperature distribution)
5. Thermal expansion (pellet and coated pipe)
6. Crack of pellet Rearrangement of cracked pellet (closeout of gap)
7. Densification of pellet (increased gap)
8. Creep down of coated pipe (closeout of gap)
9. Swelling (expansion produced by solid or gaseous FP-disclosure of gap)
10. FP gas discharge-diffusion model (deterioration of gap heat transfer coefficient, and increased internal pressure)
11. Particle growth Sudden discharge of FP gas (discharge accompanied by particle growth and diffusion discharge
12. Mixture of FP gas and plenum gas (Flowing axial direction gas due to inter diffusion and pressure gradient)
13. During power up
14. Irradiation
15. Early period
16. Latter period
17. Irradiation middle period

4. Research of Fuel Failure

The leak-generation rate of light-water reactor fuel is extremely low, say, one fuel rod per million fuel rods when a power reactor is driven for 1 year, producing no serious hazardous problems. This figure is said to be lower by one column to several columns compared with foreign countries. However, several fuel failures which are known to other countries have occurred in the past in Japan. Many attempts have been made to clarify and eliminate the causes of the troubles.

In the early stage of the practical application of light-water reactors, the coated pipe was subjected to the generation of local hydrogen due to the moisture quite frequently contained in the fuel rod. The Halden project clarified the limit of moisture which creates such failures. JAERI has used this limit of moisture effectively in an attempt to understand the basic mechanism by measuring the internal pressure of a fuel rod where moisture is intentionally added to the pellets, by irradiating the rod in the Halden reactor and by finding the absorption of the moisture by the coated pipe in a few days. JAERI has resolved such fuel failures by minimizing the volume or humidity in a fuel rod.

The fuel failures due to the densification of pellet (irradiation sintering) are well known internationally. The failures occurred around 1972. The Halden project tried the irradiation tests in a timely manner by installing device to measure pellet differential length on a fuel rod, thus clarifying the effects of the measurement, such as the density of pellet or sintering conditions. JAERI also tried to study the characteristics of densification of domestic fuel in a joint study with domestic fuel makers and clarified the manufacturing requirements for a pellet whose densification is small by carrying out instrumentation fuel irradiation tests in the Halden reactor.

With regard to the failures or troubles due to the interaction between pellet and coating, JAERI has carried out research for the past decade, using about 40 test fuel rods and 6 sets of irradiation units exclusively designed in the Halden reactor. Three irradiation units out of the six units were especially used to provide the burnup required for the test fuel. The rest of the irradiation units were used to measure the diameter in the reactor during sudden power up and power cycle, the whole length of the fuel rod, the internal pressure of the fuel rod, the central temperature of fuel and to carry out non-destructive tests about the coated pipe. Figure 7 shows the comparison between the deformation (elongation) of a fuel rod which leaked fission products due to the interaction between pellet and coating and a fuel rod which maintained safety even when it was subjected to the same power up. The figure clearly indicates that a fuel rod, when it was subjected to extreme deformation, produced a leakage. The data served to establish a detection method to predict failures due to the interaction between pellet and coating. JAERI confirmed that no fuel failure would occur under the service conditions of the present power reactor, judging from the results of the tests or the data bases for other pellet-coating interaction.

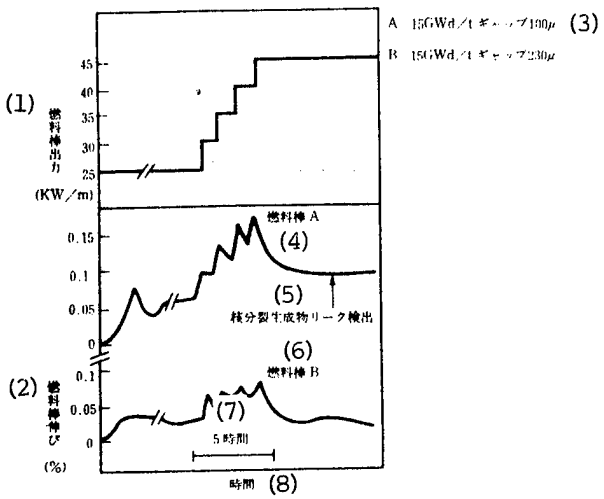


Figure 7. Elongation Comparison Between Fuel Rod (A) which Generated Leakage of Fission Product during Power Up Test and Safety Fuel Rod (B) The burnup of fuel rods is specified as 15,000MWd/t. The manufactured gap of the fuel rod A was 100 microns while the manufactured gap of the fuel rod B was 230 microns.

Key:
1. Output of fuel rod
2. Elongation of fuel rod
3. Gap
4. Fuel rod A
5. Leak detection of fission product
6. Fuel rod B
7. 5 hours
8. Time

5. Conclusion

The accomplishment of JAERI in terms of safety research can be summarized in the following:

a) Confirmation of domestic fuel performance

b) Clarification of incore behavior of fuel rod, investigation into the causes of fuel failures and research of their mechanism

c) Development of the FEMAXI fuel-behavior code

d) Contribution to safety examinations in Japan

With regard to the irradiation of the Halden reactor, JAERI has conducted irradiation tests about 14 fuel rods of JAERI and 24 fuel rods based on a joint research with PNC (Power reactor and Nuclear Fuel Development Corporation) and fuel makers since it joined the Halden project in 1967. They are considered to have contributed greatly toward the improvement in the fuel technology in Japan.

In the future JAERI is going to proceed with safety research of high burnup fuel, including the corrosion problems of coated pipe according to the safety research annual program decided by the state and conduct safety research of a new type reactor under the review of JAERI currently as well, if it materializes.

Pressurized Reactor Vessel Safety

43062019b Tokyo PROMETEUSU in Japanese Dec 88 pp 34-38

[Article by Tatsuo Kondo, manager of Fuel Material Engineering Department of Tokai Research Institute, Japan Atomic Energy Research Institute: "Research of Defect Growth Speed of Nuclear Reactor Pressure Vessel"]

[Text] 1. Introduction

Various discussions or many research tests about the safety of light water reactors are conducted based on scenarios about "virtual accidents." The starting point of these scenarios includes the damage of component materials of the pressure walls which separate the primary system or more specifically, the high-pressure water or vapor system designed to circulate through the incore and make connections with the heat application system from the outside atmospheric pressure system. The equipment in the pressure section covers nuclear reactor pressure vessels, pumps, main circulating piping, and valves. In the case of a pressure water reactor, a pressure generator of pressurizer is also included. When the main piping is subject to damage, it is feared that the damage will lead to considerable leakage and finally bring about an accident due to loss of cooling water.

The pressure vessel is defined as typical pressure-resisting equipment in the list stated above. The pressure-resisting section means that no possible rupture is allowed to occur under any circumstances. Therefore, no rupture is included in the accident scenario. This logic is supported by a series of technological backgrounds, such as the materials, design, manufacture, inspection, and maintenance management, required for the pressure vessel. We have tried to reinforce this technological background for the past decades in the field of light-water reactor technology.

The author wonders what level of damage the power plants currently in service could afford to allow. To put it in other words, we have accumulated through our efforts the know-how to accurately predict the degree of safety or durability by aggressively pursuing knowledge of what conditions must be necessary to bring about rupture. The research on defect growth is one such effort.

2. Vanguard Process to Fracture

As described above, the pressure vessel steel is extremely strong and it is very difficult to break down the vessel intentionally. Cutting in the shop is not an easy task. On the other hand, the steel can be broken down with relative ease if an attempt is made to cut the steel sharply and lower its temperature down to several tens of degrees below zero for cooling. This is brittle fracture which is produced when both "cold embrittlement" and "notch effect" are satisfied. This may be compared to cutting an apple where the hardness of the flesh and cut of a table knife

help each other, if one is allowed to explain in familiar terms. In other words, the brittleness of the flesh stands for the condition of the material and the cut of a knife for the easy propagation of a crack.

Back to the pressure vessel, the brittle fracture of a pressure vessel, which must be extremely strong, is made an object of strict research because it is associated with the cold brittleness as stated above. The embrittlement of a brand new nuclear reactor pressure vessel will never occur unless it is cooled down to several tens of degrees below zero. However, as illustrated conceptionally in Figure 1, the change point (transition temperature) from ductility to brittleness will move to the higher temperature side gradually if an attempt is made to apply fast neutrons in great quantities. If the change point rises to operating temperature, the vessel will lose its inherent toughness. The degree of change greatly depends on the material, as described later. The degree of change also differs with reactor type, the in-core position, and with the depth from the surface of a material. For example, this presents no vital problem primarily to a small boiling waterreactor which produces a small amount of neutron irradiation. Even in the case of a pressure water type reactor, it is a problem around the surface of the shell of the vessel near the core. Its effect is drastically attenuated with the depth inside the shell.

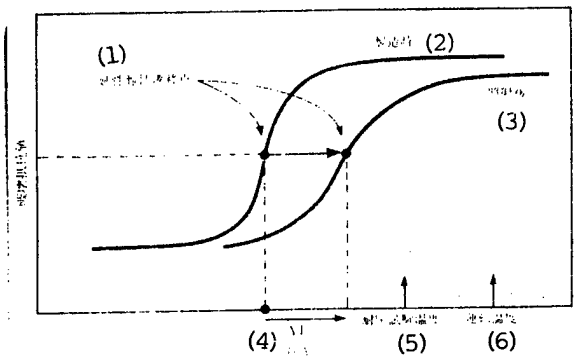


Figure 1. Concept Diagram Denoting the Change in Fracture Characteristics by Fast Neutron Irradiation of Steel for Nuclear Reactor Pressure Vessel

Key:
1. Transition point between ductility and brittleness
2. During manufacture
3. After irradiation
4. Temperature
5. Pressure test temperature
6. Operating temperature

Then, the author will discuss the method of understanding quantitatively the relation between the potentiality of fracture and this irradiation brittleness. As it is not required to recall the metaphor of an apple stated above, it would be better to think of a "critical defective dimensions" safety standard aimed at clarifying the magnitude of defects which can break down the vessel. When this concept is applied to the

the relation between embrittleness due to the irradiation of the pressure vessel and fracture-resisting force, it will be equivalent to the drop in the defective dimensions allowable to the structural members, if the materials are subject to deterioration by neutrons.

According to this concept, the lifetime of the pressure vessel can be obtained at the intersection of the two factors as illustrated in Figure 2。 This clearly indicates that the vanguard process called "crack growth" is required to bring about fracture.

The opening of the two curves at a certain time in the figure is equivalent to the margin of safety from fracture. For their quantification, incore surveillance tests provide practical data on the subject of materials while non-destructive tests during fixed intervals provide concrete data on the subject of defects, respectively.

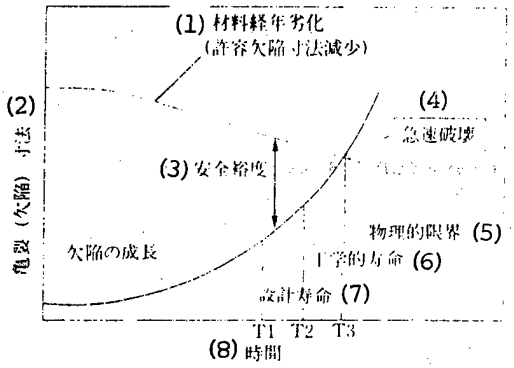


Figure 2. Concept Diagram Denoting the Lifetime of Light-Water Reactor Pressure Vessel

The lifetime of the pressure vessel can be considered based on the deterioration produced by radiation or heat and defective dimensions as its basic criterion

Key:

1. Material deterioration (reduction in allowable fault dimensions)
2. Crack (fault) dimensions
3. Margin of safety
4. Rapid fracture
5. Physical limit
6. Engineering lifetime
7. Design lifetime
8. Time

3. Real Situations of Defective Growth and Its Driving Factors

The author wonders how the reactor could operate when it is cracked. If a significant defect signal is obtained by period inspection, repair work will be called for without doubt. Before starting the operation, the required margin of safety will be evaluated through strict numerical analysis. Therefore, as indicated in the previous figure, there will be no situation which allows the operation to continue for several decades when there are comprehensive cracks. Therefore, this figure puts stress on an assumed premise that the defect is based on maintenance.

As a matter of fact, this is a serious problem which concerns the establishment of safety assurance between two periodic inspections. On the other hand, however, we must recognize the limit of accuracy in terms of fault defection. More specifically, we must recognize the premise that there always exists the seed of a defect less than the critical value as well.

According to the newest information available to us, a long period time ranging from several tens of years to over a hundred years will be required before a defect produced on the surface of a pressure vessel of a reactor proceeding with routine operations grows to one-quarter of the thickness of the shell, which is a roughly defined standard for fracture. The results of our researches for the past two decades allow us to accurately predict material fracture with some confidence. Accomplishment of the above research are thanks to the contribution of an internationally organized joint research group called ICCGR (chaired by the author) organized by the leading scientists of 11 countries and 50 organizations in the world. On the domestic scene, 16 organizations carried out joint tests with JAERI as a central figure under the management of Hideo Kitagawa, honorable professor of Tokyo University.

The main objective of the research is to measure and analyze the defect-growth mechanism, dominance factors and speeds produced on a pressure vessel, in conditions as close as possible to actual reactor environmental conditions, thereby establishing a common evaluation basis.

Historically, the research originated from an experiment attempted by a JAERI research group, including the author as a member, triggered by a crack produced on the internal surface of a pressure vessel in a 1968 JPDR accident. An indentation or crack, if produced on the internal surface of a pressure vessel due to corrosion, will accelerate "fatigue" generated by fluctuation loads accompanied by the start and stop cycle of the reactor. The fatigue-based crack growth itself is the mechanism of defect growth stated above. We have not found any significant clue which connects the microscopic defect except for this.

The reason why the Japanese experimental results attracted worldwide attention and spurred an attempt to revise the defect evaluation procedure of the ASME Design Code (1974 version) and carry out international cooperation over such wide areas as stated above, comes from our discovery that crack growth is many times faster in the cooling water of a nuclear reactor than in the atmosphere. If we simply interpret this fact, it means that if we allow the fatigue to proceed when a large defect is discovered, during the safety period which we could allow fractures may occur. However, actual events are not so simple. It is true that the above subject created a major reactor engineering problem at that time, including the truth or falsehood of the experiment. This phenomenon is generally called corrosion fatigue. High-temperature subaqueous crack growth research, which made a seemingly sensational start, has made dramatic progress for the past two decades, adding several important discoveries. The author will introduce some of those important discoveries.

4. Quality of Steel and Fracture Resistance

In 1968 it was reported that a rise in steel brittleness transition temperature produced by the irradiation of neutrons is governed by copper and phosphorus, which are impurities contained in steel. However, the aforesaid joint research both here and abroad clarified the fact that high-temperature subaqueous crack growth is accelerated by sulfur contained in steel. Solutions of the above problems especially in the past few years are quite dramatic. When the phenomena associated with the quality of two kinds of steels are considered together, it must be admitted that we are forced to drastically change our recognition about the fracture resistance and lifetime of a steel-made reactor pressure vessel. The then technological level of the United States that had developed a light-water reactor had a marked influenceon the design guidance or the accident scenario which serves as the design standard for nuclear reactors and the standard for safety evaluation. They still have considerable significance because old-type U.S. reactors are the standard. The author wonders whether the standards should apply to Japan, which is a newly developed country in terms of reactors. From the standpoint of steel technology, the steel materials used in Japan's nuclear power plants, even in the early stage of development, were considered as the best according to international standards. For the past decade, clean materials which may be defined as exceeding quality standards have occasionally been used. With regard to the speed or high-temperature subaqueous crack growth, we find it necessary to examine the actual state of Japan's power plants currently in service once again. Irrespective of their failures, for better or for worse, we firmly believe that they will provide information which is very important to examine for the future development of nuclear reactors. As a matter of fact, the quality level of the reactor pressure vessels manufactured by the state-of-art-steelmaking technology has been improved to such an extent that almost no brittleness or crack-growth acceleration, which have been discussed so far, will occur.

Figure 3 shows an example of the difference produced by the sulfur contents of high-temperature subaqueous fatigue crack growth behavior. Apparently, sulfur affects the speed of corrosion-fatigue progress and fracture after the test. It is found out that the quality of steel is a key to the solution of the above problems. Even the medium-sulfur materials, (corresponding to JPDR generation) which were used for a relatively old power plant, rank below the specified upper limit of domestic ASME-code-based materials. The low-sulfur materials used for almost every power plant in Japan are substantially free of growth acceleration and fail to reach the atmospheric reference line specified by ASME in most cases, which illustrates the actual situation in Japan.

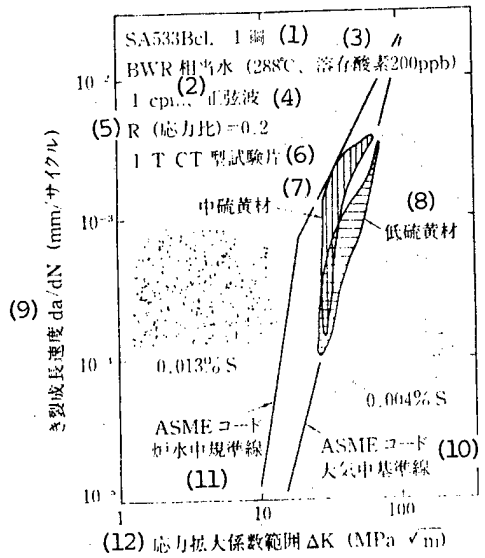


Figure 3. Relation Between Fatigue Crack Growth Speed and Stress Expansion Coefficient Range in the Steel of Pressure Vessel in BWR Primary Cooling System Using Nearby Water
K is increased in proportion to the enlargement of the crack, (Derived from the common material test results based on the cooperation between the seven domestic organizations in JAERI's structural safety research institute/corrosion fatigue expert committee)

Key:
1. Steel
2. BWR equivalent water
3. Dissolved oxygen
4. Sine wave
5. Stress ratio
6. IT-CT type test piece
7. Medium-sulfur material
8. Low-sulfur material
9. Crack-growth speed da/dN (mm/cycle)
10. ASME code atmospheric criterion line
11. ASME code reactor subaqueous criterion line
12. Stress expansion coefficient range

Figure 4 provides a more interesting topic。 The figure shows the study of the upper limit of the impurity analytical value in terms of pressure vessel steel for nuclear reactors (including those for export) produced in Japan. The selection of the axes of ordinate and abscissa is made based on the two great events associated with the breakdown described above. The quality of materials has made dramatic progress with the passage of time. What is attracting most attention is the approach to the area denoted with slanting lines in the lower left side of the figure。 The materials near this area no longer present any serious problems of neutron irradiation-based brittleness and defect-growth acceleration used currently. This at least is the fact we must understand when we consider the engineering safety of nuclear reactors.

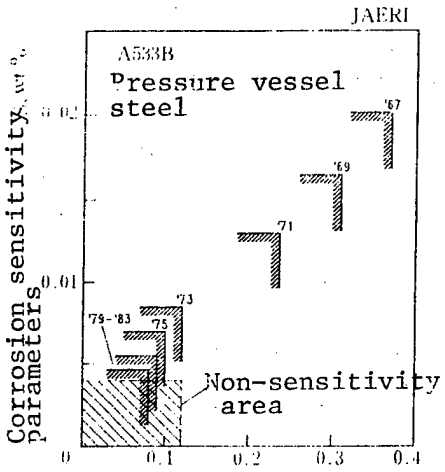


Figure 4. Age-Based Transition About the Relation Between the Sensitivity of Irradiation Brittleness and Corrosion Crack Sensitivity Parameter Calculated Based on the Maximum Contents of Impurity Elements in the Pressure Vessel Steels Manufactured in Japan.

5. Relation With Extension of Plant Lifetime

The lifetime of a nuclear power plant can be determined mostly based on the pressure vessel, which is very difficult to replace. The author believes this assumption is not off the point. Figure 5 shows an example of prediction about the engineering lifetime or the breakdown of a nuclear reactor pressure vessel estimated based on conservative values as much as possible, using the up-to-date information and understanding available. Assuming that we were to continue operation, neglecting a crack as deep as 5 mm, we would never find any sign of breakdown until 50 years had passed according to the old data. To be more precise, it would take more than a hundred years before the reactor becomes subject to breakdown according to the data currently available. It is certain that there is no rule without exception. However, what the author has discussed so far is one conclusion in terms of reactor safety through many critical processes. The lifetime of a power plant involves many important aspects of the subsequent nuclear engineering problems to be solved, such as reduction of radioactive waste and the conservation of resources. It may be inferred that the recognition of the role of research about the damage and breakdown of materials has changed slightly step by step.

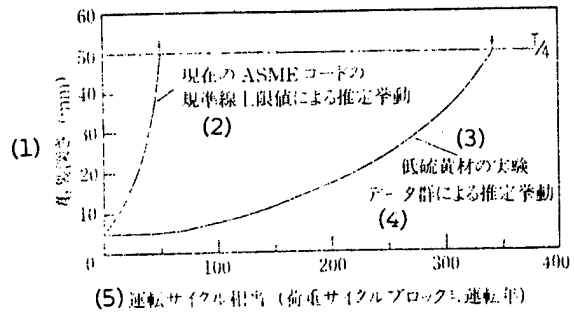


Figure 5. Calculation Example on the Relation Between the Residual Lifetime of LWR Pressure Vessel where a 5-mm Deep Defect Is Neglected at the Nozzle and the Reference Value About Crack-Generation Speed Based on the Experimental Data

Key:
1. Depth of crack
2. Estimated behavior based on the upper limit of the reference line according to the newest version of ASME code
3. Experiment of low-sulfur materials
4. Estimated behavior based on data groups
5. Equivalent to operation cycle (load cycle block = operation years)

6. Conclusion

It is greatly due to our scientific progress in the past decade that we can handle structural and material breakdowns with satisfactory accuracy in a quantitative manner. What the author has discussed in the previous chapter may give an optimistic impression. However, it is a conclusion which has undergone many criticisms, starting with the criticism or doubt of established concepts.

- END -

10

22161
45
NTIS
ATTN: PROCESS 103
5285 PORT ROYAL RD
SPRINGFIELD, VA 22161

Foreign Broadcast Information Service (FBIS) and Joint Publications Research Service (JPRS) publications contain political, economic, military, and sociological news, commentary, and other information, as well as scientific and technical data and reports. All information has been obtained from foreign radio and television broadcasts, news agency transmissions, newspapers, books, and periodicals. Items generally are processed from the first or best available source; it should not be inferred that they have been disseminated only in the medium, in the language, or to the area indicated. Items from foreign language sources are translated; those from English-language sources are transcribed, with personal and place names rendered in accordance with FBIS transliteration style.

Headlines, editorial reports, and material enclosed in brackets [ ] are supplied by FBIS/JPRS. Processing indicators such as [Text] or [Excerpts] in the first line of each item indicate how the information was processed from the original. Unfamiliar names rendered phonetically are enclosed in parentheses. Words or names preceded by a question mark and enclosed in parentheses were not clear from the original source but have been supplied as appropriate to the context. Other unattributed parenthetical notes within the body of an item originate with the source. Times within items are as given by the source. Passages in boldface or italics are as published.

SUBSCRIPTION/PROCUREMENT INFORMATION

The FBIS DAILY REPORT contains current news and information and is published Monday through Friday in eight volumes: China, East Europe, Soviet Union, East Asia, Near East & South Asia, Sub-Saharan Africa, Latin America, and West Europe. Supplements to the DAILY REPORTs may also be available periodically and will be distributed to regular DAILY REPORT subscribers. JPRS publications, which include approximately 50 regional, worldwide, and topical reports, generally contain less time-sensitive information and are published periodically.

Current DAILY REPORTs and JPRS publications are listed in *Government Reports Announcements* issued semimonthly by the National Technical Information Service (NTIS), 5285 Port Royal Road, Springfield, Virginia 22161 and the *Monthly Catalog of U.S. Government Publications* issued by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402.

The public may subscribe to either hardcover or microfiche versions of the DAILY REPORTs and JPRS publications through NTIS at the above address or by calling (703) 487-4630. Subscription rates will be provided by NTIS upon request. Subscriptions are available outside the United States from NTIS or appointed foreign dealers. New subscribers should expect a 30-day delay in receipt of the first issue.

U.S. Government offices may obtain subscriptions to the DAILY REPORTs or JPRS publications (hardcover or microfiche) at no charge through their sponsoring organizations. For additional information or assistance, call FBIS, (202) 338-6735,or write to P.O. Box 2604, Washington, D.C. 20013. Department of Defense consumers are required to submit requests through appropriate command validation channels to DIA, RTS-2C, Washington, D.C. 20301. (Telephone: (202) 373-3771, Autovon: 243-3771.)

Back issues or single copies of the DAILY REPORTs and JPRS publications are not available. Both the DAILY REPORTs and the JPRS publications are on file for public reference at the Library of Congress and at many Federal Depository Libraries. Reference copies may also be seen at many public and university libraries throughout the United States.