Understanding AEMS2

What does "anytime" mean? Why is it important in an online solver?

The **anytime** means: an algorithm that can return a valid solution to a problem even if it is interrupted before it ends. That is the anytime algorithm, unlike the traditional (offline) algorithm apply to artificial intelligence or AI take a long time to complete results, is designed to complete in a shorter amount of time.

Since the online solver's input is piece-by-piece in a serial fashion, and the end user actually sometime are waiting the outcomes, the program or algorithm need to return the answer before user lost their patient. In this kind of situation, an anytime algorithm is able to return a partial answer, whose quality depends on the amount of computation they were able to perform. The answer generated by this algorithm is an approximation of the correct answer.

So anytime algorithm is important in online solver because it can give intelligent systems the ability to make results of better quality in return for turn-around time. And it also maintains the last result, so that if they are given more time, they can continue from where they left off to obtain an even better result.

How is the error calculated in AEMS2? Is it just the difference between upper bound and lower bound? Write the formula.

The error calculated in AEMS2 is not just the difference between upper bound and lower bound. Because of these 3 factors:

- The discount factor γ in the optimal value function means reduce the contribution of this error on the parent belief stat's value.
- The $\max_{a\in A}$ … operator in the optimal value function indicates that we need to consider only the values of belief states that can be reached by doing a sequence of optimal actions
- The probability of a certain fringe belief state can be reached by a sequence of optimal actions

So, the error formula of the expected error in AEMS2 is:

$$E\left(b^d\right) = \gamma^d P\left(b^d\right)\hat{\epsilon}\left(b^d\right) \tag{1}$$

In this equation, b^d is the belief state at depth d on the current belief state b_0, and γ is the discount factor. While $P\left(b^d\right)$ is the probability of reaching a certain fringe belief state b^d at depth d (detail in equation 2) and $\hat{\epsilon}\left(b^d\right)$ is the difference between upper and lower bound (detail in equation ()).

$$P\left(b^d\right) = \prod_{i=0}^{d-1} P\left(o^i \middle| b^i, a^i\right) P\left(a^i \middle| b^i\right) \tag{2}$$

$$\hat{\epsilon}(b) = U(b) - L(b) \tag{3}$$

For equation 2, $P\left(o^i \middle| b^i, a^i\right)$ is the probability of observing o^i after doing action a^i in belief state b^i (detail in equation 4) and $P\left(a^i \middle| b^i\right)$ is the current action with the highest upper bound as the optimal action (detail in equation 5).

$$P(o|b,a) = \sum_{s\in S} O(o,a,s') \sum_{s\in S} T\,(s,a,s')b(s) \tag{4}$$

$$P(a|b) = \begin{cases} 1 \;\; if\; a = \underset{a'\in A}{argmax}\, U(a',b) \\ 0 \;\; otherwise \end{cases} \tag{5}$$

For equation 3, $U(b)$ is the upper bound on $V^*(b)$ and $L(b)$ is the upper bound on $V^*(b)$.

Write down the error term that AEMS2 would use for the unexpanded nodes in the And-Or tree figure (b_2 to b_8). (Discount factor $\gamma = 0.95$)

From the chat, we can know that, for $\hat{\epsilon}(\mathrm{b})$:

$$\hat{\epsilon}(\mathrm{b}_2) = U(b_2) - L(b_2) = 20 - 15 = 5$$

$$\hat{\epsilon}(\mathrm{b}_3) = U(b_3) - L(b_3) = 15 - 9 = 6$$

$$\hat{\epsilon}(\mathrm{b}_4) = U(b_4) - L(b_4) = 18 - 10 = 8$$

$$\hat{\epsilon}(\mathrm{b}_5) = U(b_5) - L(b_5) = 14 - 6 = 8$$

$$\hat{\epsilon}(\mathrm{b}_6) = U(b_6) - L(b_6) = 12 - 9 = 3$$

$$\hat{\epsilon}(\mathrm{b}_7) = U(b_7) - L(b_7) = 20 - 11 = 9$$

$$\hat{\epsilon}(\mathrm{b}_8) = U(b_8) - L(b_8) = 12 - 10 = 2$$

From the chat, we can know that, for $U(a|b)$:

$$\begin{cases} U(a_1|b_0) = 17.9 \\ U(a_2|b_0) = 18.7 \end{cases} \qquad \begin{cases} U(a_1|b_1) = 11.5 \\ U(a_2|b_1) = 16.9 \end{cases}$$

So, for $P(a|b)$ we have:

$$\begin{cases} P(a_1|b_0) = 0 \\ P(a_2|b_0) = 1 \end{cases} \qquad \begin{cases} P(a_1|b_1) = 0 \\ P(a_2|b_1) = 1 \end{cases}$$

From the chat, we can know that, for $P(o|b,a)$:

$$\begin{cases} P(z_1|b_0,a_1) = 0.7 \\ P(z_2|b_0,a_1) = 0.3 \end{cases} \qquad \begin{cases} P(z_1|b_0,a_2) = 0.5 \\ P(z_2|b_0,a_2) = 0.5 \end{cases}$$

$$\begin{cases} P(z_1|b_1,a_1) = 0.6 \\ P(z_2|b_1,a_1) = 0.4 \end{cases} \qquad \begin{cases} P(z_1|b_1,a_2) = 0.2 \\ P(z_2|b_1,a_2) = 0.8 \end{cases}$$

With these values, for depth 1, we can get:

$$P(b_1) = \prod_{i=0}^{1-1} P\left(o^i \middle| b^i, a^i\right) P\left(a^i \middle| b^i\right) = P(z_1|b_0,a_1) * \;(a_1|b_0) = 0$$

$$P(b_2) = \prod_{i=0}^{1-1} P\left(o^i \middle| b^i, a^i\right) P\left(a^i \middle| b^i\right) = P(z_2|b_0,a_1) * \;(a_1|b_0) = 0$$

$$P(b_3) = \prod_{i=0}^{1-1} P\left(o^i \middle| b^i, a^i\right) P\left(a^i \middle| b^i\right) = P(z_1|b_0, a_2) * \ (a_1|b_0) = 0.5$$

$$P(b_4) = \prod_{i=0}^{1-1} P\left(o^i \middle| b^i, a^i\right) P\left(a^i \middle| b^i\right) = P(z_2|b_0, a_2) * \ (a_1|b_0) = 0.5$$

With these values, for depth 2, we can get:

$$P(b_5) = \prod_{i=0}^{2-1} P\left(o^i \middle| b^i, a^i\right) P\left(a^i \middle| b^i\right) = P(z_1|b_1, a_1) * \ (a_1|b_1) * P(z_1|b_0, a_1) * \ (a_1|b_0) = 0$$

$$P(b_6) = \prod_{i=0}^{2-1} P\left(o^i \middle| b^i, a^i\right) P\left(a^i \middle| b^i\right) = P(z_2|b_1, a_1) * \ (a_1|b_1) * P(z_1|b_0, a_1) * \ (a_1|b_0) = 0$$

$$P(b_7) = \prod_{i=0}^{2-1} P\left(o^i \middle| b^i, a^i\right) P\left(a^i \middle| b^i\right) = P(z_1|b_1, a_2) * \ (a_2|b_1) * P(z_1|b_0, a_1) * \ (a_1|b_0) = 0$$

$$P(b_8) = \prod_{i=0}^{2-1} P\left(o^i \middle| b^i, a^i\right) P\left(a^i \middle| b^i\right) = P(z_1|b_1, a_2) * \ (a_2|b_1) * P(z_1|b_0, a_1) * \ (a_1|b_0) = 0$$

With all the values before, now we can compute the error term for each unexpanded node:

$$\mathrm{E}(b_2) = \gamma^1 \mathrm{P}(b_2)\hat{\epsilon}(b_2) = 0.95 * 0 * 5 = 0$$

$$\mathrm{E}(b_3) = \gamma^1 \mathrm{P}(b_3)\hat{\epsilon}(b_3) = 0.95 * 0.5 * 6 = 2.85$$

$$\mathrm{E}(b_4) = \gamma^1 \mathrm{P}(b_4)\hat{\epsilon}(b_4) = 0.95 * 0.5 * 8 = 3.8$$

$$E(b_5) = \gamma^2 P(b_5)\hat{\epsilon}(b_5) = 0.95^2 * 0 * 8 = 0$$

$$E(b_6) = \gamma^2 P(b_6)\hat{\epsilon}(b_6) = 0.95^2 * 0 * 3 = 0$$

$$E(b_7) = \gamma^2 P(b_7)\hat{\epsilon}(b_7) = 0.95^2 * 0 * 9 = 0$$

$$E(b_8) = \gamma^2 P(b_8)\hat{\epsilon}(b_8) = 0.95^2 * 0 * 2 = 0$$