Running Ex: Pos Tossing
Given: - Data: $D=\left\{\left(x^{(i)}, y^{(i)}\right)\right\}_{i=1}^{N}$

$$
\begin{array}{llll}
x^{(i)}=f l_{1 i e s}^{x_{1}^{(i)}} & l_{\text {ike }}^{(i)} & x_{3}^{(i)} & x_{1}^{(i)} \\
y^{(i)} & = & y_{1}^{(i)} & V \\
y_{i}^{(i)} & y_{i}^{(i)} & y_{3}^{(i)} & y_{4}^{(i)}
\end{array}
$$

 (specific to $\vec{x}^{(i)}$ )

- Loss: $l\left(y^{*}, \hat{y}\right): Y_{\vec{x}} \times Y_{\vec{x}} \rightarrow \mathbb{R}$ possible outputs for
- Hypothesis space: If st. $\forall h \in H, h(\vec{x}) \in Y_{\vec{x}} \forall \vec{x} \in X$

Goal: Minimize Empirical Risk

$$
\begin{aligned}
& \hat{h}=\operatorname{argmin}_{h e H} \frac{1}{N} \sum_{i=1}^{N} l\left(y^{(i)}, h\left(\vec{x}^{(i)}\right)\right) \\
& \hat{\theta}=\operatorname{argmin} \frac{1}{N} \sum_{i=1}^{N} l\left(y^{(i)}, h_{\theta}\left(\vec{x}^{(i)}\right)\right)
\end{aligned}
$$

Structured Prediction as Search

$X_{1: T}=$| $x_{1}$ | $x_{2}$ | $x_{3}$ | $x_{4}$ | flies like |
| :---: | :---: | :---: | :---: | :---: | | a |
| :--- |$\quad$ plant $\quad$ Hamming Loss

Of $x_{1: T}=$ Induced search space for $\quad \vec{x}$ is $: \quad \mu\left(y_{1: T}^{*}, \hat{y}_{1: T}\right)=\sum_{t=1}^{T} 1\left(y_{t}^{*} \neq \hat{y}_{t}\right)$



Def: a trajectory is a path through the search space which is a sequence of output labels $\hat{y}_{1: T}$
Def: each state $s_{t} \in V_{x}$ corresponds to a partial trajectory starting at initial state $s_{0} \quad\left[\begin{array}{l}\hat{y}_{1: t} \\ \text { for } \\ 1 \leq t \leq T\end{array}\right.$
Def: a policy mops from (observation sequence $X_{1: T}$ and a partial trajectory $y 1: t)$ a state sett a a next state $s_{t+1}$ (i.e. next $l_{a b l} y_{t+1}$ )
Def: a action $a_{t}$ is a next label $y_{t}$

