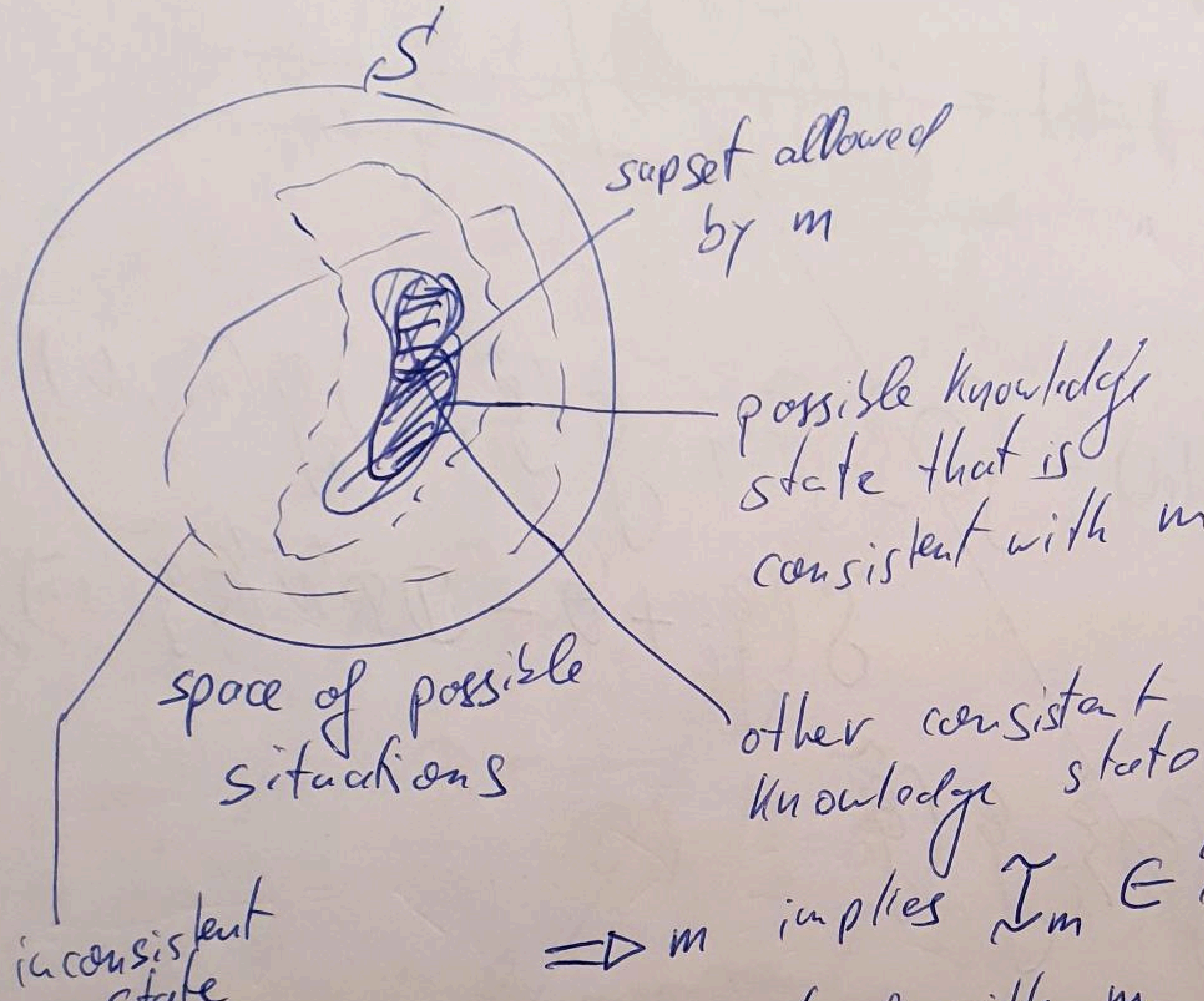
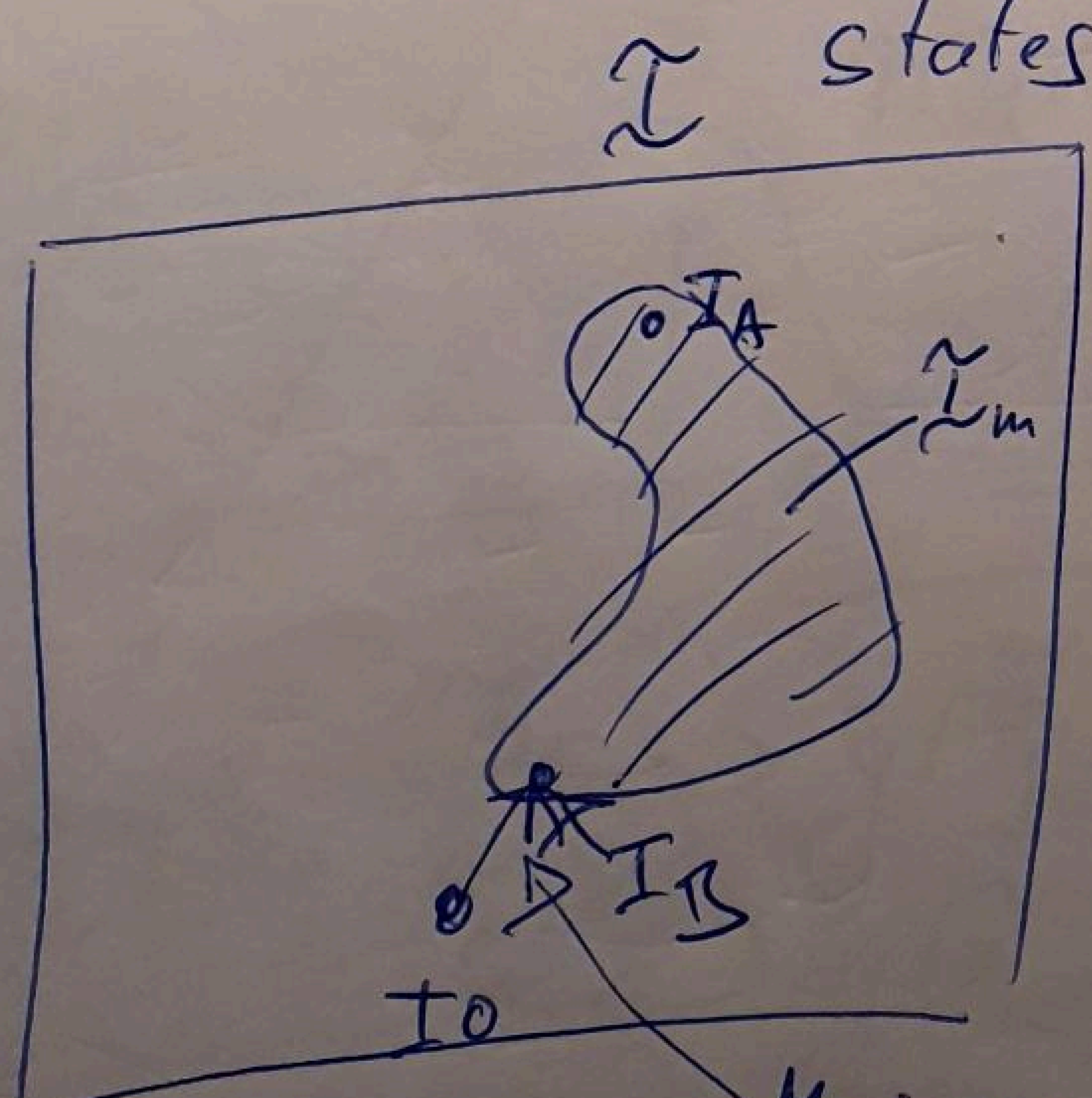
Entropic Communication

Alice choses a message $m \in \mathcal{M}$, which lets Bob update his knowledge $I_0 \xrightarrow{m} I_B$



$\Longrightarrow$ m implies $\tilde{I}_m \in \mathcal{I}$
a set of with m consistent knowledge states



$$I_B = \underset{I_B,\lambda}{\operatorname{argmin}}\left[D_{S}(I_B, I_0) + \lambda m \right]$$

Lagrange multiplier

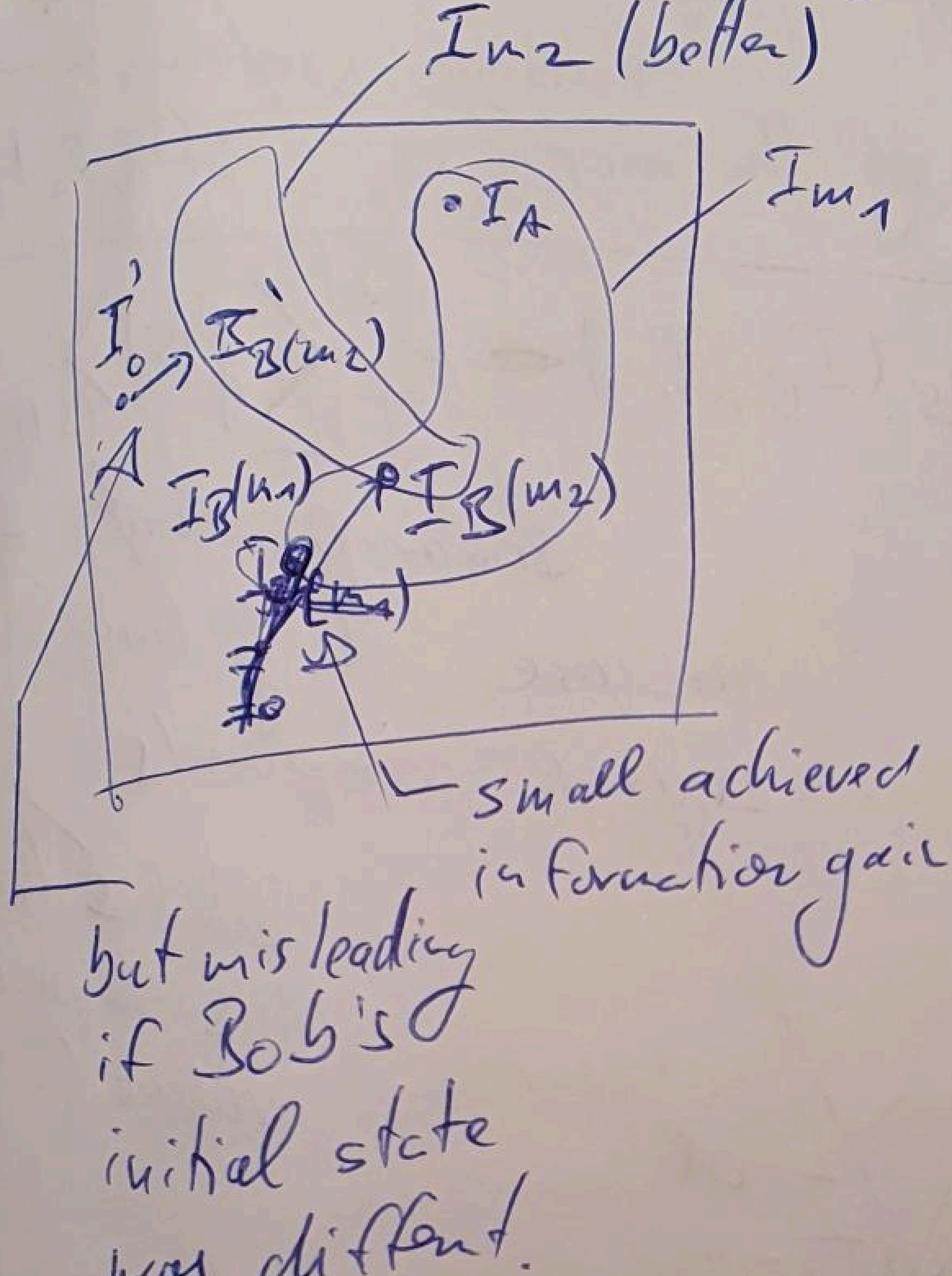
message in form of an equation

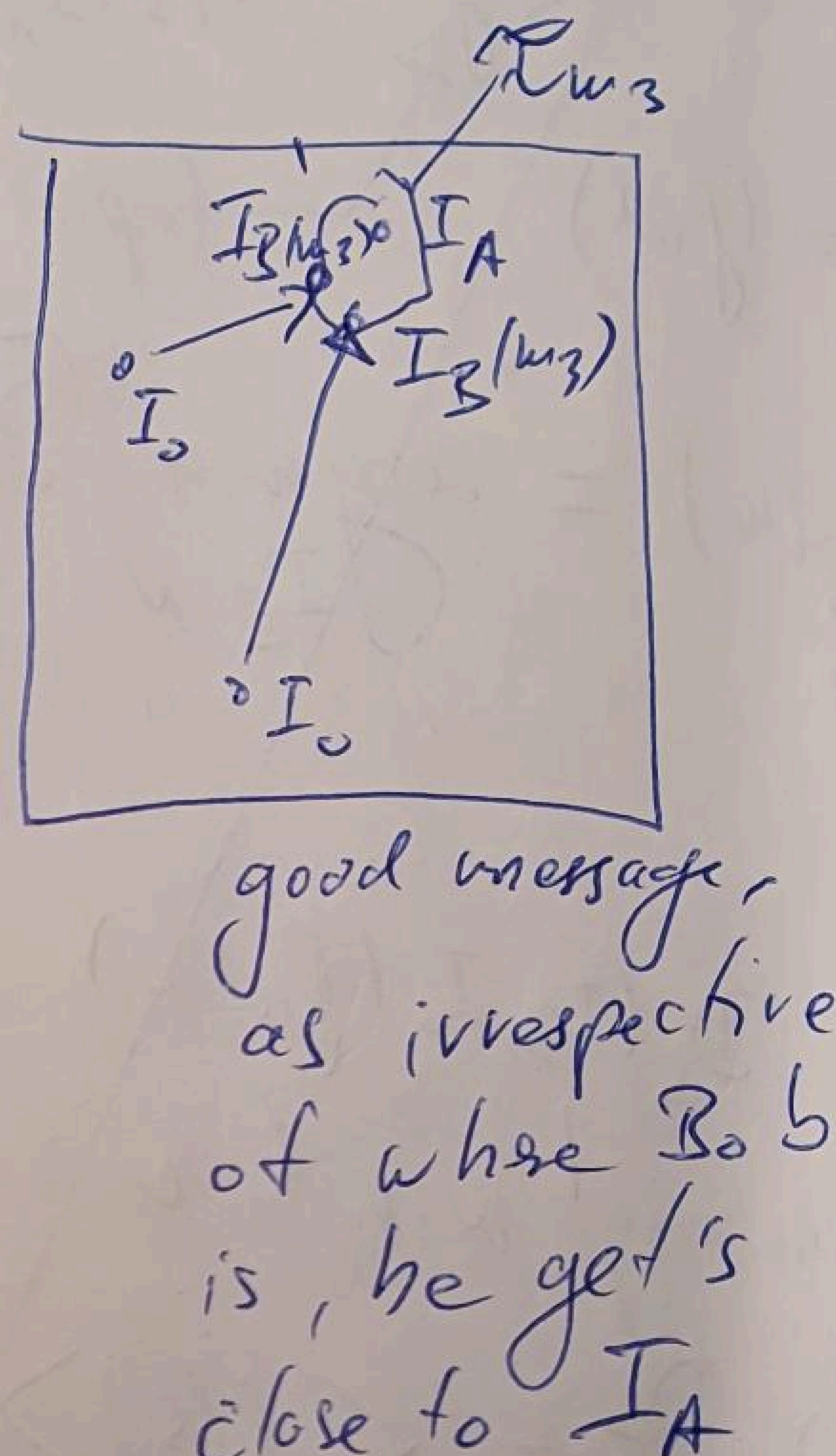
Maximum Entropy update, minimal added information

Which message should Alice chose?

That, which gets $D_S(I_A, I_B(m))$ minimal

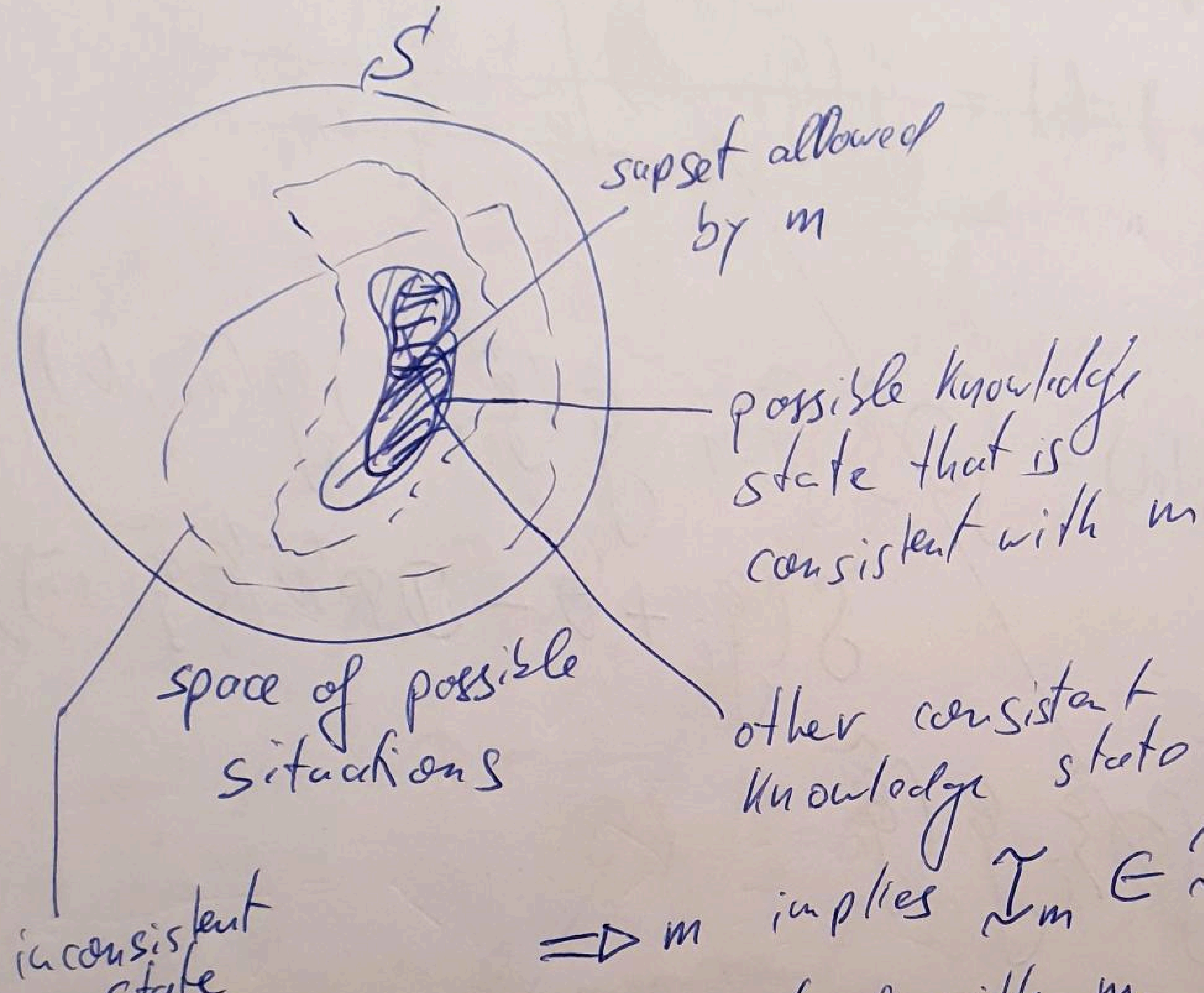
$$m = \underset{m}{\operatorname{argmin}}\, D_S\left(I_A, \underset{I_B,\lambda}{\operatorname{argmin}}\left[D_S(I_B, I_0) + \lambda m\right]\right)$$





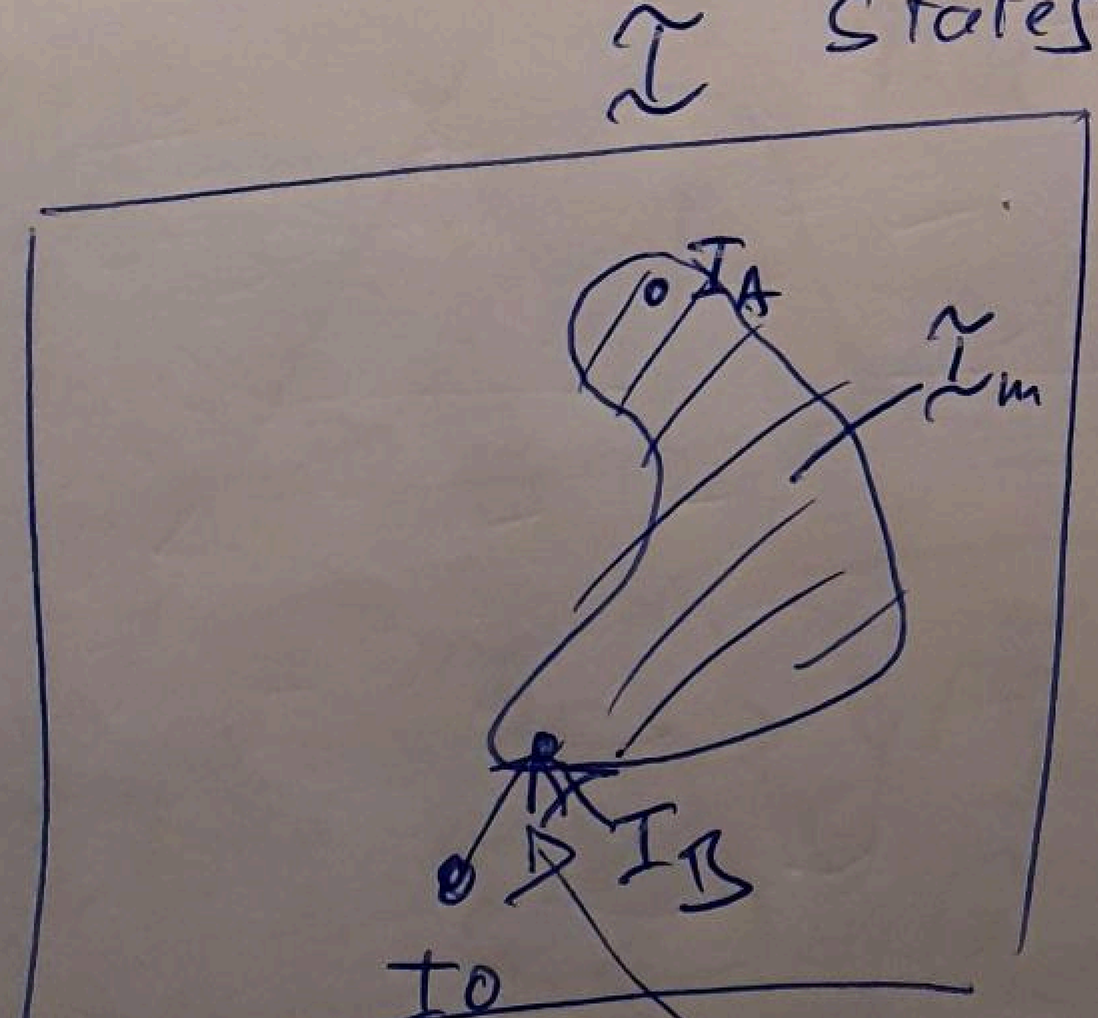
Entropic Communication

Alice choses a message $m \in \mathcal{M}$, which lets Bob update his knowledge $I_0 \xrightarrow{m} I_B$



$\Rightarrow$ m implies $\tilde{I}_m \in \mathcal{I}$

a set of with m consistent knowledge states



$$I_B = \underset{I_B,\lambda}{\operatorname{argmin}}\left[D_S(I_B, I_0) + \lambda m \right]$$

Lagrange multiplier

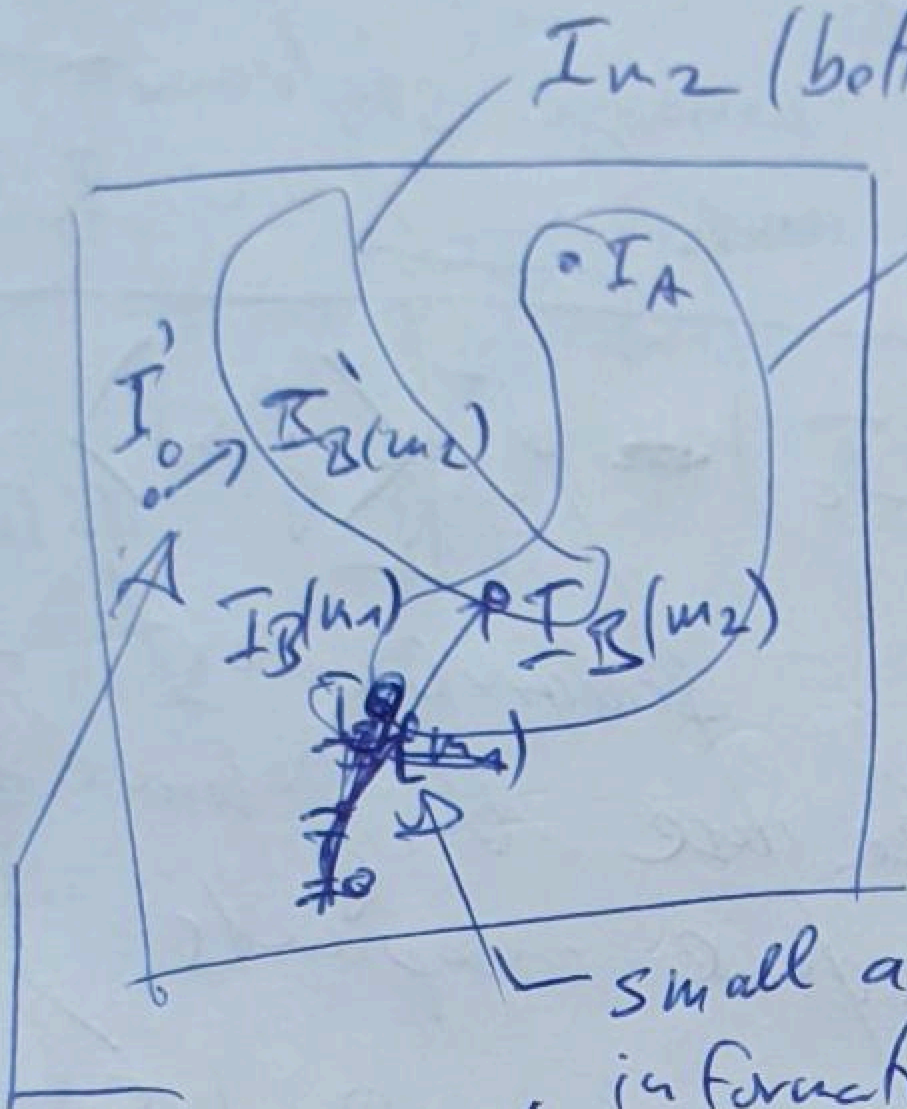
message in form of an equation

Maximum Entropy update, minimal added information

Which message should Alice chose?

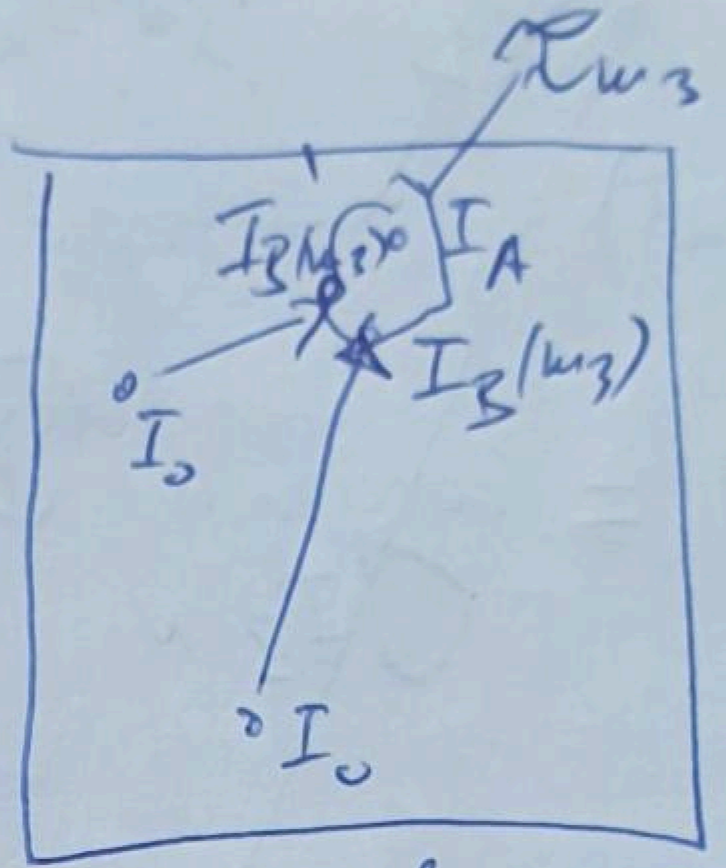
That, which gets $D_S(I_A, I_B(m))$ minimal

$$m = \arg\min_{m} D_S\left(I_A, \arg\min_{I_B,\lambda}\left[D_S(I_B, I_0) + \lambda m\right]\right)$$



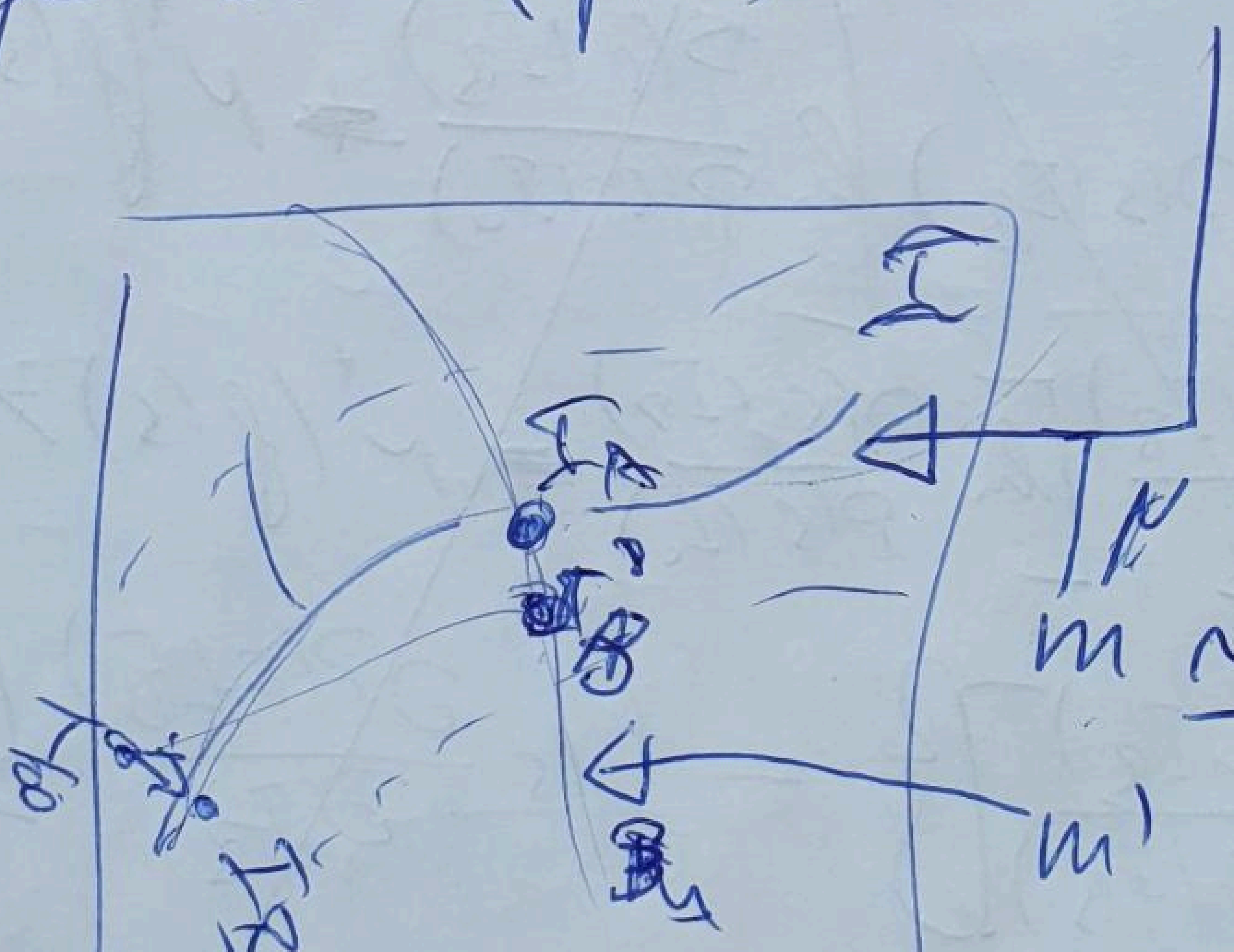
small achieved information gain

but misleading if Bob's initial state was diffent.



good message, as irrespective of where Bob is, he get's close to I_A

message $m = (f, d)$: $d = \langle f(s) \rangle_{(s|I_A)}$



$m \simeq \{I : \langle f(s) \rangle_{(s|I)} = d\}$

m'

Decoding a message

1) soft data constraint

$$m \hat{=} \langle f(s) \rangle_{(s|I_A)} = d \qquad f(s) = \begin{pmatrix} f_1(s) \\ \vdots \\ f_n(s) \end{pmatrix}, \; d = \begin{pmatrix} d_1 \\ \vdots \\ d_n \end{pmatrix}$$

d.

$$m = (f, d) \qquad \mathcal{D}_s(I_B, I_0 | \mu^\dagger m)$$

d — message data

f — message topic

$$I_B(m) = \underset{I_B,\, \mu}{\arg\min} \Big[\underbrace{\mathcal{D}_s(I_B, I_0) - \mu^\dagger \big[\langle f(s) \rangle_{(s|I_B)} - d \big]}_{\mathcal{D}_s(I_B, I_0 | \mu^\dagger m)} \Big]$$

Bob wants to constrain his knowledge to m

$$0 \overset{!}{=} \frac{\partial \mathcal{D}_s(I_B, I_0 | \mu^\dagger m)}{\partial \mu} = - \langle f(s) \rangle_{(s|I_B)} + d \overset{!}{=} 0$$

$$\langle f(s) \rangle_{(s|I_B)} = d = \langle f(s) \rangle_{(s|I_A)}$$

$\Rightarrow$ moment matching

$$0 \overset{!}{=} \frac{\partial \mathcal{D}_s(I_B, I_0 | \mu^\dagger m)}{\partial I_B} = \frac{\partial}{\partial I_B} \Big\{ \int ds\, P_s(s|I_B) \Big[\ln \frac{P_s(s|I_B)}{P_s(s|I_0)} - \mu^\dagger f(s) \Big] - d \Big\}$$

$$= \int ds\, \frac{\partial P_s(s|I_B)}{\partial I_B} \Big\{ \ln \frac{P_s(s|I_B)}{P(s|I_0)} - \mu^\dagger f(s) \Big\}$$

$$+ \int ds\, P(s|I_B) \underbrace{\frac{\partial}{\partial I_B} \ln \frac{P(s|I_B)}{P(s|I_0)}}_{1/P(s|I_B)\, \frac{\partial P(s|I_B)}{\partial I_B}}$$

$$= \int ds \frac{\partial P(s|I_B)}{\partial I_B} \left\{ \ln \frac{P(s|I_B)}{P(s|I_0)} - \mu^\dagger f(s) + \underbrace{\frac{P(s|I_B)}{P(s|I_B)}}_{1} \right\} \overset{!}{=} 0$$

solved if

$$\ln \frac{P(s|I_B)}{P(s|I_0)} = \mu^\dagger f(s) - 1$$

$$P(s|I_B) = P(s|I_0)\, e^{\mu^\dagger f(s) - 1}$$

We did not ensure normalization!

$$\langle 1 \rangle_{(s|I_B)} = 1$$

Let's do that by adding a zeroth entry to $f \to \tilde{f}$ and $d \to \tilde{d}$: $f_0(s) = 1$, $d_0 = 1$

$$\Rightarrow \frac{\partial D_s(I_B, I_0 | \mu^\dagger u)}{\partial \mu_0} = -\langle f_0(s) \rangle_{(s|I_B)} + d_0 = 0$$

$$\Rightarrow \langle 1 \rangle_{(s|I_B)} = 1 \checkmark$$

$$\tilde{\mu}^\dagger \tilde{f}(s) = \mu_0 f_0(s) + \underbrace{\sum_{i=1}^{n} \mu_u f_u(s)}_{\mu^\dagger f(s)}$$

$$\boxed{P(s|I_B) = P(s|I_0) \underbrace{e^{\mu_0 - 1}}_{=: Z^{-1}(\mu)} e^{\mu^\dagger f(s)}}$$

<u>Normalization</u> : $\langle 1 \rangle_{(s|I_B)} = 1$

$$1 = \int ds \, \frac{P(s|I_0)}{Z(\mu)} e^{\mu^\dagger f(s)}$$

$$\Rightarrow \boxed{\begin{array}{l} Z(\mu) := \int ds \, P(s|I_0) \, e^{\mu^\dagger f(s)} \\ \text{Zustandssumme, partition function} \end{array}}$$

<u>soft data constraints:</u>

chose μ such that $\langle f(s) \rangle_{(s|I_B)} = d$

How? Use Z!

$$d \overset{!}{=} \langle f(s) \rangle_{(s|I_B)} = \int \frac{ds \, f(s) \, P(s|I_0) \, e^{\mu^\dagger f(s)}}{Z(\mu)}$$

$$= \frac{1}{Z(\mu)} \frac{\partial}{\partial \mu} \int ds \, P(s|I_0) \, e^{\mu^\dagger f(s)}$$

$$= \frac{1}{Z(\mu)} \frac{\partial}{\partial \mu} Z(\mu) = \boxed{\frac{\partial}{\partial \mu} \ln Z(\mu) \overset{!}{=} d}$$

<u>Apparent information gain</u>

$$D_s(I_B, I_0) = \int ds \, P(s|I_B) \ln \frac{P(s|I_B)}{P(s|I_0)}$$

$$= \left\langle \ln \frac{e^{\mu^\dagger f(s)}}{Z(\mu)} \right\rangle_{(s|I_B)} = \langle \mu^\dagger f(s) \rangle_{(s|I_B)} - \ln Z(\mu)$$

$$\boxed{\mathcal{D}_s(I_B, I_0) = \mu^\dagger d - \ln Z(\mu)}$$

Max Ent decoding

message $m = (f, d)$: "$\langle f(s) \rangle_{(s|I_A)} = d$"

1) calculate $Z(\mu) = \int ds\, P(s|I_0)\, e^{\mu^\dagger f(s)}$

2) determine μ: $\dfrac{\partial \ln Z(\mu)}{\partial \mu} \overset{!}{=} d$

3) assign $P(s|I_B) = \dfrac{P(s|I_0)\, e^{\mu^\dagger f(s)}}{Z(\mu)}$

4) enjoy information gain

$$\mathcal{D}_s(I_B, I_0) = \mu^\dagger d - \ln Z(\mu)$$

Example: Coin tossing

$I = $ "$s \in \{0,1\}$"

$P(s|I_0) = 1/2$ [Bob assumes initially a fair coin, due to Laplace principle of insufficient reason (to favour any case)]

irrespective of s

$m = \cancel{(s, \dots)}$ $\langle s \rangle_{(s|I_A)} = P(0|I_A) \cdot 0 + P(1|I_A) \cdot 1$

$= (s, d)$ $= P(1|I_A) \stackrel{!}{=} d_A$

frequency of heads

$\cancel{P(s|I_B) = P(s|I_0)\, e}$

$$Z(\mu) = \sum_{s \in \{0,1\}} P(s|I_0)\, e^{\mu s} = \frac{1}{2} e^{\mu 0} + \frac{1}{2} e^{\mu 1}$$

$$\boxed{Z(\mu) = \frac{1}{2}(1 + e^{\mu})}$$

$$\mu: \frac{\partial \ln Z}{\partial \mu} = \frac{e^{\mu}}{1 + e^{\mu}} \stackrel{!}{=} d$$

$$e^{\mu} = (1 + e^{\mu})\, d$$

$$e^{\mu}(1 - d) = d \qquad e^{\mu} = \frac{d}{1-d}$$

$$\mu = \ln \frac{d}{1-d}$$

$$\Rightarrow P(s|I_B) = \frac{1/2\, e^{\mu s}}{1/2(1 + e^{\mu})} \Bigg|_{e^{\mu} = \frac{d}{1-d}} = \frac{\left(\frac{d}{1-d}\right)^s}{1 + \frac{d}{1-d}}$$

Hard data constraint as a soft one

$$s \in \mathcal{S} = \mathbb{R}^n \qquad d \in \mathcal{D} = \mathbb{R}^m$$

$$s' = (d, s) \in \mathcal{S}' = \mathcal{S} \otimes \mathcal{D} = \mathbb{R}^{m+n}$$

message: $m = $ "$d = d^*$"

topic $f(s') = \delta(d - d^*)$

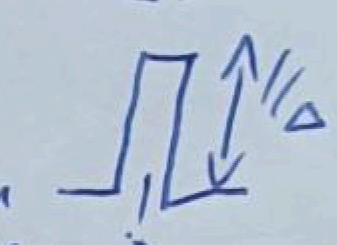
value $\tilde{d} = \langle f(s') \rangle_{(s'|I_A)} = \int ds \int dd\, \delta(d-d^*)\, P(d, s|I_A)$

$$= \int ds\, P(d^*, s|I_A)$$

$$= P(d^*|I_A)$$ ~~(not communicated)~~

$$= 1$$

$\delta(x) = \lim_{\Delta \to 0}$



$\Rightarrow e^{\mu\delta} = \lim_{\Delta\to 0} \delta \cdot \underbrace{e^{\mu\Delta}}_{g(\mu)}$

$$m = (\delta(d-d^*), 1)$$

partition function

$$Z(\mu) = \int ds'\, P(s'|I_0)\, e^{\mu f(s')} = \int ds'\, P(s'|I_0)\, e^{\mu \delta(d-d^*)}$$

$$= \int dd \int ds\, P(d,s|I_0)\, \underbrace{e^{\dots}}_{\delta(d-d^*)\cdot g(\mu)}$$

$$= \int ds\, P(d^*, s|I_0)\, g(\mu) = P(d^*|I_0) \cdot g(\mu)$$

$$\ln Z = \ln P(d^*|I_0) + \ln g(\mu)$$

$$\frac{\partial \ln Z}{\partial \mu} = 0 \Rightarrow \mu = \text{something}$$

$$P(d,s \mid I_B) = \frac{P(d,s \mid I_0)\, g(\mu)\, \delta(d-d^*)}{\int dd \int ds^* P(d,s \mid I_0)\, g(\mu)\, \delta(d-d^*)}$$

$$= \left. \frac{\delta(d-d^*)\, P(d,s \mid I_0)}{P(d \mid I_0)} \right|_{d=d^*}$$

$$\overset{\text{Bayes}}{=} \delta(d-d^*)\, P(s \mid d, I_0)$$

$$\Rightarrow I_B = (d, I_0)$$

Bayes