Introduction: The best course of action to reach $p$?

Action $c$ is better than $b$. 
Introduction: The best course of action to reach $p$?

action $d$ is better than $c$. 
Introduction: The best course of action to reach \( p \)?

Action \( a \) is better than \( d \).
An example

“Try your best to reach $p$”
An example

Policy $\pi_1$
Policy $\pi_2$ clearly worse than $\pi_1$!
An example

Policy $\pi_3$

worse than $\pi_1$! but $\pi_2$?
An example

Policy $\pi_4$

worse than $\pi_2$ and $\pi_3$
An example

Policy $\pi_5$

worse than $\pi_3$
An example

Policy $\pi_6$

worse than $\pi_4, \pi_5$
An example

Policy $\pi_7$
Really bad!
An example

Which is the best Policy?
How do we express "best policy"?
Existing Logic

- LTL: The property of a sequence of states besides the final state (if it exists);
Existing Logic

- **LTL**: The property of a sequence of states besides the final state (if it exists);

- **CTL**: LTL + properties of all paths from each state;
Existing Logic

- **LTL**: The property of a sequence of states besides the final state (if it exists);

- **CTL***: LTL + properties of all paths from each state;

- **π-CTL***: CTL* + properties of all paths in the policy from a state.
Linear Temporal Logic LTL

- Linear time: sequence of states
- Operators:
  \( \square p = \text{always } p \)
  \( \Diamond p = \text{eventually } p \)
  \( \bigcirc p = \text{next } p \)
  \( p \cup q = p \text{ true until } q \)
Branching Temporal logic CTL

- Branching time
- New operators for paths
Branching Temporal logic CTL* 

- Branching time
- New operators for paths
Branching Temporal logic CTL*

- Branching time
- New operators for paths
Branching Temporal logic CTL*

- Branching time
- New operators for paths
Branching Temporal logic CTL*

- Branching time
- New operators for paths
Branching Temporal logic CTL*

\[ A\phi = \text{for any path, } \phi \text{ holds} \]
\[ E\phi = \text{for some path, } \phi \text{ holds} \]
Branching Temporal logic CTL*

Examples:
$A \Diamond p = \text{all paths reach } p$
$E \Box p = \text{in some path, always } p$
Branching Temporal logic CTL* 

Syntax:

\[ \langle p \rangle = \text{propositional formula}; \]
\[ \langle sf \rangle = \text{“state” formula}; \]
\[ \langle pf \rangle = \text{“path” formula} \]
**Branching Temporal logic CTL**

Syntax:

\[ \langle p \rangle = \text{propositional formula}; \]

\[ \langle sf \rangle = \text{“state” formula}; \]

\[ \langle pf \rangle = \text{“path” formula} \]

\[ \langle sf \rangle ::= \langle p \rangle | \langle sf \rangle \land \langle sf \rangle | \langle sf \rangle \lor \langle sf \rangle | \neg \langle sf \rangle | E\langle pf \rangle | A\langle pf \rangle \]
Branching Temporal logic CTL*

Syntax:
\(\langle p \rangle =\) propositional formula;
\(\langle sf \rangle =\) “state” formula;
\(\langle pf \rangle =\) “path” formula

\[
\langle sf \rangle ::= \langle p \rangle | \langle sf \rangle \land \langle sf \rangle | \langle sf \rangle \lor \langle sf \rangle | \neg \langle sf \rangle | E \langle pf \rangle | A \langle pf \rangle
\]

\[
\langle pf \rangle ::= \langle sf \rangle | \langle pf \rangle \lor \langle pf \rangle | \neg \langle pf \rangle | \langle pf \rangle \land \langle pf \rangle | \langle pf \rangle \lor \langle pf \rangle | \langle pf \rangle \land \langle pf \rangle
\]
The extension of CTL*: $\pi$-CTL*

Syntax:

\[
\langle sf \rangle ::= \langle p \rangle | \langle sf \rangle \land \langle sf \rangle | \langle sf \rangle \lor \langle sf \rangle | \neg \langle sf \rangle | \\
E\langle pf \rangle | A\langle pf \rangle | A_\pi\langle pf \rangle | E_\pi\langle pf \rangle
\]

\[
\langle pf \rangle ::= \langle sf \rangle | \langle pf \rangle \lor \langle pf \rangle | \neg \langle pf \rangle | \langle pf \rangle \land \langle pf \rangle | \\
\langle pf \rangle \cup \langle pf \rangle | \Box \langle pf \rangle | \Diamond \langle pf \rangle | \square \langle pf \rangle
\]
The extension of CTL*: $\pi$-CTL*

- group the set of paths from the initial state that all correspond to the same policy:
The extension of CTL*: $\pi$-CTL*

- group the set of paths from the initial state that all correspond to the same policy:
  - $A_\pi pf$: ‘for all paths that agree with the policy $\pi$, $pf$ holds’;
  - $E_\pi pf$: ‘there exists a path that agrees with the policy $\pi$ for which $pf$ holds’.

- By policy, we mean the mapping from states to actions.

- We now illustrate some goals in $\pi$-CTL*
**Weak Plan**

The weakest reachability goal “from the initial state there is a possibility that \( p \) can be reached” is expressed by \( E_\pi \Diamond p \).

\((s_1, c)\) is a weak plan
strong plan

A stronger goal “from the initial state $p$ must be reached” is expressed as $A_\pi \Diamond p$.

$(s_1, a)$ is a strong plan
"All along the trajectory there is always a possible path to \( p \) by following the policy" is expressed as \( A_\pi \square (E_\pi \Diamond p) \).

\((s_1, d)\) is a strong cyclic plan.
Some goals in $\pi$-CTL$^*$

weak plan: The weakest reachability goal “from the initial state there is a possibility that $p$ can be reached” is expressed by $E_{\pi}\Diamond p$. 
Some goals in $\pi$-CTL*

weak plan: The weakest reachability goal “from the initial state there is a possibility that $p$ can be reached” is expressed by $E_{\pi} \Diamond p$. From $s_1$, all policies but $\pi_7$ satisfy the goal.
strong plan: A stronger goal “from the initial state $p$ must be reached” is expressed as $A_{\pi} \diamond p$. For $s_1$, no policy makes it true.
strong plan: A stronger goal “from the initial state $p$ must be reached” is expressed as $A_{\pi} \Diamond p$. For $s_1$, no policy makes it true. But, for instance, for $s_2$ the policy $\{(s_2, a_2)\}$ satisfies the goal.
“All along the trajectory there is always a possible path to $p$ by following the policy” is expressed as $A_{\pi} \Box (E_{\pi} \Diamond p)$. For $s_1$, no policy.
“All along the trajectory there is always a possible path to $p$ by following the policy” is expressed as $A_\pi \Box (E_\pi \Diamond p)$. For $s_1$, no policy. For $s_2$, policies $\{(s_2, a_2)\}$ and $\{(s_2, a_7)\}$ satisfy this goal.
However, policy \( \{(s_2, a_5)\} \) does not, (we could go to \( s_5 \) from where \( p \) can not be reached).
More examples

- $A_\pi(E \diamond p) = \text{“All along the trajectory there is always a possible path to } p\text{, but this path is not necessary abide the policy the agent taken”}$. 
More examples

- $A_\pi(E \diamond p) = \text{“All along the trajectory there is always a possible path to } p, \text{ but this path is not necessary abide the policy the agent taken”}. $

- $A(E_\pi \Diamond p) = \text{“For any state that is reachable from the initial state, there is always a path to } p \text{ by following the policy.”} $
More examples

- \( A_{\pi}(E\diamond p) = \) “All along the trajectory there is always a possible path to \( p \), but this path is not necessary abide the policy the agent taken”.

- \( A(E_{\pi}\diamond p) = \) “For any state that is reachable from the initial state, there is always a path to \( p \) by following the policy.”

- \( E\diamond p \rightarrow E_{\pi}\diamond p = \) “from the initial state, if it is possible to reach \( p \), the agent should possibly reach \( p \)”.
  Useful to allow the agent to pursue an alternative goal when it realizes that its initial goal is no longer achievable.
More examples

- $A_\pi(E \diamond p) = \text{“All along the trajectory there is always a possible path to } p, \text{ but this path is not necessary abide the policy the agent taken”}.$

- $A(E_\pi \diamond p) = \text{“For any state that is reachable from the initial state, there is always a path to } p \text{ by following the policy.”}$

- $E \diamond p \rightarrow E_\pi \diamond p = \text{“from the initial state, if it is possible to reach } p, \text{ the agent should possibly reach } p”. \text{ Useful to allow the agent to pursue an alternative goal when it realizes that its initial goal is no longer achievable.}$

- $A_\pi \Box(E \diamond p \rightarrow E_\pi \diamond p) = \text{idem, but now from any state in the trajectory (not only initial one).}$
To find the best policy, the comparison of policies is necessary. For example:

All along your trajectory
if from any state $p$ can be achieved for sure,
then the policy being executed must achieve $p$,
else ......
To find the best policy, the comparison of policies is necessary. For example:

All along your trajectory
if from any state $p$ can be achieved for sure,
then the policy being executed must achieve $p$,
else ......

- $\mathcal{AP}$: ‘for all policies from the state, the property is hold’;
- $\mathcal{EP}$: ‘there exist a policy from the state such that the property is hold in the policy’.
P-CTL* 

Syntax:

\[
\begin{align*}
\langle sf \rangle & ::= \langle p \rangle | \langle sf \rangle \land \langle sf \rangle | \langle sf \rangle \lor \langle sf \rangle | \neg \langle sf \rangle | \\
                     & \ E\langle pf \rangle | \ A\langle pf \rangle | \ A_\pi \langle pf \rangle | \ E_\pi \langle pf \rangle | \ AP\langle sf \rangle | \ EP\langle sf \rangle \\
\langle pf \rangle & ::= \langle sf \rangle | \langle pf \rangle \lor \langle pf \rangle | \neg \langle pf \rangle | \langle pf \rangle \land \langle pf \rangle | \\
                     & \langle pf \rangle \ U \langle pf \rangle | \ \Box \langle pf \rangle | \ \Diamond \langle pf \rangle | \ \square \langle pf \rangle
\end{align*}
\]
Goals in P-CTL*: Based on the weak plan “from the initial state, if there is a policy such that $p$ is possibly reached, then in the policy chosen by the agent, $p$ is possibly reached” is expressed by $(\mathcal{E}\mathcal{P}E_{\pi}\Diamond p) \rightarrow (E_{\pi}\Diamond p)$.
Based on the strong plan
“from the initial state, if there is a policy such that $p$ must be reached, then in the policy chosen by the agent, $p$ must be reached” is expressed by $(\mathcal{E} \mathcal{P} A_{\pi} \lozenge p) \rightarrow (A_{\pi} \lozenge p)$.
Based on the strong cyclic plan: “from the initial state, if there is a policy such that all along the trajectory there is always a possible path to $p$ by following the policy, then in any state of the chosen policy, there is always a possible path to $p$” is expressed as $\mathcal{E}\mathcal{P}(A_{\pi} \square (E_{\pi} \diamond \neg p)) \rightarrow (A_{\pi} \square (E_{\pi} \diamond p))$. 
one version of *Try your best to reach* \( p \)

“In any state, if there is a policy that is possibly reach \( p \), then the agent should possibly reach \( p \); if there is a policy that guarantees to reach \( p \), then the agent should guarantee to reach \( p \); if there is a policy such that in any state of the policy, there is a path to \( p \), then in the policy chosen by the agent, there is always a path to \( p \).”

It is expressed as

\[
A_\pi \Box ( (\mathcal{E} \mathcal{P} E_\pi \Diamond p) \rightarrow (E_\pi \Diamond p) ) \\
\wedge A_\pi \Box ( (\mathcal{E} \mathcal{P} A_\pi \Diamond p) \rightarrow (A_\pi \Diamond p) ) \\
\wedge A_\pi \Box ( \mathcal{E} \mathcal{P} (A_\pi \Box (E_\pi \Diamond p)) \rightarrow (A_\pi \Box (E_\pi \Diamond p)) )
\]
Policy $\pi_1$ in the previous example is the “Best” policy.
• In State $s_1$: There is a policy that has path to $p$, but no policy can guarantee to reach $p$

• In state $s_2$: There is an action ($a_5$) that has path to $p$, there is an action ($a_7$) that in any state of any path in the policy, there is always a hope of reaching $p$, and there is an action ($a_2$) that guarantees to reach $p$.

• In state $s_3$: There is an action ($a_3$) that has path to $p$

• In state $s_4$: $p$ is reached

• In state $s_5$: No policy has path to $p$, give up.
<table>
<thead>
<tr>
<th>Goal presentation</th>
<th>Satisfiable policies</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E_\pi \diamond p$</td>
<td>$\pi_1, \pi_2, \pi_3, \pi_4, \pi_5, \pi_6$</td>
</tr>
<tr>
<td>$A_\pi \Box (E\mathcal{P}A_\pi \Box (E_\pi \diamond p) \rightarrow A_\pi \Box (E_\pi \diamond p))$</td>
<td>$\pi_1, \pi_2, \pi_3, \pi_4, \pi_7$</td>
</tr>
<tr>
<td>$A_\pi \Box (E\mathcal{P}E_\pi \diamond p \rightarrow E_\pi \diamond p)$</td>
<td>$\pi_1, \pi_3, \pi_5$</td>
</tr>
<tr>
<td>$E_\pi \diamond p \land A_\pi \Box (E\mathcal{P}A_\pi \Box (E_\pi \diamond p) \rightarrow A_\pi \Box (E_\pi \diamond p))$</td>
<td>$\pi_1, \pi_2, \pi_3, \pi_4$</td>
</tr>
<tr>
<td>$A_\pi \Box (E\mathcal{P}A_\pi \Box \diamond p \rightarrow A_\pi \diamond p)$</td>
<td>$\pi_1, \pi_2, \pi_7$</td>
</tr>
<tr>
<td>$E_\pi \diamond p \land A_\pi \Box (E\mathcal{P}A_\pi \Box \diamond p \rightarrow A_\pi \diamond p)$</td>
<td>$\pi_1, \pi_2$</td>
</tr>
<tr>
<td>$A_\pi \Box ((E\mathcal{P}E_\pi \diamond p \rightarrow E_\pi \diamond p) \land (E\mathcal{P}A_\pi \Box (E_\pi \diamond p) \rightarrow A_\pi \Box (E_\pi \diamond p)))$</td>
<td>$\pi_1, \pi_3$</td>
</tr>
<tr>
<td>$A_\pi \Box ((E\mathcal{P}E_\pi \diamond p \rightarrow E_\pi \diamond p) \land (E\mathcal{P}A_\pi \Box \diamond p \rightarrow A_\pi \diamond p))$</td>
<td>$\pi_1$</td>
</tr>
<tr>
<td>$A_\pi \Diamond p$</td>
<td>$\emptyset$</td>
</tr>
</tbody>
</table>
More examples

• \( A_\pi \square ((\mathcal{E}\mathcal{P}A_\pi \diamond p) \cup q) \) = “reach \( q \) but want to make sure that all along the path if necessary it can make a new policy that can guarantee to reach \( q \)”.
More examples

- $A_\pi \Box ((EPA_\pi \Diamond p) U q) = "reach q but want to make sure that all along the path if necessary it can make a new policy that can guarantee to reach q".$

- $A_\pi \Box (APE_\pi \neg \Box p \rightarrow A_\pi (q U p) \land EPA_\pi \Box p \rightarrow A_\pi \Box p) = "Maintain p true and if that is not guaranteeedly possible, then it must maintain q true until p becomes true."$
Conclusions

• We extended $\pi$-CTL* to capable of comparing policies

• P-CTL* is a proper superest of mentioned existing languages

• P-CTL* is capable of capturing several degrees of “trying the best of reaching $p$”