Decision Networks & Hidden Markov Models

CS 4804 Fall 2020
Virginia Tech
Today’s Topics

• Decision Networks
• The Value of Information
• Hidden Markov Models
Decision Networks

• Bayesian Networks
• Node types for actions and utilities
• **Chance node:** Random variable
• **Decision node:** Points where the decision maker has a choice of actions
• **Utility node:** utility function and is deterministic
Example: Decision Networks

Diagram:

- Airport Site
- Air Traffic
- Litigation
- Construction
- Safety
- Quietness
- Frugality
- U
Example: Simplified Representation of DN

- The chance nodes are omitted
- The utility node is connected directly to the current-state node and the decision node
- The utility node represents the expected utility associated with each action
- The utility node is associated with an action-utility function (Q-function)
DN Evaluation

1. Set the evidence variables for the current state
2. For each possible value of the decision node:
   a) Set the decision node to that value
   b) Calculate the **posterior probabilities** for the parent nodes of the utility node, using a standard probabilistic inference algorithm
   c) Calculate the resulting utility for the action
3. Return the action with the highest utility
Maximum Expected Utility

Umbrella = leave

\[ EU(\text{leave}) = \sum_w P(w)U(\text{leave}, w) \]
\[ = 0.7 \cdot 100 + 0.3 \cdot 0 = 70 \]

Umbrella = take

\[ EU(\text{take}) = \sum_w P(w)U(\text{take}, w) \]
\[ = 0.7 \cdot 20 + 0.3 \cdot 70 = 35 \]

Optimal decision = leave

\[ \text{MEU}(\emptyset) = \max_a EU(a) = 70 \]

---

* This example is adopted from UC Berkeley CS 188
Maximum Expected Utility

Umbrella = leave

\[
EU(\text{leave}|\text{bad}) = \sum_{w} P(w|\text{bad})U(\text{leave}, w) \\
= 0.34 \cdot 100 + 0.66 \cdot 0 = 34
\]

Umbrella = take

\[
EU(\text{take}|\text{bad}) = \sum_{w} P(w|\text{bad})U(\text{take}, w) \\
= 0.34 \cdot 20 + 0.66 \cdot 70 = 53
\]

Optimal decision = take

\[
\text{MEU}(F = \text{bad}) = \max_{a} EU(a|\text{bad}) = 53
\]

* This example is adopted from UC Berkeley CS 188
Inference in Ghostbusters

- A ghost is in the grid somewhere
- Sensor readings tell how close a square is to the ghost
  - On the ghost: red
  - 1 or 2 away: orange
  - 3 or 4 away: yellow
  - 5+ away: green

- Sensors are noisy, but we know \( P(\text{Color} \mid \text{Distance}) \)

<table>
<thead>
<tr>
<th></th>
<th>( P(\text{red} \mid 3) )</th>
<th>( P(\text{orange} \mid 3) )</th>
<th>( P(\text{yellow} \mid 3) )</th>
<th>( P(\text{green} \mid 3) )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.05</td>
<td>0.15</td>
<td>0.5</td>
<td>0.3</td>
</tr>
</tbody>
</table>

Slide credits: http://ai.berkeley.edu
Value of Information

• Information value theory:
  – Enables an agent to choose what information to acquire
  – Involves a simplified form (belief state) of sequential decision making
  – Agent’s action will be affected by the value of observation
  – Observing new evidence almost always has some cost
Value of Perfect Information (VPI)

• Depends on the current state of information
• It can change as more information is acquired
• Mathematically quantifies the amount an agent’s maximum expected utility is expected to increase if it observes some new evidence
• Is it worth to get more evidence to help decide action to make?
  – Cost of get new evidence
  – New MEU after get this new evidence
Value of Perfect Information (VPI)

• Current maximum utility given current evidence $e$: 

$$MEU(e) = \max_a \sum_s P(s|e)U(s,a)$$

• Assume we observed new evidence $e'$:

$$MEU(e,e') = \max_a \sum_s P(s|e,e')U(s,a)$$

• We don’t know what new evidence $e'$ we’ll get. We compute MEU for $E'$

$$MEU(e,E') = \sum_{e'} P(e'|e)MEU(e,e')$$

• VPI: How much MEU goes up by reveling $E'$ then acting, over acting now

$$VPI(E'|e) = MEU(e,E') - MEU(e)$$
Example: VPI

- MEU with no evidence:
  \[ MEU(\emptyset) = \max\{0.7 \times 100 + 0.3 \times 0, 0.7 \times 20 + 0.3 \times 70\} = \max\{70, 35\} = 70 \]
- MEU if forecast is bad:
  \[ MEU(F=\text{bad}) = \max\{0.34 \times 20 + 0.66 \times 70, 0.34 \times 100 + 0.66 \times 0\} = 53 \]
- MEU if forecast is good:
  \[ MEU(F=\text{good}) = 95 \]
- \[ MEU(e,E') = MEU(F) = P(F = \text{good}) \times MEU(F = \text{good}) + P(F = \text{bad}) \times MEU(F = \text{bad}) = 0.59 \times 95 + 0.41 \times 53 = 77.78 \]
- \[ VPI(F) = MEU(F) - MEU(\emptyset) = 77.78 - 70 = 7.78 \]
VPI Properties

• **Nonnegative**: \( \forall j \ VPI(E_j) \geq 0 \)

Observing new information always allows you to make a *more informed* decision, and so your maximum expected utility can only increase (or stay the same if the information is irrelevant for the decision you must make)

• **Nonadditive**: \( VPI(E_j, E_k) \neq VPI(E_j) + VPI(E_k) \)

In general, observing some new evidence \( E_j \) might change how much we care about \( E_k \); therefore we can’t simply add the VPI of observing \( E_j \) to the VPI of observing \( E_k \) to get the VPI of observing both of them

• **Order-independent**:\[
\begin{align*}
VPI(E_j, E_k) &= VPI(E_j) + VPI(E_k|E_j) = VPI(E_k) + VPI(E_j|E_k) = VPI(E_k, E_j)
\end{align*}
\]

Observing multiple new evidences yields the same gain in maximum expected utility regardless of the order of observation
POMDPs

• MDPs have:
  – States $S$
  – Actions $A$
  – Transition function $P(s' |s,a)$ (or $T(s,a,s')$)
  – Rewards $R(s,a,s')$

• POMDPs add:
  – Observations $O$
  – Observation function $P(o|s)$ (or $O(s,o)$)

• POMDPs are MDPs over belief states $b$ (distributions over $S$)
Example: Ghostbusters

- In (static) Ghostbusters:
  - Belief state determined by evidence to date \( \{e\} \)
  - Tree really over evidence sets
  - Probabilistic reasoning needed to predict new evidence given past evidence

- Solving POMDPs
  - One way: use truncated expectimax to compute approximate value of actions
  - What if you only considered busting or one sense followed by a bust?
  - You get a VPI-based agent!

Slide credits: http://ai.berkeley.edu
POMDPs as Decision Networks

MDPs have:
- States $S$
- Actions $A$
- Transition function $P(s' | s, a)$ (or $T(s, a, s')$)
- Rewards $R(s, a, s')$

POMDPs add:
- Observations $O$
- Observation function $P(o | s)$ (or $O(s, o)$)

Slide credits: http://ai.berkeley.edu
Time and Uncertainty

- Things are not always fixed, it could change rapidly over time
- Assess the current state from the history of evidence and to predict the outcomes of actions
- Examples:
  - Tracking the location of a robot
  - Tracking the economic activity
  - Speech recognition
  - Medical monitoring
Markov Models

- $X_t$: the set of state variables at time $t$
- **Transition model** $P(X_t|X_{t-1})$: specify how the state evolves over time
- **Stationarity** assumption: transition probabilities are the same at all times
- A (growable) BN: We can always use generic BN reasoning on it if we truncate the chain at a **fixed** length
Markov Assumption

- Past and future independent given the present
  \(-X_{t+1}\) is independent of \(X_0,...,X_{t-1}\) given \(X_t\)
- The current state depends on only a finite **fixed** number of previous states
- Each time step only depends on the previous
- Processes satisfying this assumption are called Markov processes or Markov chains
- Simplest: first-order Markov process
  \(-The current state depends only on the previous state and not on any earlier states\)
- Joint distribution \(P(X_{0:t}) = P(X_0) \prod_{i=1:t} P(X_i \mid X_{i-1})\)
Example: Markov Chains (Markov Processes)

- States: \( X = \{ \text{rain, sun} \} \)
- Initial distribution: 1.0 sun
- CPT \( P(X_t \mid X_{t-1}) \):

<table>
<thead>
<tr>
<th>( X_{t-1} )</th>
<th>( X_t )</th>
<th>( P(X_t \mid X_{t-1}) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>sun</td>
<td>sun</td>
<td>0.9</td>
</tr>
<tr>
<td>sun</td>
<td>rain</td>
<td>0.1</td>
</tr>
<tr>
<td>rain</td>
<td>sun</td>
<td>0.3</td>
</tr>
<tr>
<td>rain</td>
<td>rain</td>
<td>0.7</td>
</tr>
</tbody>
</table>

* This example is adopted from UC Berkeley CS 188
Example: Markov Chains (Markov Processes)

- Initial distribution: 1.0 sun
- What is the probability distribution after one step?

\[ P(X_t | X_{0:t-1}) = P(X_t | X_{t-1}) \]

\[
P(X_2 = \text{sun}) = \sum_{x_1} P(x_1, X_2 = \text{sun}) = \sum_{x_1} P(X_2 = \text{sun}|x_1)P(x_1)
\]

\[
P(X_2 = \text{sun}) = P(X_2 = \text{sun}|X_1 = \text{sun})P(X_1 = \text{sun}) + P(X_2 = \text{sun}|X_1 = \text{rain})P(X_1 = \text{rain})
\]

\[
0.9 \cdot 1.0 + 0.3 \cdot 0.0 = 0.9
\]
Example: Markov Chains (Markov Processes)

• Initial distribution: 1.0 sun
• What is the probability distribution after two step?

\[ P(X_t | X_{0:t-1}) = P(X_t | X_{t-1}) \]

\[
P(X_3 = sun) = \sum_{x_2} P(x_2, X_3 = sun) = \sum_{x_2} P(X_3 = sun|x_2)P(x_2)
\]

\[
P(X_3 = sun) = P(X_3 = sun|X_2 = sun)P(X_2 = sun) + P(X_3 = sun|X_2 = rain)P(X_2 = rain)
\]

\[0.9 \times 0.9 + 0.3 \times 0.1 = 0.84\]
Mini-Forward Algorithm

• Question: What’s $P(X)$ on some day $t$?

\[ P(x_1) = \text{known} \]
\[ P(x_t) = \sum_{x_{t-1}} P(x_{t-1}, x_t) \]
\[ = \sum_{x_{t-1}} P(x_t \mid x_{t-1}) P(x_{t-1}) \]

Forward simulation
Example: Mini-Forward Algorithm

- From initial observation of sun

\[
\begin{align*}
\mathbb{P}(X_1) &= \begin{pmatrix} 1.0 \\ 0.0 \end{pmatrix} & \mathbb{P}(X_2) &= \begin{pmatrix} 0.9 \\ 0.1 \end{pmatrix} & \mathbb{P}(X_3) &= \begin{pmatrix} 0.84 \\ 0.16 \end{pmatrix} & \mathbb{P}(X_4) &= \begin{pmatrix} 0.804 \\ 0.196 \end{pmatrix} & \mathbb{P}(X_\infty) &= \begin{pmatrix} 0.75 \\ 0.25 \end{pmatrix}
\end{align*}
\]

- From initial observation of rain

\[
\begin{align*}
\mathbb{P}(X_1) &= \begin{pmatrix} 0.0 \\ 1.0 \end{pmatrix} & \mathbb{P}(X_2) &= \begin{pmatrix} 0.3 \\ 0.7 \end{pmatrix} & \mathbb{P}(X_3) &= \begin{pmatrix} 0.48 \\ 0.52 \end{pmatrix} & \mathbb{P}(X_4) &= \begin{pmatrix} 0.588 \\ 0.412 \end{pmatrix} & \mathbb{P}(X_\infty) &= \begin{pmatrix} 0.75 \\ 0.25 \end{pmatrix}
\end{align*}
\]

- From yet another initial distribution \(\mathbb{P}(X_1):\)

\[
\begin{align*}
\mathbb{P}(X_1) &= \begin{pmatrix} p \\ 1 - p \end{pmatrix} & \ldots & \mathbb{P}(X_\infty) &= \begin{pmatrix} 0.75 \\ 0.25 \end{pmatrix}
\end{align*}
\]

Slide credits: http://ai.berkeley.edu
Stationary Distributions

• For most chains:
  – Influence of the initial distribution gets less and less over time.
  – The distribution we end up in is independent of the initial distribution

• Stationary distribution:
  – The distribution we end up with is called the stationary distribution $P_{\infty}$ of the chain
  – It satisfies

$$P_{\infty}(X) = P_{\infty+1}(X) = \sum_{x} P(X|x)P_{\infty}(x)$$
Example: Stationary Distributions

• Question: What’s $P(X)$ at time $t = \infty$?

\[
P_\infty(\text{sun}) = P(\text{sun}|\text{sun})P_\infty(\text{sun}) + P(\text{sun}|\text{rain})P_\infty(\text{rain})
\]
\[
P_\infty(\text{rain}) = P(\text{rain}|\text{sun})P_\infty(\text{sun}) + P(\text{rain}|\text{rain})P_\infty(\text{rain})
\]

\[
P_\infty(\text{sun}) = 0.9P_\infty(\text{sun}) + 0.3P_\infty(\text{rain})
\]
\[
P_\infty(\text{rain}) = 0.1P_\infty(\text{sun}) + 0.7P_\infty(\text{rain})
\]

\[
P_\infty(\text{sun}) = 3P_\infty(\text{rain})
\]
\[
P_\infty(\text{rain}) = 1/3P_\infty(\text{sun})
\]

Also: $P_\infty(\text{sun}) + P_\infty(\text{rain}) = 1$  \hspace{2cm} $P_\infty(\text{sun}) = 3/4$

| $X_{t-1}$ | $X_t$ | $P(X_t|X_{t-1})$ |
|-----------|-------|------------------|
| sun       | sun   | 0.9              |
| sun       | rain  | 0.1              |
| rain      | sun   | 0.3              |
| rain      | rain  | 0.7              |
Hidden Markov Models

- Underlying Markov chain over states $X$
- You observe evidence $E$ at each time step
- $X_t$ is a single discrete variable; $E_t$ may be continuous and may consist of several variables
Example: Weather HMM

• An HMM is defined by:
  – Initial state model: $P(X_0)$
  – Transition model: $P(X_t | X_{t-1})$
  – Sensor model: $P(E_t | X_t)$

| $R_{t-1}$ | $R_t$ | $P(R_t | R_{t-1})$ |
|-----------|-------|-------------------|
| +r        | +r    | 0.7               |
| +r        | -r    | 0.3               |
| -r        | +r    | 0.3               |
| -r        | -r    | 0.7               |

| $R_t$ | $U_t$ | $P(U_t | R_t)$ |
|-------|-------|---------------|
| +r    | +u    | 0.9           |
| +r    | -u    | 0.1           |
| -r    | +u    | 0.2           |
| -r    | -u    | 0.8           |
HMM as Probability Model

- Joint distribution for Markov model: \( P(X_{0:t}) = P(X_0) \prod_{i=1:t} P(X_i \mid X_{i-1}) \)
- Joint distribution for hidden Markov model: 
  \( P(X_{0:t}, E_{1:t}) = P(X_0) \prod_{i=1:t} P(X_i \mid X_{i-1}) P(E_i \mid X_i) \)
- HMMs have two important independence properties:
  - Markov hidden process: future depends on past via the present
  - Current observation independent of all else given current state
Inference in Temporal Models

• Filtering: $P(X_t|e_{1:t})$
  – belief state: input to the decision process of a rational agent

• Prediction: $P(X_{t+k}|e_{1:t})$ for $k > 0$
  – evaluation of possible action sequences; like filtering without the evidence

• Smoothing: $P(X_k|e_{1:t})$ for $0 \leq k < t$
  – better estimate of past states, essential for learning

• Most likely explanation: $\arg \max_{x_{1:t}} P(x_{1:t} | e_{1:t})$
  – speech recognition, decoding with a noisy channel
Filtering / State Estimation

- Filtering is the task of computing the belief state $P(X_t|e_{1:t})$
  - The posterior distribution over the most recent state given all evidence to date
- $B(X_t) = P(X_t|e_{1:t})$: belief distribution at time $t$ with all evidence $e_1, \ldots, e_t$ observed up to date
- $B'(X_{t+1}) = P(X_{t+1}|e_{1:t+1})$: belief distribution at time $t+1$ with all evidence $e_1, \ldots, e_{t+1}$ observed up to date
- $e_{1:t} = e_1, \ldots, e_t$
Inference: Find State Given Evidence

• We are given evidence at each time and want to know $B'(X_{t+1}) = P(X_{t+1}|e_{1:t+1})$

$B(X_t) = P(X_t|e_{1:t})$  \hspace{1cm} $B'(X_{t+1})$

![Diagram of states and evidence]
Inference: Base Cases

The current state distribution is projected forward from $t$ to $t+1$

$$P(X_2) = \sum_{x_1} P(x_1, X_2)$$

$$P(X_2) = \sum_{x_1} P(X_2|x_1)P(x_1)$$

The current state distribution is updated using the new evidence $e_{t+1}$

$$P(X_1|e_1) = \frac{P(X_1, e_1)}{\sum_{x_1} P(x_1, e_1)}$$

$$P(X_1|e_1) = \frac{P(e_1|X_1)P(X_1)}{\sum_{x_1} P(e_1|x_1)P(x_1)}$$

The current state distribution is updated using the new evidence $e_{t+1}$
Passage of Time

- Assume we have current belief $P(X | \text{evidence to date})$
  $B(X_t) = P(X_t|e_{1:t})$

- After one time step passes:

  $P(X_{t+1}|e_{1:t}) = \sum_{x_t} P(X_{t+1}, x_t|e_{1:t})$ \hspace{1cm} \text{marginalization}$

  $= \sum_{x_t} P(X_{t+1}|x_t, e_{1:t}) P(x_t|e_{1:t})$ \hspace{1cm} \text{condition on } x_t$

  $= \sum_{x_t} P(X_{t+1}|x_t) P(x_t|e_{1:t})$ \hspace{1cm} \text{Apply conditional independence}$

  $B'(X_{t+1}) = \sum_{x_t} P(X'|x_t) B(x_t)$
Observation

• Assume we have current belief $P(X \mid \text{previous evidence})$:

$$B'(X_{t+1}) = P(X_{t+1} \mid e_{1:t})$$

• Then, after evidence comes in:

$$P(X_{t+1} \mid e_{1:t+1}) = P(X_{t+1} \mid e_{1:t}, e_{t+1})$$

diving up the evidence

$$= \alpha P(e_{t+1} \mid X_{t+1}, e_{1:t}) P(X_{t+1} \mid e_{1:t})$$

Bayes’ rule, given $e_{1:t}$

$$= \alpha P(e_{t+1} \mid X_{t+1}) P(X_{t+1} \mid e_{1:t})$$

Markov assumption

• $B(X_{t+1}) = \alpha P(e_{t+1} \mid X_{t+1}) B'(X_{t+1})$
Filtering algorithm

- $P(X_{t+1}|e_{1:t+1}) = \alpha P(e_{t+1}|X_{t+1}) P(X_{t+1}|e_{1:t})$
  
  $= \alpha P(e_{t+1}|X_{t+1}) \sum_{x_t} P(X_{t+1}|x_t, e_{1:t}) P(x_t|e_{1:t})$
  
  $= \alpha P(e_{t+1}|X_{t+1}) \sum_{x_t} P(X_{t+1}|x_t) P(x_t|e_{1:t})$

- $P(X_{t+1}|e_{1:t+1}) = f(e_{t+1}, P(x_t|e_{1:t}))$ recursive estimation
- $f_{1:t+1} = \text{Forward}(f_{1:t}, e_{t+1})$
- Prediction step: as time passes, uncertainty “accumulates”
- Update step: as we get observations, beliefs get reweighted, uncertainty “decreases”
Example: Weather HMM

Predict from t=0 to t=1

\[ P(R_1) = \sum_{r_0} P(R_1 | r_0) P(r_0) \]

\[ = \langle 0.7, 0.3 \rangle \times 0.5 + \langle 0.3, 0.7 \rangle \times 0.5 = \langle 0.5, 0.5 \rangle \]

\[ B(+r) = 0.5 \quad B'(+r) = 0.5 \]
\[ B(-r) = 0.5 \quad B'(-r) = 0.5 \]

| \( R_t \) | \( R_{t+1} \) | \( P(R_{t+1} | R_t) \) |
|---|---|---|
| +r | +r | 0.7 |
| +r | -r | 0.3 |
| -r | +r | 0.3 |
| -r | -r | 0.7 |

| \( R_t \) | \( U_t \) | \( P(U_t | R_t) \) |
|---|---|---|
| +r | +u | 0.9 |
| +r | -u | 0.1 |
| -r | +u | 0.2 |
| -r | -u | 0.8 |
Example: Weather HMM

\begin{align*}
B'(\text{+r}) &= 0.5 \\
B'(\text{-r}) &= 0.5 \\
B(\text{+r}) &= 0.5 \\
B(\text{-r}) &= 0.5 \\
B'(\text{+r}) &= 0.818 \\
B'(\text{-r}) &= 0.182
\end{align*}

$$B(\text{+r}) = 0.818$$
$$B(\text{-r}) = 0.182$$

Update

\begin{align*}
P(R_1 | u_1) &= \alpha P(u_1 | R_1)P(R_1) = \alpha \langle 0.9, 0.2 \rangle \langle 0.5, 0.5 \rangle \\
&= \alpha \langle 0.45, 0.1 \rangle \approx \langle 0.818, 0.182 \rangle.
\end{align*}

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|}
\hline
R_t & U_t & P(U_t | R_t) \\
\hline
+\text{r} & +\text{u} & 0.9 \\
+\text{r} & -\text{u} & 0.1 \\
-\text{r} & +\text{u} & 0.2 \\
-\text{r} & -\text{u} & 0.8 \\
\hline
\end{tabular}
\end{table}
Example: Weather HMM

\[ B'(+r) = 0.5 \]
\[ B'(-r) = 0.5 \]
\[ B(+r) = 0.5 \]
\[ B(-r) = 0.5 \]

\[ B(+r) = 0.818 \]
\[ B(-r) = 0.182 \]

\[ B'(+r) = 0.627 \]
\[ B'(-r) = 0.373 \]

Predict from \( t=1 \) to \( t=2 \)

\[
P(R_2 | u_1) = \sum_{r_1} P(R_2 | r_1)P(r_1 | u_1)
\]
\[
= \langle 0.7, 0.3 \rangle \times 0.818 + \langle 0.3, 0.7 \rangle \times 0.182 \approx \langle 0.627, 0.373 \rangle
\]

| \( R_t \) | \( R_{t+1} \) | \( P(R_{t+1} | R_t) \) |
|---|---|---|
| +r | +r | 0.7 |
| +r | -r | 0.3 |
| -r | +r | 0.3 |
| -r | -r | 0.7 |
Example: Weather HMM

\[ P(R_2 | u_1, u_2) = \alpha P(u_2 | R_2)P(R_2 | u_1) = \alpha \langle 0.9, 0.2 \rangle \langle 0.627, 0.373 \rangle = \alpha \langle 0.565, 0.075 \rangle \approx \langle 0.883, 0.117 \rangle. \]

Update

| \( R_t \) | \( U_t \) | \( P(U_t | R_t) \) |
|---|---|---|
| +r | +u | 0.9 |
| +r | -u | 0.1 |
| -r | +u | 0.2 |
| -r | -u | 0.8 |
Other HMM Queries

Filtering: \( P(X_t | e_{1:t}) \)

Prediction: \( P(X_{t+k} | e_{1:t}) \)

Smoothing: \( P(X_k | e_{1:t}), k<t \)

Explanation: \( P(X_{1:t} | e_{1:t}) \)
Real HMM Examples

• Speech recognition HMMs:
  – Observations are acoustic signals (continuous valued)
  – States are specific positions in specific words (so, tens of thousands)

• Machine translation HMMs:
  – Observations are words (tens of thousands)
  – States are translation options

• Robot tracking:
  – Observations are range readings (continuous)
  – States are positions on a map (continuous)

• Molecular biology:
  – Observations are nucleotides ACGT
  – States are coding/non-coding/start/stop/splice-site etc.
Reading and Next Class

- Decision Networks: AIMA 16.5-16.6
- HMMs: AIMA 14.1-14.3
- Next: Particle Filtering 14.4-14.6