APPENDIX A

Additional Theory on HMMs

Some more theory on HMM are presented in this chapter. The Baum-Welsch training procedure for HMMs is also presented in this chapter. Most of the information in this chapter is taken from (Rabiner, 1989).

A.1 Assumptions of the hidden Markov model

It was mentioned earlier that HMMs have a rich mathematical structure. This is again evident in the following assumptions. These assumptions are made to keep mathematical calculations tractable[1].

1. **The Markov assumption:** The Markov assumption is that the probability of transition from one state to another is only dependent upon the current state.

$$a_{ij} = P(q_{t+1} = j | q_t = i) \tag{A.1}$$

   This is actually called a first order HMM, where a second order HMM would be dependent upon the current state and the previous state. Naturally calculations become increasingly more complex.

2. **The stationarity assumption:** This implies that the state transition matrix is invariant with time. This means that :

$$P(q_{t_1+1} = j | q_{t_1} = i) = P(q_{t_2+1} = j | q_{t_2} = i) \tag{A.2}$$

   This holds for any t_1 and t_2.

[1]These definitions are from Narada Warakagoda website at http://jedlik.phy.bme.hu/ gerjanos/HMM/node3.htm

3. **The output Independence assumption:** This assumption means that the current observation is statistically independent of any previous observation. Let : $O = \{o_1, o_2, o_3, \ldots, o_n\}$ Then for a specific HMM model, λ:

$$P(O|q_1, q_2, q_3, \ldots, q_n, \lambda) = \prod_{t=1}^{T}(o_t|q_t, \lambda) \tag{A.3}$$

A.2 Training the hidden Markov model

Together with the forward procedure, another two parameters need to be introduced before the Baum-Welch re-estimation procedure can be explained. The first is the backward procedure which is very similar to the forward procedure. This is also calculated recursively and works as follows:

Define:

$$\beta_i(i) = P(O_{t+1}, O_{t+2}, \ldots O_T|q_t = S_i, \lambda) \tag{A.4}$$

We can solve for β inductively:

1. Initialisation:

$$\beta_t(i) = 1, \quad 1 \leq i \leq N \tag{A.5}$$

2. Induction:

$$\beta_t(i) = \sum_{j=1}^{N} a_{ij}b_j(O_{t+1}\beta_{t+1}(j)), \qquad t = T-1, T-2, \ldots, 1, \quad 1 \leq i \leq N \tag{A.6}$$

In order to describe the rest of the procedure for re-estimation of the HMM parameters, we first define $\xi_t(i, j)$, the probability of being in state S_j at time t, and state S_j at time $t + 1$, given the model and the observation sequence e.g.

$$\xi_t(i, j) = P(q_t = S_i, q_{t+1} = S_j|O, \lambda) \tag{A.7}$$

This parameter, ξ can now be written in terms of α and β

$$\xi(i, j) = \frac{\alpha_t(i)a_{ij}b_j(O_{t+1})\beta_{t+1}(j)}{\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_t(i)a_{ij}b_j(O_{t+1})\beta_{t+1}(j)} \tag{A.8}$$

A final quantity, γ needs to be defined. This is the probability of being in state S_i at time t, given the observation sequence O, and the model λ.

$$\gamma_t(i) = P(q_t = S_i|O, \lambda) \tag{A.9}$$

In terms of α and β this is:

$$\gamma_t(i) = \frac{\alpha_t(i)\beta_t(i)}{\sum_{j=1}^{N} \alpha_t(i)\beta_t(i)} \tag{A.10}$$

If γ is summed over t, the result is the expected number of transitions made from state S_j. Similarly, summing of $\xi_t(i,j)$ over t (from $t = 1$ to $t = T - 1$) can be interpreted as the expected number of transitions from state S_i to state S_j. Using the above formulae and the counting of event occurrences a method for re-estimation can be given:

$$\bar{\pi}_i = \text{expected number of times in state } \mathrm{S_i} \text{ at time } (\mathrm{t} = 1) \; = \; \gamma_1(\mathrm{i})$$

$$\bar{a}_{ij} = \frac{\sum_{t=1}^{T-1} \xi_t(i,j)}{\sum_{t=1}^{T-1} \gamma_t(i)} \tag{A.11}$$

$$\bar{b}_j(k) = \frac{\sum_{t=1}^{T} \gamma_t(j) \qquad (\text{the number of times in state j and observing symbol } \mathrm{v_k})}{\sum_{t=1}^{T} \gamma_t(j)} \tag{A.12}$$

APPENDIX B

Training of the HMMs

Some results on the training of the HMM are given in this chapter. Convergence plots of the training is shown here.

The training algorithm of the HMM was set to allow no more than 10 iterations. Figure B.1 shows 20 convergence-curves from the training of HMM on sharp tool data and worn tool data. The Baum-Welch algorithm quickly converges to a local optimum. Because the initialisation of the state transition matrix and the state probabilities are done randomly, the end result differs slightly after each training. A basic trend can however be detected.

The y-axis of the figure denotes "Logarithmic likelihood", this is the calculated for the whole training batch. The differences between the training outcomes can be ascribed to the influence of the following:

- The data quality of the training samples. (which were selected randomly)
- The initialisation of the parameters. (which were also selected randomly)

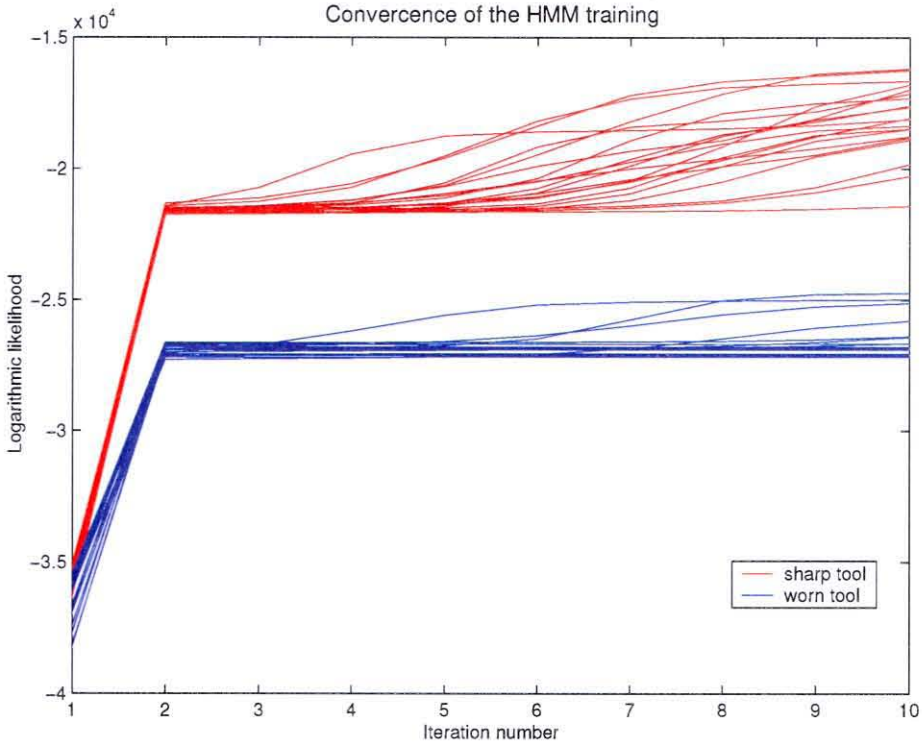

Figure B.1: Some convergence histories of the model training

APPENDIX C

Measurement of tool wear

This appendix will show the state of wear on the tool nose.

C.1 Nose wear

It was mentioned on page 39 the dominant wear mode that was detected in this experiment was nose wear. Nose wear is common at slow cutting speeds.

Tool wear usually has a slow initialisation phase followed by a an almost constant wear phase. The last phase is a very rapid wear growth followed by tool breakage.

In figures C.2 to C.3 the nose of the tool insert is shown. The angle from which is was taken is shown in figure C.1. The edge of the nose at the upper left corner is where nose wear would be seen. It can be seen that there is very little change between the two photos although they were separated by approximately 60 minutes of cutting time. This small change is an indication that the tool is still in the first phase of wear. A photo of a metal ruler calibrated in *millimetres* is shown is figure C.4. This photos was taken at the same magnification as the rest rest of the photos.

The scale that is presented in figure C.4 can also be used for figure 4.7 in the chapter on the experimental set up.

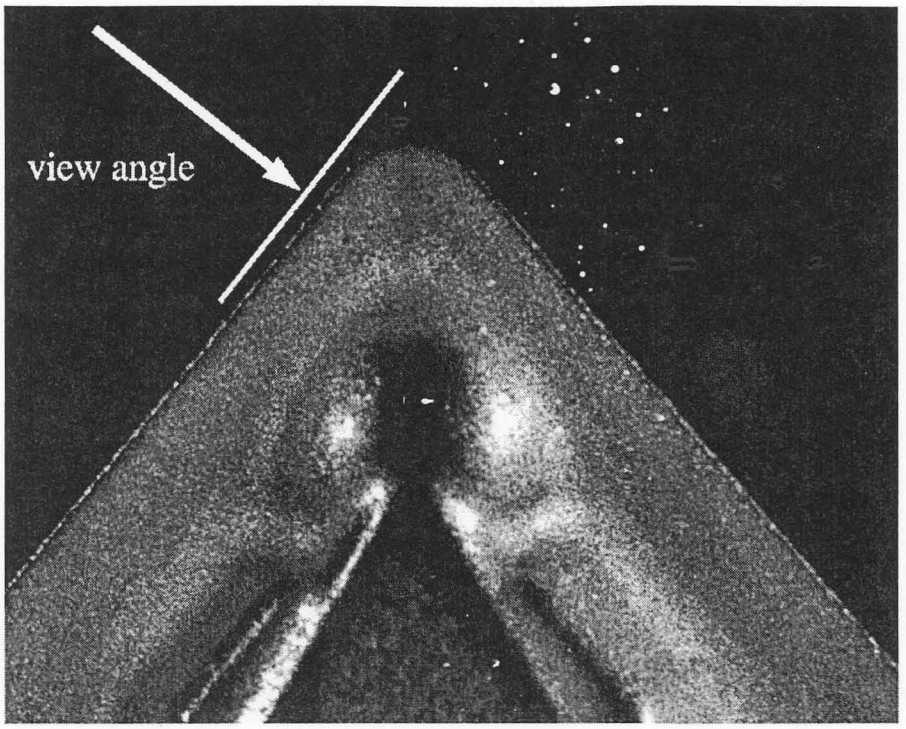

Figure C.1: The photo angle for figures C.2 and C.3.

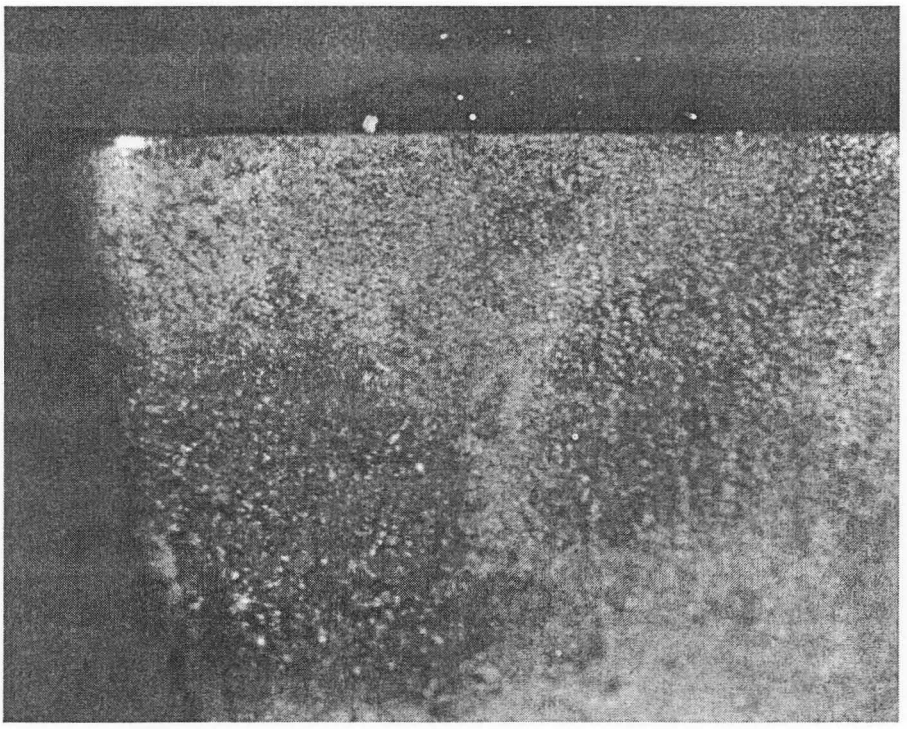

Figure C.2: Nose of a sharp tool

Figure C.3: Nose of a tool where wear has started

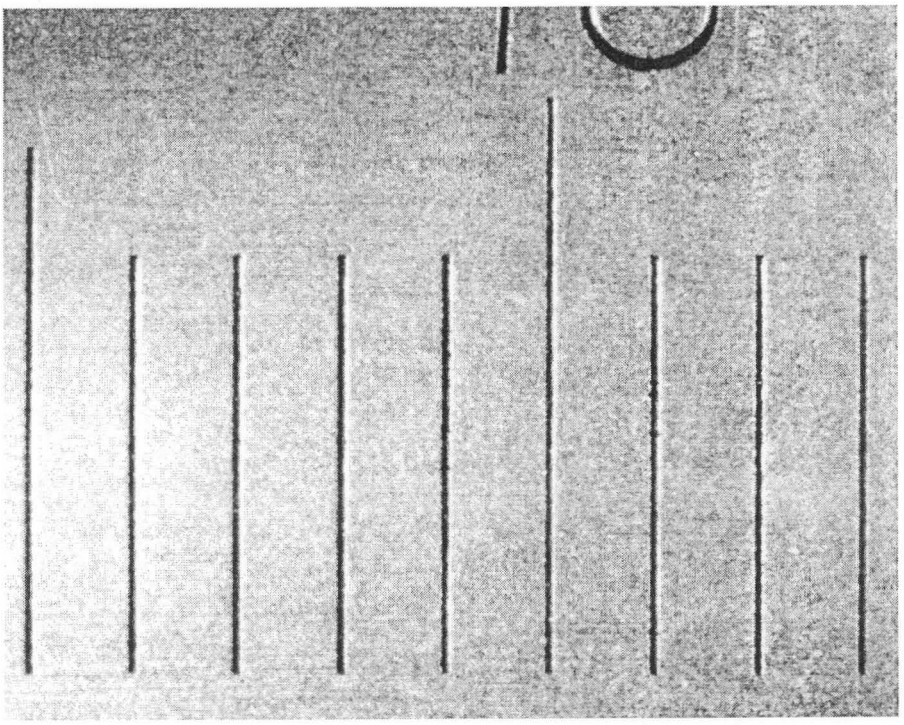

Figure C.4: A ruler calibrated in millimetres.

APPENDIX D

The setup

A few photos of the setup is shown in this chapter. Some more examples of cutting chips that were produced will also be shown.

The photos of the setup shown in figure 4.3 on page 34 will now be shown in this chapter. Figure D.1 shows the cutting tool in action. The cables from the strain gauges are contained in a shielded cable. This cable carries the wires to the strain gauge amplifiers. The strain gauge amplifier together with the anti-alias filters are housed in a metal

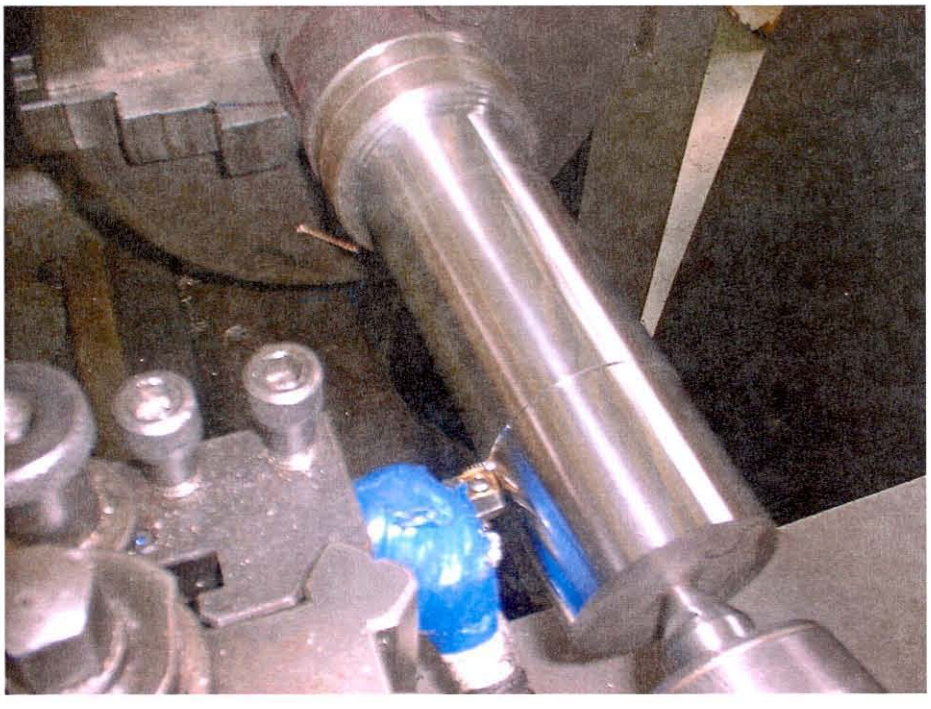

Figure D.1: The cutting tool in action.

container. This is shown in figure D.2. The output of the anti-alias filters are fed into the

PC via the National Instruments A/D card. The computer with the outside connectors for the output from the filters are shown in figure D.3.

Figure D.2: The housing for the strain gauges and filters

Figure D.3: The PC with the outside connectors shown in the upper right half