Discrete Bayesian networks

The parameters of a Bayesian network with multinomial conditional probability distributions can basically be represented as a set of tables that list the probabilities of discrete events for each combination of the discrete values of the parent nodes.

The figure -- on the last slide -- shows a very simple example of a discrete Bayesian network with binary units that can take on the values TRUE and FALSE (indicating, for example, whether or not the grass is wet or whether the sprinkler is used).

Gene expression data can -- for example -- be discretised into three different expression values:
1 - low expression value
2 - medium expression value
3 - high expression value

Bayesian network methodology

- Marriage between graph theory and probability theory.
- Directed acyclic graph (DAG) representing conditional independence relations.
- It is possible to score a network in light of the data: $P(D|M)$, $D$: data, $M$: network structure.
- We can infer how well a particular network explains the observed data.

$$P(A, B, C, D, E, F) = P(A) \cdot P(B|A) \cdot P(C|A) \cdot P(D|B, C) \cdot P(E|D) \cdot P(F|C, D)$$
Bayesian networks

\[ P(\text{graph} \mid \text{data}) = \frac{P(\text{data} \mid \text{graph}) \cdot P(\text{graph})}{P(\text{data})} \]

* \( P(\text{graph} \mid \text{data}) \) ist die posterior (aposteriori) Wahrscheinlichkeit des Graphen (i.S.v. Models) gegeben die beobachteten Daten \( \text{data} \).

* \( P(\text{data} \mid \text{graph}) \) ist die marginale Likelihood des Graphen.

* \( P(\text{graph}) \) ist die prior (apriori) Wahrscheinlichkeit des Graphen (i.S.v. Models).

* \( P(\text{data}) \) ist die Wahrscheinlichkeit der Daten und entspricht einer Normalsierungskonstanten. Es gilt:

\[ P(\text{data}) = \sum_{\text{graphs}} P(\text{data}, \text{graph}) = \sum_{\text{graphs}} P(\text{data} \mid \text{graph}) \cdot P(\text{graph}) \]

\[ P(\text{graph} \mid \text{data}) \propto P(\text{data} \mid \text{graph}) \cdot P(\text{graph}) \]

\( q \) ist ein Vektor der (unbekannten) Parameter, die zur Parametrisierung (i.S.v. statistischer Modellierung) des Graphen benötigt werden.

\[ P(\text{data} \mid \text{graph}) = \int P(\text{data}, q \mid \text{graph}) dq \]

\[ \int P(\text{data} \mid \text{graph}, q) dq \]

[Graph and Parameter Diagrams]
Der Graph spezifiziert durch die Kanten die Faktorisierung:
\[ P(C,S,R,W) = P(C) P(S|C) P(R|C) P(W|S,R) \]
Hier ist zusätzlich ein statistisches Modell (inklusive der Parameter \( q \)) gegeben.

Beispiel:

\[ \text{Der Graph spezifiziert durch die Kanten die Faktorisierung:} \]
\[ P(C,S,R,W) = P(C) P(S|C) P(R|C) P(W|S,R) \]
Hier ist zusätzlich ein statistisches Modell (inklusive der Parameter \( q \)) gegeben.

\[ P(D|G) = \int P(D,q,G) dq = \int P(D,q,G) \cdot P(q|G) dq \]

where \( D \) is the \( n \)-by-\( m \) data set matrix, \( G \) is a directed acyclic graph (DAG), and \( q \) is the parameter vector

\[ P(D|G) = \prod_{i=1}^{n} \prod_{j=1}^{m} P(X_i = D_{i,j} | \text{pa}(X_i)) = D_{\text{pa}(X_i),j}^{q_i} \]

where \( n \) is the number of nodes, \( m \) is the number of observations (realisations), \( D_{i,j} \) is the \( j \)-th realisation of the \( i \)-th domain variable \( X_i \), \( q_i \) is the set of parameters associated with the \( i \)-th local distribution, \( \text{pa}(X_i) \) is the set of parent nodes of \( X_i \), \( D_{\text{pa}(X_i),j} \) is the \( j \)-th realisation of the parent nodes of \( X_i \).
Parameter independence and modularity

Assumption 1 -- Parameter independence --

\[ P(q_i | G) = \prod_{i=1}^{n} P(q_i | G) \]

Parameter independence means that the distributions of the parameters \( q_i \) (for the local probability distributions) are stochastically independent

Assumption 2 -- Parameter modularity --

\[ P(q_i | G) \geq P(q_i | X_i, \text{pa}(X_i)) \quad (i=1,...,n) \]

Parameter modularity means that the density of the parameters \( q_i \) (of each local distribution) depends on the local structure (i.e. the parents of \( X_i \)) only

Assumptions 1&2: \[ \prod_{i=1}^{n} P(q_i | \text{pa}(X_i)) \]

Bayesian networks

\[ P(D | G) = \int P(D, q | G) dq = \int P(D | q, G) \cdot P(q | G) dq \]

Then we obtain -- under parameter independence/modularity --

\[ = \prod_{i=1}^{n} \left( P(q_i | \text{pa}(X_i)) \cdot \prod_{j=1}^{m} P(X_j = D_{i,j} | \text{pa}(X_j), q_i) \right) dq_i, ..., dq_n \]

\[ = \prod_{i=1}^{n} \Psi_i (\text{pa}(X_i), D^{(X_i)}_{\text{pa}(X_i)}) \]

Bayesian networks

\[ P(D | G) = \int P(D, q | G) dq = \int P(D | q, G) \cdot P(q | G) dq \]

Then we obtain -- under parameter independence/modularity --

\[ = \prod_{i=1}^{n} \left( P(q_i | \text{pa}(X_i)) \cdot \prod_{j=1}^{m} P(X_j = D_{i,j} | \text{pa}(X_j), q_i) \right) dq_i, ..., dq_n \]

Closed form solution -- of this functional form -- for the BDe scoring metric

Model assumption 1: Given the value combination \( j \) of its parent nodes in \( \text{pa}(X) \) each domain variable \( X_i \) is multinomial distributed \( M(\theta_{i,j,1},...\theta_{i,j,r}) \) distributed with unknown parameters \( \theta_{i,j,1},...\theta_{i,j,r} \) where \( \theta_{i,j,1}+...+\theta_{i,j,r}=1 \) and the values 1,...,r reflect the \( r \) discrete values the variable \( X_i \) can take on.

Model assumption 2: The unknown parameters \( \theta_{i,j,1},...\theta_{i,j,r} \) are Dirichlet distributed distributed whereby the 'exponents of the Dirichlet distribution (i.e. the hyperparameters) \( \alpha_{i,j,1},...\alpha_{i,j,r} > 0 \) must be specified by the user.

Density function of the Dirichlet distribution

\[ P(\theta_{i,j,1},...\theta_{i,j,r}) = \frac{\Gamma(\sum_{i=1}^{r} \alpha_{i,j,i})}{\prod_{i=1}^{r} \Gamma(\alpha_{i,j,i})} \prod_{i=1}^{r} \theta_{i,j,i}^{\alpha_{i,j,i}-1} \]
(Marginal) BDe Likelihood

\[ P(D \mid G) = \prod_{i=1}^{n} \prod_{j=1}^{q_i} \prod_{k=1}^{r_i} \frac{\Gamma(N_{i,j} + \alpha_{i,j,k})}{\Gamma(N_{i,j} + \alpha_{i,j})} \frac{\Gamma(N_{i,j,k} + \alpha_{i,j,k})}{\Gamma(N_{i,j,k})} \]

Where \( i \) indicates the node \( X_i \), \( j \) indicates a value combination of the parent nodes of \( X_i \) (\( j = 1, \ldots, q_i \)), and \( k \) is a possible realisation of \( X_i \) (\( k = 1, \ldots, r_i \)).

Moreover:

- \( N_{i,j,k} \) is the number of observations in the data set \( D \) in which \( X_i = k \) and the parent nodes have taken their \( j \)-th value combination.
- \( N_{i,j} \) is the number of observations in the data set \( D \) for which the parent nodes of \( X_i \) have taken their \( j \)-th value combination.

And the hyperparameters can/must be set as follows to ensure that equivalent DAGs have the same likelihood score:

\[ \alpha_{i,j,k} = \frac{\alpha}{q_i \cdot r_i}, \quad \alpha = \sum_{k=1}^{r_i} \alpha_{i,j,k} \]

The hyperparameters can be interpreted as pseudocounts with 'total prior precision' \( \alpha \) – for example \( \alpha = 1 \).

Example:

\[ D = \begin{pmatrix}
T & F & F & F & T & T & T \\
F & T & T & T & F & F & T \\
F & T & T & F & F & T & T \\
F & T & T & F & F & T & T \\
\end{pmatrix} \]

Example:

\[ D = \begin{pmatrix}
T & F & F & T \\
F & T & T & F \\
F & T & T & F \\
F & T & T & F \\
\end{pmatrix} \]
### Example:

**Cloudy**

<table>
<thead>
<tr>
<th>Sprinkler</th>
<th>Rain</th>
<th>Wet grass</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th><strong>We set:</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Cloudy = $X_1$ TRUE = 1 and FALSE = 2</td>
</tr>
</tbody>
</table>

And there is only one possible parent value combination, as there are no parent nodes, that is $q_1=1$

$$D = \begin{pmatrix} T & F & F & T & F & T & T & T & T & T \\ F & T & T & F & T & T & F & F & T & T \\ F & T & F & F & F & F & T & T & T & T \\ F & T & T & F & F & F & T & T & T & T \end{pmatrix}$$

**Combinations of parents**

<table>
<thead>
<tr>
<th></th>
<th><strong>Cloudy</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TRUE</strong></td>
<td>$N_{C,1,T}$ $\Rightarrow 6$</td>
</tr>
<tr>
<td><strong>FALSE</strong></td>
<td>$N_{C,1,F}$ $\Rightarrow 4$</td>
</tr>
</tbody>
</table>

### Example:

We set:

Cloudy = $X_1$, TRUE = 1 and FALSE = 2

And there is only one possible parent value combination, as there are no parent nodes, that is $q_1=1$

$$D = \begin{pmatrix} T & F & F & F & T & F & T & T & T & T \\ F & T & T & T & F & F & T & F & T & T \\ F & T & F & F & F & F & T & F & T & T \\ F & T & T & F & F & F & T & F & T & T \end{pmatrix}$$

**Combinations of parents**

<table>
<thead>
<tr>
<th></th>
<th><strong>Cloudy</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TRUE</strong></td>
<td>$N_{C,1,T}$ $\Rightarrow 6$</td>
</tr>
<tr>
<td><strong>FALSE</strong></td>
<td>$N_{C,1,F}$ $\Rightarrow 4$</td>
</tr>
</tbody>
</table>

### Example:

**We have:**

$q_1=1$, and $r_1=2$ and we set the total prior precision to $\alpha=1$:

$$\alpha_{i,j,k} = \frac{\alpha}{q_1 \cdot r_1} = \frac{1}{1 \cdot 2} = 0.5$$

$$D = \begin{pmatrix} T & F & F & T & F & T & T & T & T & T \\ F & T & F & T & T & F & F & T & T & T \\ F & T & F & F & F & F & T & T & T & T \\ F & T & T & F & F & F & T & T & T & T \end{pmatrix}$$

**Example:**

<table>
<thead>
<tr>
<th><strong>Cloudy</strong></th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th></th>
<th><strong>Cloudy</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TRUE</strong></td>
<td>$N_{C,1,T}$ $\Rightarrow 6$</td>
</tr>
<tr>
<td><strong>FALSE</strong></td>
<td>$N_{C,1,F}$ $\Rightarrow 4$</td>
</tr>
</tbody>
</table>

### Example:

**We set:**

Cloudy = $X_1$, TRUE = 1 and FALSE = 2

And there is only one possible parent value combination, as there are no parent nodes, that is $q_1=1$

$$D = \begin{pmatrix} T & F & F & F & T & F & T & T & T & T \\ F & T & T & F & T & F & F & T & T & T \\ F & T & F & F & F & F & T & F & T & T \\ F & T & T & F & F & F & T & T & T & T \end{pmatrix}$$

**Combinations of parents**

<table>
<thead>
<tr>
<th></th>
<th><strong>Cloudy</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TRUE</strong></td>
<td>$N_{C,1,T}$ $\Rightarrow 6$</td>
</tr>
<tr>
<td><strong>FALSE</strong></td>
<td>$N_{C,1,F}$ $\Rightarrow 4$</td>
</tr>
</tbody>
</table>

### Example:

<table>
<thead>
<tr>
<th><strong>No parents</strong></th>
<th><strong>Cloudy</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>$j=1$</td>
<td>$N_{i,1,j}=6$</td>
</tr>
<tr>
<td>$k=1$</td>
<td>$N_{i,1,k}=4$</td>
</tr>
</tbody>
</table>

**No parents**

<table>
<thead>
<tr>
<th></th>
<th><strong>Cloudy</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TRUE</strong></td>
<td>$N_{C,1,T}$ $\Rightarrow 6$</td>
</tr>
<tr>
<td><strong>FALSE</strong></td>
<td>$N_{C,1,F}$ $\Rightarrow 4$</td>
</tr>
</tbody>
</table>

**pseudocounts**

<table>
<thead>
<tr>
<th></th>
<th><strong>Cloudy</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TRUE</strong></td>
<td>$\alpha_{i,1,k}=0.5$</td>
</tr>
<tr>
<td><strong>FALSE</strong></td>
<td>$\alpha_{i,1,k}=0.5$</td>
</tr>
</tbody>
</table>
\[
\Psi_i(\{\}, D^{(C)}_{\{\}}) = \frac{1}{(N_{i,j} + \alpha_{i,j,k})} \prod_{k=1}^{2} \frac{\Gamma(N_{i,j,k} + \alpha_{i,j,k})}{\Gamma(\alpha_{i,j,k})} 
\]

<table>
<thead>
<tr>
<th>No parents</th>
<th>Cloudy</th>
<th>(q_i=1)</th>
<th>Pseudocounts (a_{1,1,k})</th>
</tr>
</thead>
<tbody>
<tr>
<td>(j=1)</td>
<td>(k=1)</td>
<td>(k=2)</td>
<td>(j=1) (k=1) (0.5) (0.5)</td>
</tr>
<tr>
<td></td>
<td>(N_{i,1,1}=6) (N_{i,1,2}=4)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[
\Psi_i(\{\}, D^{(C)}_{\{\}}) = \frac{1}{(N_{i,j} + \alpha_{i,j,k})} \prod_{k=1}^{2} \frac{\Gamma(N_{i,j,k} + \alpha_{i,j,k})}{\Gamma(\alpha_{i,j,k})} 
\]

<table>
<thead>
<tr>
<th>No parents</th>
<th>Cloudy</th>
<th>(q_i=1)</th>
<th>Pseudocounts (a_{1,1,k})</th>
</tr>
</thead>
<tbody>
<tr>
<td>(j=1)</td>
<td>(k=1)</td>
<td>(k=2)</td>
<td>(j=1) (k=1) (0.5) (0.5)</td>
</tr>
<tr>
<td></td>
<td>(N_{i,1,1}=6) (N_{i,1,2}=4)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[
\Psi_i(\{\}, D^{(C)}_{\{\}}) = \frac{\Gamma(\alpha_{i,1,k})}{\Gamma(N_{i,1} + \alpha_{i,1,k})} \left( \frac{\Gamma(N_{i,1,1} + \alpha_{i,1,1,k})}{\Gamma(\alpha_{i,1,1,k})} \right) 
\]

<table>
<thead>
<tr>
<th>No parents</th>
<th>Cloudy</th>
<th>(q_i=1)</th>
<th>Pseudocounts (a_{1,1,k})</th>
</tr>
</thead>
<tbody>
<tr>
<td>(j=1)</td>
<td>(k=1)</td>
<td>(k=2)</td>
<td>(j=1) (k=1) (0.5) (0.5)</td>
</tr>
<tr>
<td></td>
<td>(N_{i,1,1}=6) (N_{i,1,2}=4)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[
\Psi_i(\{\}, D^{(C)}_{\{\}}) = \frac{1}{(N_{i,j} + \alpha_{i,j,k})} \prod_{k=1}^{2} \frac{\Gamma(N_{i,j,k} + \alpha_{i,j,k})}{\Gamma(\alpha_{i,j,k})} 
\]

<table>
<thead>
<tr>
<th>No parents</th>
<th>Cloudy</th>
<th>(q_i=1)</th>
<th>Pseudocounts (a_{1,1,k})</th>
</tr>
</thead>
<tbody>
<tr>
<td>(j=1)</td>
<td>(k=1)</td>
<td>(k=2)</td>
<td>(j=1) (k=1) (0.5) (0.5)</td>
</tr>
<tr>
<td></td>
<td>(N_{i,1,1}=6) (N_{i,1,2}=4)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Example:

$$D = \begin{pmatrix} T & F & F & T & F & F & T & T & T & T \\ T & T & T & F & F & F & T & T & T & T \\ F & T & T & F & F & F & T & T & T & T \\ F & T & T & F & F & F & T & T & T & T \end{pmatrix}$$

Sprinkler
Rain
Wet grass

$$D = \begin{pmatrix} T & F & F & T & F & F & T & T & T & T \\ T & T & T & F & F & F & F & F & F & F \\ F & T & T & T & F & F & F & T & T & T \\ F & T & T & F & T & T & F & T & T & T \end{pmatrix}$$

Sprinkler
Rain
Wet grass
### Example:

<table>
<thead>
<tr>
<th>Combinations of parents</th>
<th>Wet grass</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sprinkler</strong></td>
<td><strong>Rain</strong></td>
</tr>
<tr>
<td>TRUE</td>
<td>TRUE</td>
</tr>
<tr>
<td>FALSE</td>
<td>TRUE</td>
</tr>
<tr>
<td>FALSE</td>
<td>FALSE</td>
</tr>
</tbody>
</table>

\[
D = \begin{pmatrix}
T & F & F & T & F & T & T \\
F & T & F & T & F & T & T \\
T & F & F & F & F & T & T \\
F & T & F & F & F & T & T \\
\end{pmatrix}
\]

### Example:

<table>
<thead>
<tr>
<th>Combinations of parents</th>
<th>Wet grass</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sprinkler</strong></td>
<td><strong>Rain</strong></td>
</tr>
<tr>
<td>TRUE</td>
<td>TRUE</td>
</tr>
<tr>
<td>TRUE</td>
<td>FALSE</td>
</tr>
<tr>
<td>FALSE</td>
<td>TRUE</td>
</tr>
<tr>
<td>FALSE</td>
<td>FALSE</td>
</tr>
</tbody>
</table>

\[
D = \begin{pmatrix}
T & F & F & F & T & T & T \\
F & T & T & T & F & F & T \\
T & F & F & T & F & T & T \\
F & T & F & F & T & T & T \\
\end{pmatrix}
\]

### Example:

**We set:**

Wet grass = \( X_4 \) and furthermore \( \text{TRUE} = 1 \) and \( \text{FALSE} = 2 \)

And there are are four possible parent value combinations:

- \( \text{TRUE,TRUE} = 1 \)
- \( \text{TRUE,FALSE} = 2 \)
- \( \text{FALSE,TRUE} = 3 \)
- \( \text{FALSE,FALSE} = 4 \)
Example:

We set:

Wet grass = $X_i$ TRUE = 1 and FALSE = 2

And there are four possible parent value combinations:

(TRUE,TRUE)=1, (TRUE;FALSE)=2, (FALSE;TRUE)=3, and (FALSE,FALSE)=4

We have:

$q_4=4$, and $r_4=2$ and we set the total prior precision to $\alpha=1$:

$$\alpha_{q_4} = \frac{\alpha}{q_4 \cdot r_4} = \frac{1}{4 \cdot 2} = \frac{1}{8} = 0.125$$

Example:

<table>
<thead>
<tr>
<th>Combinations of parents</th>
<th>Wet grass $i=4$ pseudocounts $\alpha_{4,j,k}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$k=1$</td>
</tr>
<tr>
<td>j=1</td>
<td>0.125</td>
</tr>
<tr>
<td>j=2</td>
<td>0.125</td>
</tr>
<tr>
<td>j=3</td>
<td>0.125</td>
</tr>
<tr>
<td>j=4</td>
<td>0.125</td>
</tr>
</tbody>
</table>

Example:

The fourth local score can be computed from these values...

Example:

We have:

$q_4=4$, and $r_4=2$ and thus with total prior precision $\alpha=1$:

$$\alpha_{4,j} = \frac{\alpha}{q_4 \cdot r_4} = \frac{1}{4 \cdot 2} = \frac{1}{8} = 0.125$$

The fourth local score can be computed from these values...

<table>
<thead>
<tr>
<th>Combinations of parents</th>
<th>Wet grass $i=4$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$k=1$</td>
</tr>
<tr>
<td>j=1</td>
<td>$N_{4,1,1}=2$</td>
</tr>
<tr>
<td>j=2</td>
<td>$N_{4,2,1}=3$</td>
</tr>
<tr>
<td>j=3</td>
<td>$N_{4,3,1}=1$</td>
</tr>
<tr>
<td>j=4</td>
<td>$N_{4,4,1}=0$</td>
</tr>
</tbody>
</table>
Summary – BDe score

\[
P(\text{graph} \mid \text{data}) = \frac{P(\text{data} \mid \text{graph}) \cdot P(\text{graph})}{P(\text{data})} \propto P(\text{data} \mid \text{graph}) \cdot P(\text{graph})
\]

\[
= P(\text{graph}) \cdot \int P(\text{data}, \theta(\text{graph}) \mid \text{graph}) d\theta(\text{graph})
\]

\[\text{BDe metric: } \rightarrow \text{ closed form solution}\]

\[= \text{score}_{\text{BDe}}(\text{graph} \mid \text{data})\]

Why Bayesian approaches?

(Marginal) Likelihood -- Bayesian network approach:

\[
P(D \mid G) = \int P(D, q \mid G) dq = \int P(D \mid q, G) \cdot P(q \mid G) dq
\]

→ This approach guards against over-fitting, as it includes a penalty for unnecessary complexity!

(Maximal) Likelihood – 'Frequentist' approach:

\[
P(D \mid \hat{q}_{ML}, G)
\]

Idea: For a given network structure G compute the maximum likelihood estimators for the unknown parameters and plug them into the likelihood function to obtain the (maximal) likelihood value.

Why Bayesian approaches?

Consider a simple 1-dimensional example: A statistical model M with one parameter, namely q:

\[
P(D \mid M) = \int P(D \mid q, M) \cdot P(q \mid M) dq
\]

We assume that the parameter prior \(P(q \mid M)\) is nearly a uniform distribution so that it can be approximated by a uniform distribution on an interval of length \(\Delta q_{\text{prior}}\) and that the likelihood function \(L(q) = P(D \mid q, M)\) is unimodal and peaked around its maximum.
Why Bayesian approaches?

\[ P(D \mid M) = \int P(D \mid q, M) \cdot P(q \mid M) dq \]

over an interval of length \( \Delta q_{\text{post}} \)
and 0 outwith this interval

The larger \( \Delta q_{\text{post}} \) the higher the marginal likelihood. That is, the larger the area in which the parameter \( q \) explains the data well, the higher the marginal likelihood of the model.

Complex models tend to have greater likelihoods but only for a particular parameter, that is in a tiny area of length \( \Delta q_{\text{post}} \) around the maximum and so the complexity is penalized, as this leads to a low marginal likelihood value.

\[ P(D \mid M) \approx P(D \mid \hat{q}_{\text{ML}}, M) \cdot \frac{\Delta q_{\text{post}}}{\Delta q_{\text{prior}}} \]

‘Occams factor’