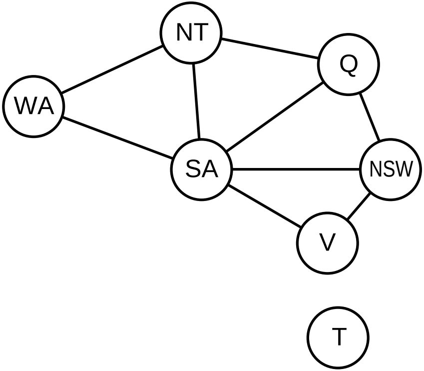
Nodes from a goal set

$$G_n \approx \{\text{nodes with distance } n \text{ from } G\}$$



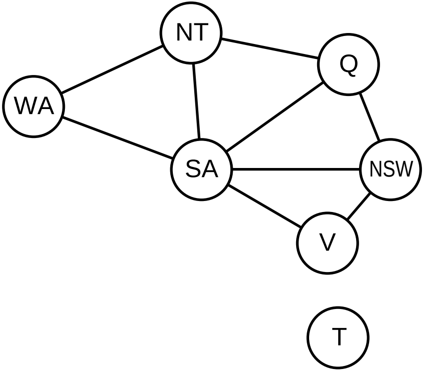
$G_0 := G$

$\{\mathrm{V}\}_0 = \{\mathrm{V}\}$

Nodes from a goal set

$$G_n \approx \{\text{nodes with distance } n \text{ from } G\}$$

$$G_{n+1} := \{s \mid (\exists s' \in G_n)\ arc(s, s')\} - \bigcup_{i=1}^{n} G_i$$



$$G_0 := G$$

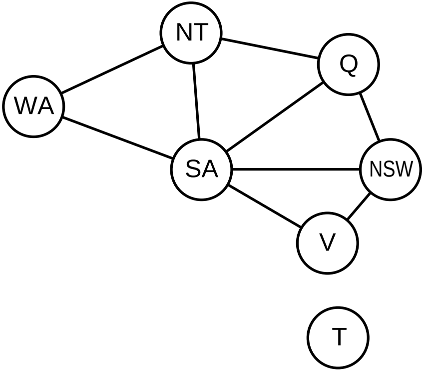
$$\{\mathrm{V}\}_0 = \{\mathrm{V}\}$$

$$\{\mathrm{V}\}_1 = \{\mathrm{SA}, \mathrm{NSW}\}$$

Nodes from a goal set

$$G_n \approx \{\text{nodes with distance } n \text{ from } G\}$$

$$G_{n+1} := \{s \mid (\exists s' \in G_n)\ arc(s, s')\} - \bigcup_{i=1}^{n} G_i$$



$G_0 := G$

$\{\mathrm{V}\}_0 = \{\mathrm{V}\}$

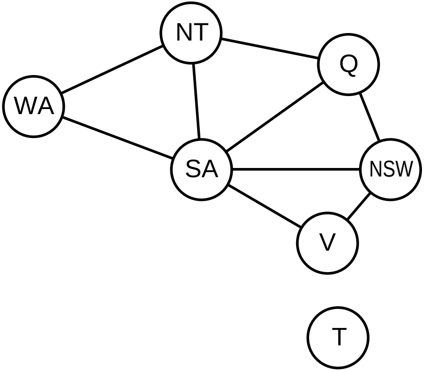
$\{\mathrm{V}\}_1 = \{\mathrm{SA}, \mathrm{NSW}\}$

$\{\mathrm{V}\}_2 = \{\mathrm{WA}, \mathrm{NT}, \mathrm{Q}\}$

Nodes from a goal set

$$G_n \approx \{\text{nodes with distance } n \text{ from } G\}$$

$$G_{n+1} := \{s \mid (\exists s' \in G_n)\ arc(s, s')\} - \bigcup_{i=1}^{n} G_i$$



$$G_0 := G$$

$$\{\mathsf{V}\}_0 = \{\mathsf{V}\}$$

$$\{\mathsf{V}\}_1 = \{\mathsf{SA}, \mathsf{NSW}\}$$

$$\{\mathsf{V}\}_2 = \{\mathsf{WA}, \mathsf{NT}, \mathsf{Q}\}$$

$$\{\mathsf{V}\}_\infty = \{\mathsf{T}\}$$

Distance d_G to minimize

$$d_G(s) \;:=\; \begin{cases} n & \text{if } s \in G_n \\ \infty & \text{otherwise} \end{cases}$$

Refine

$$\delta_G(s) \;:=\; \begin{cases} 1 & \text{if } s \in G \\ 0 & \text{otherwise} \end{cases}$$

to reward from 1 to 0 ($\approx$ distance from 0 to ∞)

Distance d_G to minimize $\rightsquigarrow$ reward r_G to maximize

$$d_G(s) \;:=\; \begin{cases} n & \text{if } s \in G_n \\ \infty & \text{otherwise} \end{cases}$$

Refine

$$\delta_G(s) \;:=\; \begin{cases} 1 & \text{if } s \in G \\ 0 & \text{otherwise} \end{cases}$$

to reward from 1 to 0 ($\approx$ distance from 0 to ∞)

$$r_G(s) \;:=\; \begin{cases} 2^{-n} & \text{if } s \in G_n \\ 0 & \text{otherwise} \end{cases}$$

halving the reward as we step back (starting at G)

Distance d_G to minimize $\rightsquigarrow$ reward r_G to maximize

$$d_G(s) \;:=\; \begin{cases} n & \text{if } s \in G_n \\ \infty & \text{otherwise} \end{cases}$$

Refine

$$\delta_G(s) \;:=\; \begin{cases} 1 & \text{if } s \in G \\ 0 & \text{otherwise} \end{cases}$$

to reward from 1 to 0 ($\approx$ distance from 0 to ∞)

$$r_G(s) \;:=\; \begin{cases} 2^{-n} & \text{if } s \in G_n \\ 0 & \text{otherwise} \end{cases}$$

halving the reward as we step back (starting at G)

$$r_G(s) \;=\; \frac{1}{2} r_G(s') \text{ if } arc(s, s') \text{ and } d_G(s') < d_G(s).$$

Rewards looking ahead

$$
\begin{aligned}
H_0(s) &:= \delta_G(s) \\
H_{n+1}(s) &:= \delta_G(s) + \frac{1}{2} \max\{H_n(s') \mid arc_=(s, s')\}
\end{aligned}
$$

where

$$arc_=(s, s') \iff arc(s, s') \text{ or } s = s'.$$

Rewards looking ahead

$$
\begin{aligned}
H_0(s) &:= \delta_G(s) \\
H_{n+1}(s) &:= \delta_G(s) + \frac{1}{2}\max\{H_n(s') \mid arc_=(s, s')\}
\end{aligned}
$$

where

$$arc_=(s, s') \iff arc(s, s') \text{ or } s = s'.$$

For $s \in G_0$,

$$H_{n+1}(s) = 1 + \frac{1}{2}H_n(s) = a_{n+1}$$

where

$$a_n := \sum_{k=0}^{n} 2^{-k} = 2(1 - 2^{-(n+1)})$$

Rewards looking ahead

$$H_0(s) := \delta_G(s)$$

$$H_{n+1}(s) := \delta_G(s) + \frac{1}{2}\max\{H_n(s') \mid arc_=(s,s')\}$$

where

$$arc_=(s,s') \iff arc(s,s') \text{ or } s = s'.$$

For $s \in G_0$,

$$H_{n+1}(s) = 1 + \frac{1}{2}H_n(s) = a_{n+1}$$

where

$$a_n := \sum_{k=0}^{n} 2^{-k} = 2(1 - 2^{-(n+1)})$$

and for $s \in G_{k+1}$, $s' \in G_k$ and $arc(s,s')$,

$$H_{k+n+1}(s) = \frac{1}{2}H_{k+n}(s')$$

Rewards looking ahead: $\lim_{n\to\infty} H_n(s) = 2r_G(s)$

$$\begin{aligned} H_0(s) &:= \delta_G(s) \\ H_{n+1}(s) &:= \delta_G(s) + \frac{1}{2}\max\{H_n(s') \mid arc_=(s,s')\} \end{aligned}$$

where

$$arc_=(s,s') \iff arc(s,s') \text{ or } s = s'.$$

For $s \in G_0$,

$$H_{n+1}(s) \;=\; 1 + \frac{1}{2}H_n(s) \;=\; a_{n+1}$$

where

$$a_n \;:=\; \sum_{k=0}^{n} 2^{-k} \;=\; 2(1 - 2^{-(n+1)})$$

and for $s \in G_{k+1}$, $s' \in G_k$ and $arc(s,s')$,

$$H_{k+n+1}(s) \;=\; \frac{1}{2}H_{k+n}(s')$$

$H = \lim_{n\to\infty} H_n$

$$H(s) \; = \delta_G(s) + \frac{1}{2}\max\{H(s') \mid arc_=(s,s')\}$$

a foolproof heuristic for the shortest solution

`Frontier = [Hd|Tl]` with $H(\texttt{Hd}) \; \geq H(s)$ for all s in `Tl`.

$H = \lim_{n \to \infty} H_n$

$$H(s) \ = \delta_G(s) + \frac{1}{2} \max\{H(s') \mid arc_=(s, s')\}$$

a foolproof heuristic for the shortest solution

`Frontier = [Hd|Tl]` with $H(\texttt{Hd}) \ \geq H(s)$ for all s in `Tl`.

What if arcs have different costs?

$H = \lim_{n\to\infty} H_n$

$$H(s) \ = \delta_G(s) + \frac{1}{2}\max\{H(s') \mid arc_=(s,s')\}$$

a foolproof heuristic for the shortest solution

`Frontier = [Hd|Tl]` with $H(\texttt{Hd}) \ \geq H(s)$ for all s in `Tl`.

What if arcs have different costs?

Modify $\delta_G(s)$ to

$$Q_0(s,s') \ := \ \begin{cases} 1 & \text{if } s = s' \in G \\ -\text{cost}(s,s') & \text{else if } arc(s,s') \\ -\max_{s_1,s_2}\text{cost}(s_1,s_2) & \text{otherwise} \end{cases}$$

and $H_{n+1}(s)$ to

$$Q_{n+1}(s,s') \ := \ Q_0(s,s') + \frac{1}{2}\max\{Q_n(s',s'') \mid arc_=(s',s'')\}$$

Discounted rewards ($0 \le \gamma < 1$)

(immediate) rewards $r_1, r_2, r_3, \ldots$ at times $1, 2, 3, \ldots$ give a γ-discounted value of

$$V \; := r_1 + \gamma r_2 + \gamma^2 r_3 + \cdots \; = \; \sum_{i \ge 1} \gamma^{i-1} r_i$$

Discounted rewards ($0 \leq \gamma < 1$)

(immediate) rewards $r_1, r_2, r_3, \ldots$ at times $1, 2, 3, \ldots$ give a γ-discounted value of

$$\begin{aligned} V \;&:=\; r_1 + \gamma r_2 + \gamma^2 r_3 + \cdots \;=\; \sum_{i \geq 1} \gamma^{i-1} r_i \\ &=\; \Big(\sum_{i=1}^{t} \gamma^{i-1} r_i\Big) \;+\; \gamma^t V_{t+1} \end{aligned}$$

where V_t is the value from time step t on ($V_1 = V$)

$$V_t \;:=\; \sum_{i \geq t} \gamma^{i-t} r_i \;=\; r_t + \gamma V_{t+1}$$

Discounted rewards ($0 \leq \gamma < 1$)

(immediate) rewards $r_1, r_2, r_3, \ldots$ at times $1, 2, 3, \ldots$ give a γ-discounted value of

$$V \;:=\; r_1 + \gamma r_2 + \gamma^2 r_3 + \cdots \;=\; \sum_{i \geq 1} \gamma^{i-1} r_i$$

$$= \; \Big(\sum_{i=1}^{t} \gamma^{i-1} r_i \Big) \; + \; \gamma^t V_{t+1}$$

where V_t is the value from time step t on ($V_1 = V$)

$$V_t \;:=\; \sum_{i \geq t} \gamma^{i-t} r_i \;=\; r_t + \gamma V_{t+1}$$

which is bound by bounds on r_i

$$m \leq r_i \leq M \;\text{ for each } i \geq t \quad \text{implies} \quad \frac{m}{1-\gamma} \leq V_t \leq \frac{M}{1-\gamma}$$

since $\sum_{i \geq 0} \gamma^i = (1-\gamma)^{-1}$

Discounted rewards ($0 \leq \gamma < 1$)

(immediate) rewards $r_1, r_2, r_3, \ldots$ at times $1, 2, 3, \ldots$ give a γ-discounted value of

$$
\begin{aligned}
V \;&:= r_1 + \gamma r_2 + \gamma^2 r_3 + \cdots \;=\; \sum_{i \geq 1} \gamma^{i-1} r_i \\
&= \;\Big(\sum_{i=1}^{t} \gamma^{i-1} r_i\Big) \;+\; \gamma^t V_{t+1}
\end{aligned}
$$

where V_t is the value from time step t on ($V_1 = V$)

$$V_t \;:=\; \sum_{i \geq t} \gamma^{i-t} r_i \;=\; r_t + \gamma V_{t+1} \quad \text{(backward induction)}$$

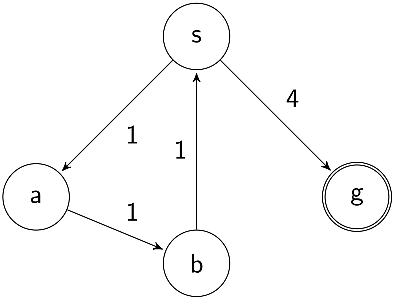
which is bound by bounds on r_i

$$m \leq r_i \leq M \;\text{ for each } i \geq t \quad \text{implies} \quad \frac{m}{1-\gamma} \leq V_t \leq \frac{M}{1-\gamma}$$

since $\sum_{i \geq 0} \gamma^i = (1-\gamma)^{-1}$

$Q = \lim_{n\to\infty} Q_n$

$$
\begin{aligned}
Q(s,s') &:= Q_0(s,s') + \frac{1}{2}\max\{Q(s',s'') \mid arc_=(s',s'')\} \\
Q_0(s,s') &:= \begin{cases} 1 & \text{if } s = s' \in G \\ -\mathrm{cost}(s,s') & \text{else if } arc(s,s') \end{cases}
\end{aligned}
$$



soln not chosen

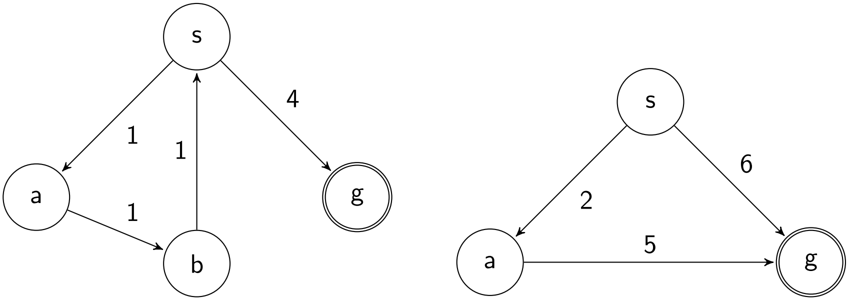
Q(s,g) = -3

Q(s,a) = -2 = Q(a,b) = Q(b,s)

$Q = \lim_{n\to\infty} Q_n$

$$Q(s,s') := Q_0(s,s') + \frac{1}{2}\max\{Q(s',s'') \mid arc_=(s',s'')\}$$

$$Q_0(s,s') := \begin{cases} 1 & \text{if } s = s' \in G \\ -\text{cost}(s,s') & \text{else if } arc(s,s') \end{cases}$$



soln not chosen	costlier soln chosen
Q(s,g) = -3	Q(s,g) = -5
Q(s,a) = -2 = Q(a,b) = Q(b,s)	Q(a,g) = -4 = Q(s,a)

Upping the reward

Adjust $Q_0(s, s')$ to

$$R(s, s') \;:=\; \begin{cases} r & \text{if } s = s' \in G \\ -\text{cost}(s, s') & \text{else if } arc(s, s') \end{cases}$$

for some reward r high enough to offset costs of reaching a goal

Upping the reward

Adjust $Q_0(s, s')$ to

$$R(s, s') \ := \ \begin{cases} r & \text{if } s = s' \in G \\ -\text{cost}(s, s') & \text{else if } arc(s, s') \end{cases}$$

for some reward r high enough to offset costs of reaching a goal

$$\text{e.g. } r \ \geq \ 2^n nc$$

for solutions up to $n - 1$ arcs with arcs costing $\leq c$

e.g. n states, and c is max arc cost.

Upping the reward

Adjust $Q_0(s, s')$ to

$$R(s, s') \;:=\; \begin{cases} r & \text{if } s = s' \in G \\ -\text{cost}(s, s') & \text{else if } arc(s, s') \end{cases}$$

for some reward r high enough to offset costs of reaching a goal

$$\text{e.g. } r \;\geq\; 2^n nc$$

for solutions up to $n-1$ arcs with arcs costing $\leq c$

e.g. n states, and c is max arc cost.

Let

$$V(s) \;:=\; \max\{Q(s, s') \mid arc_=(s, s')\}$$

so for $0 \leq i < n$, $s' \in G_i$ and $arc(s, s')$,

$$V(s') \;\geq 2^{n-i}(n-i)c$$

$$V(s) \;\geq\; -c + \frac{1}{2}V(s') \;\geq\; 2^{n-(i+1)}(n-(i+1))c \;\geq\; 2c.$$

Recap

From node s, find path to goal via s' maximizing

$$Q(s, s') \;:=\; R(s, s') + \frac{1}{2} V(s')$$

with discount $\frac{1}{2}$ on future $V(s')$, contra

$$\begin{aligned} \mathrm{cost}(s_1 \cdots s_k) \;&=\; \sum_{i=1}^{k-1} \mathrm{cost}(s_i, s_{i+1}) \\ &\neq\; \sum_{i=1}^{k-1} 2^{1-i} \mathrm{cost}(s_i, s_{i+1}) \end{aligned}$$

Recap

From node s, find path to goal via s' maximizing

$$Q(s,s') \;:=\; R(s,s') + \frac{1}{2}V(s')$$

with discount $\frac{1}{2}$ on future $V(s')$, contra

$$\begin{aligned}\mathrm{cost}(s_1 \cdots s_k) \;&=\; \sum_{i=1}^{k-1} \mathrm{cost}(s_i, s_{i+1}) \\ &\neq\; \sum_{i=1}^{k-1} 2^{1-i}\mathrm{cost}(s_i, s_{i+1})\end{aligned}$$

Trade min cost guarantee for cost-benefit analysis with chance $\frac{1}{2}$ of survival/doom.

Recap

From node s, find path to goal via s' maximizing

$$Q(s,s') \ := R(s,s') + \frac{1}{2} V(s')$$

with discount $\frac{1}{2}$ on future $V(s')$, contra

$$\begin{aligned} \mathrm{cost}(s_1 \cdots s_k) \ &= \ \sum_{i=1}^{k-1} \mathrm{cost}(s_i, s_{i+1}) \\ &\neq \ \sum_{i=1}^{k-1} 2^{1-i} \mathrm{cost}(s_i, s_{i+1}) \end{aligned}$$

Trade min cost guarantee for cost-benefit analysis with chance $\frac{1}{2}$ of survival/doom.

NEXT: more uncertainty, approached via approximations like

$$\begin{aligned} Q_n(s,s') \ &\approx \ Q(s,s') \text{ up to look ahead } n \\ Q(s,s') \ &= \ \lim_{n\to\infty} Q_n(s,s'). \end{aligned}$$