Memo No.

Date / /

Mo Tu We Th Fr So Su

ID3 举例：

起源：

Target：是否受欢迎：{Y：9，N：5}

Attributes：
- Appearance：{Ah：5=3Y+2N，Good：5=2Y+3N，Great：4=4Y}
- Income：{Low：4=2Y+2N，Good：6=4Y+2N，Great：4=3Y+1N}
- Age：{Younger：7=6Y+1N，Older：7=3Y+4N}
- Profession：{Unstable：6=3Y+3N，Steady：8=6Y+2N}

第一步：系统熵计算

$$H(D) = -\sum_{k=1}^{K} P_k \log_2 P_k = -\frac{9}{14}\log_2\frac{9}{14} - \frac{5}{14}\log_2\frac{5}{14} = 0.94$$

第二步：计算每个特征条件下的熵

Appearance：
$$H(F_{Great}) = -\frac{4}{4}\log_2\left(\frac{4}{4}\right) = 0$$
$$H(F_{Good}) = -\frac{2}{5}\log_2\frac{2}{5} - \frac{3}{5}\log_2\frac{3}{5} = 0.971$$
$$H(F_{Ah}) = -\frac{3}{5}\log_2\frac{3}{5} - \frac{2}{5}\log_2\frac{2}{5} = 0.971$$
$$H(D|F_{App}) = \frac{4}{14}H(F_{Great}) + \frac{5}{14}H(F_{Good}) + \frac{5}{14}H(F_{Ah}) = 0.693$$

Income：
$$H(F_{Great}) = -\frac{3}{4}\log_2\frac{3}{4} - \frac{1}{4}\log_2\frac{1}{4} = 0.811$$
$$H(F_{Good}) = -\frac{4}{6}\log_2\frac{4}{6} - \frac{2}{6}\log_2\frac{2}{6} = 0.918$$
$$H(F_{Low}) = -\frac{2}{4}\log_2\frac{2}{4} - \frac{2}{4}\log_2\frac{2}{4} = 1$$
$$H(D|F_{Inc}) = \frac{4}{14}H(F_{Great}) + \frac{6}{14}H(F_{Good}) + \frac{4}{14}H(F_{Low}) = 0.911$$

Mo Tu We Th Fr Sa Su

Memo No. _____

Date / /

Age: $H(F_{\text{older}}) = -\frac{3}{7}\log_2\frac{3}{7} - \frac{4}{7}\log_2\frac{4}{7} = 0.985$

$H(F_{\text{younger}}) = -\frac{6}{7}\log_2\frac{6}{7} - \frac{1}{7}\log_2\frac{1}{7} = 0.592$

$H(D|F_{\text{Age}}) = \frac{7}{14}H(F_{\text{older}}) + \frac{7}{14}H(F_{\text{younger}}) = 0.789$

Profession: $H(F_{\text{steady}}) = -\frac{6}{8}\log_2\frac{6}{8} - \frac{2}{8}\log_2\frac{2}{8} = 0.84$

$H(F_{\text{unstable}}) = -\frac{3}{6}\log_2\frac{3}{6} - \frac{3}{6}\log_2\frac{3}{6} = 1$

$H(D|F_{\text{pro}}) = \frac{6}{14}H(F_{\text{unstable}}) + \frac{8}{14}H(F_{\text{steady}}) = 0.892$

第三步：计算不同特征值的熵增量：

$G(D|F_{\text{App}}) = H(D) - H(D|F_{\text{App}}) = 0.246$

$G(D|F_{\text{Inc}}) = H(D) - H(D|F_{\text{Inc}}) = 0.029$

$G(D|F_{\text{Age}}) = H(D) - H(D|F_{\text{Age}}) = 0.151$

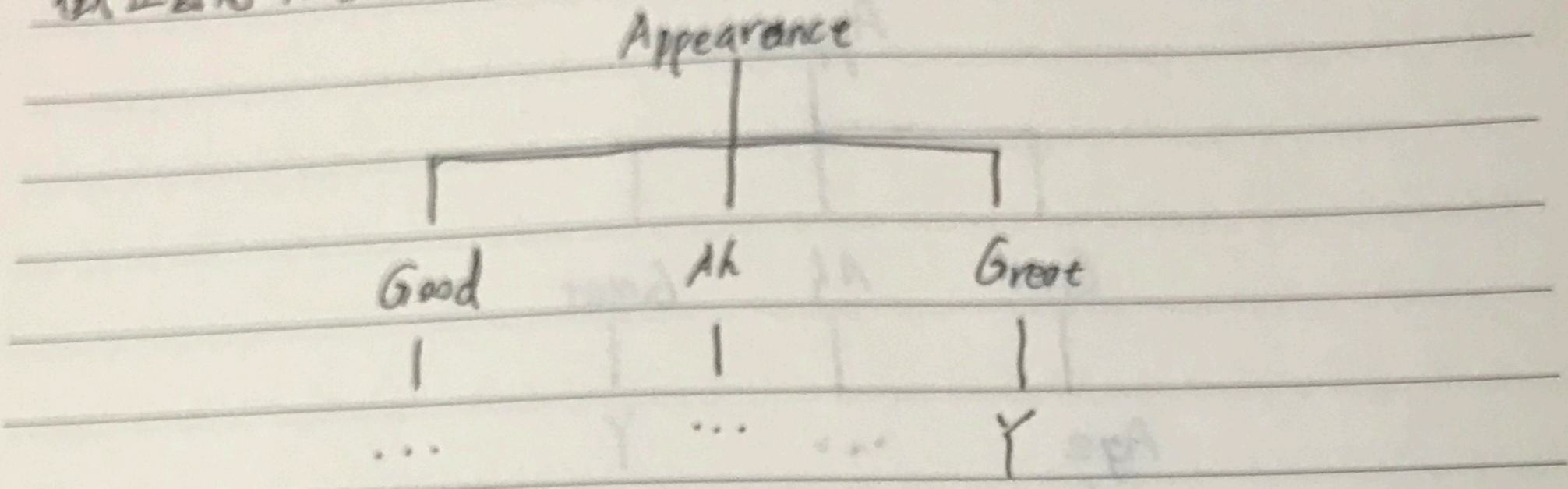
$G(D|F_{\text{pro}}) = H(D) - H(D|F_{\text{pro}}) = 0.048$

显然，上式中Apperance对应的熵增最大，故选Appearance作为决策树的根节点，又Appearance为为3种情况，即含有3个分支，其中Appearance为Great的熵增为0，故不再对Appearance的Great分支进行划分。

Memo No.

Date / /

截止目前决策树图为



接下来对 Appearance 的 Good 分支，包含情况如下：

Appearance_Good：{Y:2, N:3}

Income：{Low: 2=2N, Good: 2=1Y+1N, Great: 1=1Y}

Age：{Younger: 2=2Y, Older: 3=3N}

Profession：{Unstable: 2=1Y+1N, Steady: 3=1Y+2N}

此时，系统熵即为之前的 H(App_Good) = 0.971

类比之前的计算方法，可得：

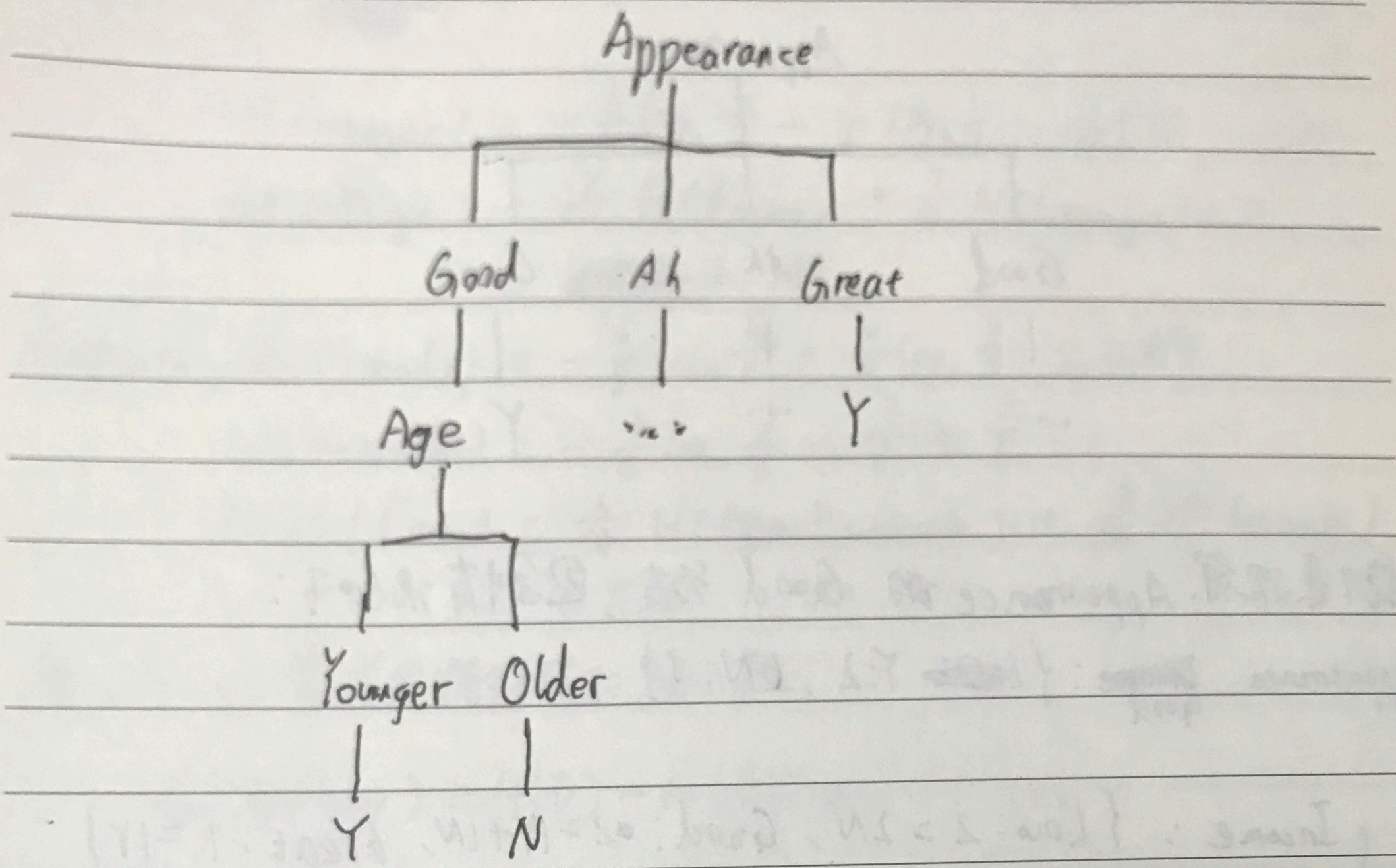
G(F_App|F_Inc) = H(F_App) − H(F_App|F_Inc) = 0.571

G(F_App|F_Age) = H(F_App) − H(F_App|F_Age) = 0.971

G(F_App|F_Pro) = H(F_App) − H(F_App|F_Pro) = 0.020

故 Appearance_Good 的分支选取 Age 作为根节点，其下属之支分支的熵均为0，故不再划分。

此时决策树对更新如下：



同理对Appearance的Ah分支进行计算，可知Profession的熵增最大，故将Profession作为Appearance_Ah的根节点，因此时Profession的两分支熵均为0，故不再划分。最终整理决策树如下：