Risk Aware Belief-dependent Constrained Simplified POMDP Planning - Supplementary Material

Andrey Zhitnikov[1] and Vadim Indelman[2]
[1]Technion Autonomous Systems Program [2]Department of Aerospace Engineering

Technion - Israel Institute of Technology, Haifa 32000, Israel
andreyz@campus.technion.ac.il, vadim.indelman@technion.ac.il

This document provides supplementary material to the paper [3]. Therefore, it should not be considered a self-contained document, but instead regarded as an appendix of [3]. Throughout this report, all notations and definitions are with compliance to the ones presented in [3].

This supplementary document contains:

1. The proofs of the claims of the main manuscript 1;
2. Discussion about calculating the posterior belief conditioned on the safe prior belief 2;
3. Discussion on applying the simplification paradigm in the context of the second form of the constraint 3;
4. Our novel algorithm for chance constrained continuous POMDP 4;
5. Probability of reaching the goal 5;
6. Proof that mean distance ro goal accounts for uncertainty 6.

1 Proofs

1.1 Proof of Lemma 1 (representation of our general constraint).

$$P\left(c(b_{k:k+L})|b_k,\pi,a_k\right) = \int_{\substack{b_{k+1:k+L}\\ z_{k+1:k+L}}} P\left(c(b_{k:k+L})|b_k,\pi,a_k,b_{k+1:k+L}\right)$$
$$\mathbb{P}(b_{k+1:k+L}|b_k,\pi,a_k,z_{k+1:k+L})p(z_{k+1:k+L}|b_k,\pi,a_k) = \tag{1}$$
$$\mathbb{E}_{z_{k+1:k+L}}\left[c(b_{k:k+L})\middle|b_k,\pi,a_k\right]. \tag{2}$$

We used the fact that $p(b_{k+1:k+L}|b_k,\pi,a_k,z_{k+1:k+L})$ is Dirac's delta function. ■

1.2 Proof of Theorem 1

Let us note that since the belief is a conditional distribution it is a random variable and should be denoted $b(x;\omega)$. For clarity we will denote sometimes $b(\omega)$. Since it holds that $l(b_{\ell+1}(\omega),b_\ell(\omega);\omega) \leq \phi(b_{\ell+1}(\omega),b_\ell(\omega);\omega)\ \forall\omega\in\Omega$ (such that $P(\Omega)=1$) so the following relation over the sets holds

$$\{\omega : l(b_{\ell+1}(\omega),b_\ell(\omega);\omega) > \delta\} \subseteq \{\omega : \phi(b_{\ell+1}(\omega),b_\ell(\omega);\omega) > \delta\} \quad \forall\delta\in\mathbb{R}. \tag{3}$$

Hence

$$P\Big(\{\omega : l(\omega) > \delta\}\Big) \leq P\Big(\{\omega : \phi(\omega) > \delta\}\Big) \quad \forall\delta\in\mathbb{R}. \tag{4}$$

This is known as *usual stochastic order*. We arrive to the desired result thanks to

$$\sum_{\ell=k}^{k+L-1} l(b_{\ell+1}(\omega),b_\ell(\omega);\omega) \leq \sum_{\ell=k}^{k+L-1} \phi(b_{\ell+1}(\omega),b_\ell(\omega);\omega)\ \forall\omega\in\Omega \tag{5}$$

■

1.3 Proof of Lemma 2 (probability of the trajectory)

$$\mathbb{P}(x_{k:k+L}|b_k,\pi_{k+1:k+L-1},a_k)=\int\limits_{z_{k+1:k+L-1}}\mathbb{P}(x_{k:k+L},z_{k+1:k+L-1}|b_k,\pi)= \tag{6}$$

$$\int\limits_{z_{k+1:k+L-1}}\mathbb{P}(x_{k+L}|z_{k:k+L-1},x_{k:k+L-1},b_k,\pi)\mathbb{P}(z_{k:k+L-1},x_{k:k+L-1}|b_k,\pi)= \tag{7}$$

$$\int\limits_{z_{k+1:k+L-1}}\mathbb{P}_T(x_{k+L}|x_{k+L-1},a_{k+L-1})\mathbb{P}(z_{k+L-1}|x_{k:k+L-1},z_{k+1:k+L-2},b_k,\pi)\mathbb{P}(x_{k:k+L-1},z_{k+1:k+L-2}|b_k,\pi)= \tag{8}$$

$$\int\limits_{z_{k+1:k+L-1}}\mathbb{P}_T(x_{k+L}|x_{k+L-1},a_{k+L-1})\mathbb{P}_Z(z_{k+L-1}|x_{k+L-1})\mathbb{P}(x_{k:k+L-1},z_{k+1:k+L-2}|b_k,\pi). \tag{9}$$

We observe the recurrence relation. Overall

$$\mathbb{P}(\tau|b_k,\pi_{k+1:k+L-1},a_k)=$$
$$\mathbb{P}_T(x_{k+1}|x_k,a_k)b_k(x_k)\int_{z_{k+1:k+L-1}}\prod_{i=k+1}^{k+L-1}\mathbb{P}_T(x_{i+1}|x_i,\pi(b_i(b_{i-1},a_{i-1},z_i)))\mathbb{P}_Z(z_i|x_i) \tag{10}$$

■

1.4 Proof of Lemma 3 (average over the safe posteriors)

$$\underbrace{P(\bigwedge_{i=k}^{k+L}\mathbf{1}\left\{x_i\in\mathcal{X}_i^{\text{safe}}\right\}|b_k,\pi)}_{(a)}=P(\mathbf{1}\left\{x_k\in\mathcal{X}_k^{\text{safe}}\right\}|b_k)\underbrace{P(\bigwedge_{i=k+1}^{k+L}\mathbf{1}\left\{x_i\in\mathcal{X}_i^{\text{safe}}\right\}|\mathbf{1}\left\{x_k\in\mathcal{X}_k^{\text{safe}}\right\},b_k,\pi)}_{(b)} \tag{11}$$

Let us focus on the expression we marked by (b)

$$P(\bigwedge_{i=k+1}^{k+L}\mathbf{1}\left\{x_i\in\mathcal{X}_i^{\text{safe}}\right\}|\mathbf{1}\left\{x_k\in\mathcal{X}_k^{\text{safe}}\right\}b_k,\pi)=$$
$$\int_{b_{k+1}}P(\bigwedge_{i=k+1}^{k+L}\mathbf{1}\left\{x_i\in\mathcal{X}_i^{\text{safe}}\right\}|b_{k+1},\mathbf{1}\left\{x_k\in\mathcal{X}_k^{\text{safe}}\right\},b_k,\pi)P(b_{k+1}|\mathbf{1}\left\{x_k\in\mathcal{X}_k^{\text{safe}}\right\}b_k,\pi)= \tag{12}$$
$$\int_{b_{k+1}}P(b_{k+1}|\mathbf{1}\left\{x_k\in\mathcal{X}_k^{\text{safe}}\right\},b_k,\pi)P(\bigwedge_{i=k+1}^{k+L}\mathbf{1}\left\{x_i\in\mathcal{X}_i^{\text{safe}}\right\}|b_{k+1},\pi) \tag{13}$$

Merging the two expressions we obtain

$$P(\bigwedge_{i=k}^{k+L}\mathbf{1}\left\{x_i\in\mathcal{X}_i^{\text{safe}}\right\}|b_k,\pi)=P(\mathbf{1}\left\{x_k\in\mathcal{X}_k^{\text{safe}}\right\}|b_k)\cdot$$
$$\int_{b_{k+1}}P(b_{k+1}|\mathbf{1}\left\{x_k\in\mathcal{X}_k^{\text{safe}}\right\},b_k,\pi)\underbrace{P(\bigwedge_{i=k+1}^{k+L}\mathbf{1}\left\{x_i\in\mathcal{X}_i^{\text{safe}}\right\}|b_{k+1},\pi)}_{(c)} \tag{14}$$

We observe that expression (a) is very similar to (c), namely

$$P(\bigwedge_{i=k+1}^{k+L}\mathbf{1}\left\{x_i\in\mathcal{X}_i^{\text{safe}}\right\}|b_{k+1},\pi)=P(\mathbf{1}\left\{x_{k+1}\in\mathcal{X}_{k+1}^{\text{safe}}\right\}|b_{k+1})$$
$$\int_{b_{k+2}}P(b_{k+2}|\mathbf{1}\left\{x_{k+1}\in\mathcal{X}_{k+1}^{\text{safe}}\right\},b_{k+1},\pi)\underbrace{P(\bigwedge_{i=k+2}^{k+L}\mathbf{1}\left\{x_i\in\mathcal{X}_i^{\text{safe}}\right\}|b_{k+2},\pi)}_{(d)} \tag{15}$$

Merging the two we got

$$P(\bigwedge_{i=k}^{k+L} \mathbf{1}\left\{x_i \in \mathcal{X}_i^{\text{safe}}\right\}|b_k,\pi) = P(\mathbf{1}\left\{x_k \in \mathcal{X}_k^{\text{safe}}\right\}|b_k)\cdot$$

$$\int_{b_{k+1}} P(b_{k+1}|\mathbf{1}\left\{x_k \in \mathcal{X}_k^{\text{safe}}\right\}, b_k, \pi)P(\mathbf{1}\left\{x_{k+1} \in \mathcal{X}_{k+1}^{\text{safe}}\right\}|b_{k+1}) \tag{16}$$

$$\int_{b_{k+2}} P(b_{k+2}|\mathbf{1}\left\{x_{k+1} \in \mathcal{X}_{k+1}^{\text{safe}}\right\}, b_{k+1}, \pi)P(\bigwedge_{i=k+2}^{k+L} \mathbf{1}\left\{x_i \in \mathcal{X}_i^{\text{safe}}\right\}|b_{k+2},\pi). \tag{17}$$

We behold the recurrence relation.

Now we show that marginalization can be done with respect to the observations. Let us assume that i is the last index ($i = k+L$)

$$\int\limits_{b_i} P(\mathbf{1}\left\{x_i \in \mathcal{X}_i^{\text{safe}}\right\}|b_i)P(b_i|\mathbf{1}\left\{x_{i-1} \in \mathcal{X}_{i-1}^{\text{safe}}\right\}, b_{i-1},\pi) = \tag{18}$$

$$\int\limits_{b_i}\int_{z_i \in \mathcal{Z}} P(\mathbf{1}\left\{x_i \in \mathcal{X}_i^{\text{safe}}\right\}|b_i)\delta(b_i - \psi(b_{i-1}, \mathbf{1}\left\{x_{i-1} \in \mathcal{X}_{i-1}^{\text{safe}}\right\}, a_{i-1}, z_i))$$

$$p(z_i|a_{i-1}, b_{i-1}, \mathbf{1}\left\{x_{i-1} \in \mathcal{X}_{i-1}^{\text{safe}}\right\}) = \tag{19}$$

$$\int_{z_i \in \mathcal{Z}}\int\limits_{b_i} P(\mathbf{1}\left\{x_i \in \mathcal{X}_i^{\text{safe}}\right\}|b_i)\delta(b_i - \psi(b_{i-1}, \mathbf{1}\left\{x_{i-1} \in \mathcal{X}_{i-1}^{\text{safe}}\right\}, a_{i-1}, z_i))$$

$$p(z_i|a_{i-1}, b_{i-1}, \mathbf{1}\left\{x_{i-1} \in \mathcal{X}_{i-1}^{\text{safe}}\right\}) = \tag{20}$$

$$\int_{z_i \in \mathcal{Z}} P(\mathbf{1}\left\{x_i \in \mathcal{X}_i^{\text{safe}}\right\}|\psi(b_{i-1}, \mathbf{1}\left\{x_{i-1} \in \mathcal{X}_{i-1}^{\text{safe}}\right\}, a_{i-1}, z_i))$$

$$p(z_i|a_{i-1}, b_{i-1}, \mathbf{1}\left\{x_{i-1} \in \mathcal{X}_{i-1}^{\text{safe}}\right\}) = \tag{21}$$

$$\mathbb{E}_{z_i}\left[P(\mathbf{1}\left\{x_i \in \mathcal{X}_i^{\text{safe}}\right\}|\psi(b_{i-1}, \mathbf{1}\left\{x_{i-1} \in \mathcal{X}_{i-1}^{\text{safe}}\right\}, a_{i-1}, z_i))\middle| a_{i-1}, b_{i-1}, \mathbf{1}\left\{x_{i-1} \in \mathcal{X}_{i-1}^{\text{safe}}\right\}\right] \tag{22}$$

We plug this result into expression for $i-1$ and do the same trick two b_{i-1} ■

2 Calculating the posterior conditioned on the safe prior

$$p(b'|b, a, \mathbf{1}\left\{x \in \mathcal{X}^{\text{safe}}\right\}) = \int_{z \in \mathcal{Z}} p(b'|b, a, z, \mathbf{1}\left\{x \in \mathcal{X}^{\text{safe}}\right\})p(z|a, b, \mathbf{1}\left\{x \in \mathcal{X}^{\text{safe}}\right\}) =$$

$$\int_{z \in \mathcal{Z}} \delta(b' - \psi(b^{\text{safe}}, a, z))p(z|a, b, \mathbf{1}\left\{x \in \mathcal{X}^{\text{safe}}\right\}) \tag{23}$$

We first calculate the propagated belief conditioned on the safe prior.

$$p(x'|b, a, \mathbf{1}\left\{x \in \mathcal{X}^{\text{safe}}\right\}) = \frac{\int_{x \in \mathcal{X}} \mathbf{1}\left\{x \in \mathcal{X}^{\text{safe}}\right\} \mathbb{P}_T(x'|x, a)b(x)}{\int_{x \in \mathcal{X}} \mathbf{1}\left\{x \in \mathcal{X}^{\text{safe}}\right\} b(x)} \tag{24}$$

b and event safe, meaning that belief supposed to be zero at non safe places. Finally,

$$p(z|a, b, \mathbf{1}\left\{x \in \mathcal{X}^{\text{safe}}\right\}) = \int_{x' \in \mathcal{X}'} \mathbb{P}_Z(z|x')p(x'|b, a, \mathbf{1}\left\{x \in \mathcal{X}^{\text{safe}}\right\}). \tag{25}$$

We can also look at the above from slightly different angle. We define b^{safe}

$$b^{\text{safe}}(x) = \frac{\mathbf{1}\left\{x \in \mathcal{X}^{\text{safe}}\right\} b(x)}{\int_{\xi \in \mathcal{X}} \mathbf{1}\left\{\xi \in \mathcal{X}^{\text{safe}}\right\} b(\xi)}. \tag{26}$$

Now the ψ is conventional belief update operator receiving as input $\psi(b^{\text{safe}}, a, z')$.

3 Application of the simplification on the second form of the constraint

Suppose $\forall\ b_{\ell+1}, b_\ell \in \mathcal{B}$ holds

$$l(b_{\ell+1}, b_\ell) \le \phi(b_{\ell+1}, b_\ell) \le u(b_{\ell+1}, b_\ell), \tag{27}$$

so

$$P\left(\prod_{\ell=k}^{k+L-1} \mathbf{1}\left\{\phi(b_{\ell+1}, b_\ell) \ge \delta\right\}|b_k, \pi\right) \ge P\left(\prod_{\ell=k}^{k+L-1} \mathbf{1}\left\{l(b_{\ell+1}, b_\ell) \ge \delta\right\}|b_k, \pi\right), \tag{28}$$

$$P\left(\prod_{\ell=k}^{k+L-1} \mathbf{1}\left\{\phi(b_{\ell+1}, b_\ell) \ge \delta\right\}|b_k, \pi\right) \le P\left(\prod_{\ell=k}^{k+L-1} \mathbf{1}\left\{u(b_{\ell+1}, b_\ell) \ge \delta\right\}|b_k, \pi\right), \tag{29}$$

The proof is similar to the proof of Theorem 1 from section 1.2 in this document. Since it holds that $l(b_{\ell+1}(\omega), b_\ell(\omega); \omega) \le \phi(b_{\ell+1}(\omega), b_\ell(\omega); \omega)\ \forall \omega \in \Omega$ (such that $P(\Omega) = 1$) so the following relation over the sets holds

$$\{\omega : l(b_{\ell+1}(\omega), b_\ell(\omega); \omega) > \delta\} \subseteq \{\omega : \phi(b_{\ell+1}(\omega), b_\ell(\omega); \omega) > \delta\} \quad \forall \delta \in \mathbb{R}. \tag{30}$$

Hence

$$\bigcap_{\ell=k}^{k+L-1} \{\omega : l(b_{\ell+1}(\omega), b_\ell(\omega); \omega) > \delta\} \subseteq \bigcap_{\ell=k}^{k+L-1} \{\omega : \phi(b_{\ell+1}(\omega), b_\ell(\omega); \omega)) > \delta\} \quad \forall \delta \in \mathbb{R}. \tag{31}$$

$$P\Bigg(\bigcap_{\ell=k}^{k+L-1} \{\omega : l(b_{\ell+1}(\omega), b_\ell(\omega); \omega) > \delta\}\Bigg) \le P\Bigg(\bigcap_{\ell=k}^{k+L-1} \{\omega : \phi(b_{\ell+1}(\omega), b_\ell(\omega); \omega)) > \delta\}\Bigg) \quad \forall \delta \in \mathbb{R}. \tag{32}$$

Another form of writing the above is

$$P\left(\prod_{\ell=k}^{k+L-1} \mathbf{1}\left\{\phi(b_{\ell+1}, b_\ell) \ge \delta\right\}|b_k, \pi\right) \ge P\left(\prod_{\ell=k}^{k+L-1} \mathbf{1}\left\{l(b_{\ell+1}, b_\ell) \ge \delta\right\}|b_k, \pi\right), \tag{33}$$

By switching the roles l to ϕ and ϕ to u; we prove the (29). ■

4 Chance constrained continuous POMDP

In this section, we discuss our baseline algorithm and it's differences with respect to PCSS (Alg. 1 in the main manuscript). We present Chance Constrained Sparse Sampling 1 (CCSS). The two of the algorithms are novel since they are dealing with challenging continuous domains. This is leaving aside the fact that PCSS to our knowledge it is the first algorithm in continuous domain dealing with our new formulated the probabilistic constraints in POMDP setting. Two of the algorithms assumes that the belief tree construction is not coupled with the values of the utility function. In other words we do not deal with anytime setting described in section 3.3. Let ϵ be zero. Let us emphasize the differences in the two algorithms. The PCSS formulation is equivalent to

$$\forall b^j_{k:k+L}(b_k, a_k, \pi_{k+1:k+L-1}, z^j_{k+1:k+L}) \quad \left(\prod_{\ell=k}^{k+L} \mathbf{1}\left\{P(x_\ell \in \mathcal{X}^{\text{safe}}_\ell | b^j_\ell) \ge \delta\right\}\right) = 1, \tag{34}$$

where $j = 1 \dots M$ and $M = m^L$. On the contrary the CCSS at each subtree averages the probability to be safe given belief updated from previous safe belief.

5 Probability of reaching the goal

Before we start this section, let us denote $f \equiv g$ for two operators, if we have $f(x) = g(x) \quad \forall x$. In this section we discuss one more important constraint appearing in [1], probability of reaching the goal. Through the main manuscript for clarity we assumed that the operator ϕ is identical for all time indices. Let us fix that. We redefine the constraint of the first form as follows

$$c(b_{k:k+L}) \triangleq \mathbf{1}\left\{\left(\sum_{\ell=k}^{k+L-1} \phi_{\ell+1}(b_{\ell+1}, b_\ell)\right) \ge \delta\right\}. \tag{35}$$

Further, let $\phi_{\ell+1}(\cdot) \equiv 0 \quad \forall \ell \ne L$ and

$$\phi_L(b_L) = P(x_L \in \mathcal{X}^{\text{goal}} | b_L). \tag{36}$$

All the discussions and comparison with conventional constraint applies here.

Algorithm 1 Chance Constrained BMDP Sparse Sampling

procedure CCSS(belief: b, belief: $b_{\text{from safe}}$, depth: d)
    $\varphi \leftarrow \phi(b)$
    $\Phi(b) \leftarrow \{\}$
    if $d = 0$ **then**
        $\Phi(b) \leftarrow \Phi(b) \cup \{\varphi\}$
        return (Null, $\rho(b)$)
    end if
    $(a^*, u^*) \leftarrow$ (Null, $-\infty$)
    for $a \in \mathcal{A}$ **do**
        u $\leftarrow$ 0.0
        r $\leftarrow \rho(b)$
        Calculate propagated belief b^{prop}
        Make $b_{\text{from safe}}$ safe
        Calculate propagated from safe belief $b^{\text{prop}}_{\text{from safe}}$
        $\Phi(ba) \leftarrow \{\}$ ▷ empty set of boolean values, this level constraints
        status $\leftarrow$ true
        for _ $\in 1 : m_d$ **do**
            Sample x from b^{prop}
            Sample Observation z from observation model using x
            Sample $x^{\text{from safe}}$ from $b^{\text{prop}}_{\text{from safe}}$
            Sample Observation $z^{\text{from safe}}$ from observation model using $x^{\text{from safe}}$
            Calculate posterior b'
            Calculate posterior $b'_{\text{from safe}}$
            _, $u' \leftarrow$ CCSS(b',$b'_{\text{from safe}}$,$d-1$)
            $u+ = (r + \gamma \cdot u')/m_d$
            $\Phi(ba) \leftarrow \Phi(ba) \cup \Phi(b')$
        end for
        if scale **then**
            if $\left(\frac{1}{|\Phi(ba)|}\sum_{\varphi' \in \Phi(ba)} \varphi'\right) < \delta^d$ **then**
                status $\leftarrow$ false
            end if
        else
            if $\left(\frac{1}{|\Phi(ba)|}\sum_{\varphi' \in \Phi(ba)} \varphi'\right) < \delta$ **then**
                status $\leftarrow$ false
            end if
        end if
        for $\varphi' \in \Phi(ba)$ **do**
            $\varphi' \leftarrow \varphi' \cdot \varphi$
        end for
        if status $\wedge\, u > u^*$ **then**
            $(a^*, u^*) \leftarrow (a, u)$
        end if
    end for
end procedure

6 Proof that mean distance to goal accounts for uncertainty

Let x be an arbitrary distributed random variable with μ_x and Σ being the expected value and covariance matrix of x, respectively.

$$\mathbb{E}\left[x^T \Lambda x\right] = \operatorname{tr}\left[\Lambda \Sigma_x\right] + \mu_x^T \Lambda \mu_x. \tag{37}$$

Since the quadratic form is a scalar quantity, $x^T \Lambda x = \operatorname{tr}(x^T \Lambda x)$.

Next, by the cyclic property of the trace operator,

$$\mathbb{E}[\operatorname{tr}(x^T \Lambda x)] = \mathbb{E}[\operatorname{tr}(\Lambda x x^T)]. \tag{38}$$

Since the trace operator is a linear combination of the components of the matrix, it therefore follows from the linearity of the expectation operator that

$$\mathbb{E}\left[\operatorname{tr}(\Lambda x x^T)\right] = \operatorname{tr}(\Lambda \mathbb{E}(x x^T)). \tag{39}$$

A standard property of variances then tells us that this is

$$\operatorname{tr}(\Lambda(\Sigma_x + \mu_x \mu_x^T)). \tag{40}$$

Applying the cyclic property of the trace operator again, we get

$$\operatorname{tr}(\Lambda\Sigma_x) + \operatorname{tr}(\Lambda\mu_x\mu_x^T) = \operatorname{tr}(\Lambda\Sigma_x) + \operatorname{tr}(\mu_x^T \Lambda\mu) = \operatorname{tr}(\Lambda\Sigma_x) + \mu_x^T \Lambda\mu_x. \tag{41}$$

■

References

1. A. Bry and N. Roy. Rapidly-exploring random belief trees for motion planning under uncertainty. In *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, pages 723–730, 2011.
2. Michael Kearns, Yishay Mansour, and Andrew Y Ng. A sparse sampling algorithm for near-optimal planning in large markov decision processes. *Machine learning*, 49(2):193–208, 2002.
3. A. Zhitnikov and V. Indelman. Risk aware belief-dependent constrained simplified pomdp planning. In *Proc. of the Intl. Symp. of Robotics Research (ISRR)*, 2022. submitted.