

# Jeroen's model

June 4, 2013

## 1 Notation

- $\Sigma$ : partially-ordered list of symbols (types).
- $\Sigma_i$ :  $i$ th symbol in partial ordering.
- $T$ : number of symbols in grammar.
- $\mathcal{P}$ : pose space.
- $\mathcal{B} = (b_1, \dots, b_N)$ : set of bricks.
- $\mathcal{P}(b) \in \mathcal{P}$ : set of poses associated with brick  $b$ .
- $\mathcal{A} = (a_1, \dots, a_M)$ ,  $\mathcal{A} \subseteq \mathcal{B}$ : set of active bricks.
- $\mathcal{A}_m = (a_1, \dots, a_m)$ ,  $m \leq M$   $\mathcal{A}_m \subseteq \mathcal{A}$ : subset of active bricks.
- $t(b)$ : type of brick  $b$ .
- $\#(C)$ : the cardinality (size) of some set  $C$ . *E.g.*  $\#(\mathcal{A}) = M$ .
- $\mathcal{R}$ : set of production rules of the form  $A \rightarrow B_1, \dots, B_k$  with  $r = A, B_1, \dots, B_k \in \Sigma$ .  $\mathcal{R}$  is acyclic.
- $n(r)$ : number of slots/children associated with rule  $r \in \mathcal{R}$
- $c(b)$ : the set of production rules with  $t(b)$  in the LHS.

- $\vec{\theta}_b \in \mathbb{R}^{\#(c(b))}$ : distribution over rules with  $t(b) \in \mathcal{B}$  in the LHS.  
conditional probability density over pose for  $B_i$  given pose for  $A$  for this rule and slot.
- $s_b \in \{0, 1\}$ : on/off state of brick  $b$ .
- $r_b \in \mathcal{R}(t(b))$ : rule used by brick  $b$ .
- $\mathbf{g}_b \in (\mathcal{B} \cup \perp)^{c(b)}$ : array of pointers, where each element can either point to a brick, or null (no child).
- $\mathbf{g}_b^{\mathcal{A}_m} \in (\mathcal{A}_m \cup \emptyset)^{c(b)}$ : array of pointers, where each element can either point to a brick in the active set defined by  $\mathcal{A}_m$ , or blank (child not yet specified) denoted by  $\emptyset$ . Note that blank is **different** than null (no child) in that blank means “there could be a child in this slot, but we don’t know if there is or which one yet” while null (no child) means “there definitely isn’t a child in this slot”.
- $\text{comp}(\mathbf{g}^{\mathcal{A}_m})$ : the elements of  $\mathbf{g}^{\mathcal{A}_m}$  for which  $g_{b,k}^{\mathcal{A}_m} = \emptyset$ . Note: we use “comp” as the identifier for this concept since these are the slots that may receive a child in the future; these are the slots to be “completed”.
- $V(b)$ : “before” set of bricks, which we define as  $V(b) = \{b' : t(b') < t(b)\}$  where  $<$  means “comes before in the partial ordering defined by  $\Sigma$ ”.
- $W(b)$ : “after” set of bricks, which we define as  $W(b) = \{b' : t(b') > t(b)\}$  where  $>$  means “comes after in the partial ordering defined by  $\Sigma$ ”.
- $\vec{\phi}_{b,k,r} \in \mathbb{R}^{\#(W(b))+1}$ : distribution over bricks to which brick  $b$ ’s  $k$  – th slot may point if brick  $b$  is using rule  $r$ . Note that brick  $b$  may only point to bricks in its “after” set. Note that the  $+1$  is present since a brick/slot may point to nothing, denoted by the symbol  $\perp$ . *edit*: this is a bit sloppy. Each slot has a particular type, and only bricks of that type may go in this slot. The set of possibilities for each slot, in general, is a subset of the “after” set, and not the entire after set.
- $q_b \in \mathcal{P}(b)$ : pose for brick  $b$ .  $q_b = \perp$  is a special “nothing” pose that does not contribute to image evidence.

- $\vec{\pi} \in \mathbb{R}^T$ : vector of mixing coefficients for the templates of each type of brick.
- $\mathbf{X} = \{\mathbf{s}, \mathbf{r}, \mathbf{g}, \mathbf{q}\}$ : collection of hidden variables of model.
- $\mathbf{X}_{-r}^{\mathcal{A}_m} = \{s_a, g_a, q_a\}, \forall a \in \mathcal{A}_m$ : collection of hidden variables of model, with rules removed, for the set of bricks in the active set.
- $I$ : the image.

## 2 Model

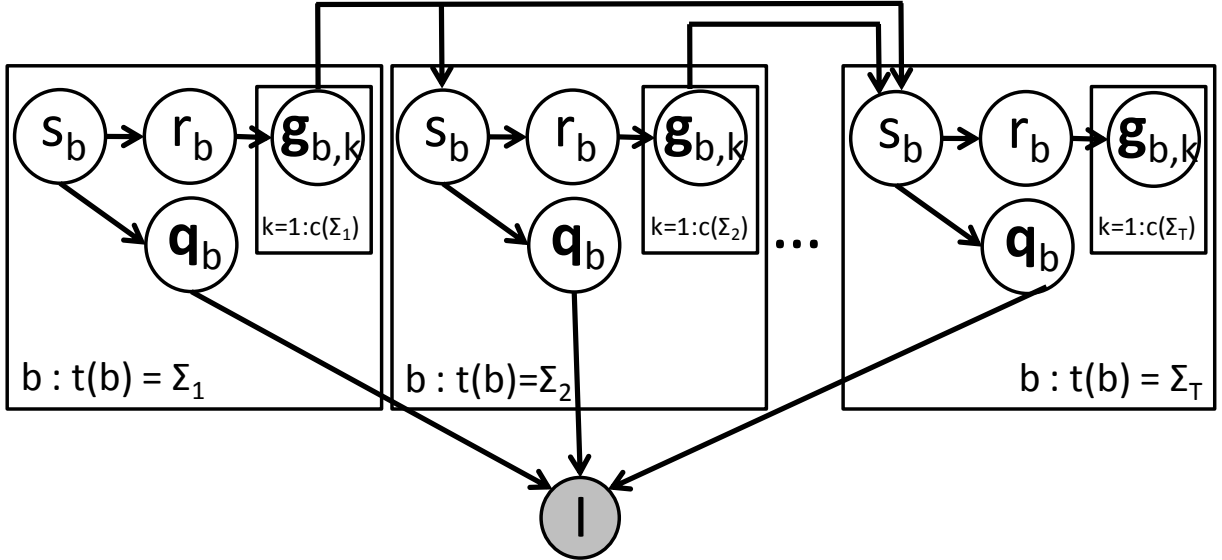


Figure 1: Graphical model

The decomposition of the joint probability is given by the graphical model in Fig. 1 and by:

$$\begin{aligned}
P(\mathbf{X}, I) &= \left( \prod_{b \in \mathcal{B}} P(s_b \mid \mathbf{g}_{V(b)}) \right) \\
&\quad \left( \prod_{b \in \mathcal{B}} P(r_b \mid s_b) \prod_{k=1}^{c(b)} P(g_{b,k} \mid r_b) \right) \\
&\quad \left( \prod_{b \in \mathcal{B}} P(\mathbf{q}_b \mid s_b) \right) \\
&\quad P(I \mid \mathbf{q})
\end{aligned} \tag{1}$$

where we define the following pieces:

$$\begin{aligned}
P(s_b = 1 \mid \mathbf{g}_{V(b)}) &= \begin{cases} 1 & \text{if } \exists \{a, k\} \text{ s.t. } g_{a,k} = b, g_{a,k} \in g_{V(b)} \\ \epsilon & \text{if } \nexists \{a, k\} \text{ s.t. } g_{a,k} = b, g_{a,k} \in g_{V(b)} \end{cases} \\
P(r_b \mid s_b) &= \text{Discrete}(\vec{\theta}_b) \\
P(g_{b,k} \mid r_b) &= \text{Discrete}(\vec{\phi}_{b,k,r_b}) \\
P(\mathbf{q}_b = q \mid s_b) &= \begin{cases} \frac{1}{\#\mathcal{P}(b)} & \text{if } s_b = 1 \text{ and } q \neq \perp, q \in \mathcal{P}(b) \\ 0 & \text{if } s_b = 1 \text{ and } q = \perp \\ 0 & \text{if } s_b = 0 \text{ and } q \neq \perp \\ 1 & \text{if } s_b = 0 \text{ and } q = \perp \end{cases} \\
P(I \mid \mathbf{q}) &= \prod_p P(I_p \mid \mathbf{q}) \\
P(I_p \mid \mathbf{q}) &\propto \sum_b \pi_{t(b)} P(I_p \mid q_b) \delta(q_b \text{ overlaps with } p) \\
P(I_p \mid q_b) &= \text{defn on where pixel } p \text{ falls in template } t(b) \text{ here, along with statement of prob. being Bernoulli}
\end{aligned}$$

where  $\delta(\cdot)$  is an indicator variable.

## Localization

We define localizations in terms of the components  $\left( \prod_{b \in \mathcal{B}} P(s_b \mid \mathbf{g}_{V(b)}) \right)$ ,  $\left( \prod_{b \in \mathcal{B}} P(r_b \mid s_b) \prod_{k=1}^{c(b)} P(g_{b,k} \mid r_b) \right)$ ,  $\left( \prod_{b \in \mathcal{B}} P(\mathbf{q}_b \mid s_b) \right)$ , and  $P(I \mid \mathbf{q})$  separately.

$$\left( \prod_{b \in \mathcal{B}} P(s_b \mid \mathbf{g}_{V(b)}) \right) \Rightarrow \left( \prod_{a \in \mathcal{A}_m} P^{\mathcal{A}_m}(s_a \mid \mathbf{g}_{V(b)}^{\mathcal{A}_m}) \right) \tag{9}$$

where  $\Rightarrow$  is used to mean “localizes to”, and we evaluate  $P^{\mathcal{A}_m}(s_a \mid \mathbf{g}_{V(b)}^{\mathcal{A}_m})$  in a similar fashion to  $P(s_b \mid \mathbf{g}_{V(b)})$  outlined above. We note that  $P^{\mathcal{A}_m}(s_a \mid \mathbf{g}_{V(b)}^{\mathcal{A}_m}) \neq P(s_a \mid \mathbf{g}_{V(b)}^{\mathcal{A}_m})$  since we do not perform the “correct” marginalizations.

Similarly,

$$\left( \prod_{b \in \mathcal{B}} P(r_b \mid s_b) \prod_{k=1}^{c(b)} P(g_{b,k} \mid r_b) \right) \Rightarrow \left( \prod_{a \in \mathcal{A}_m} P(r_a \mid s_a) \prod_{k=1}^{c(b)} P^{\mathcal{A}_m}(g_{a,k} \mid r_a) \right) \quad (10)$$

$$P^{\mathcal{A}_m}(g_{a,k} \mid r_a) = \begin{cases} 1 - \sum_{a' \in \mathcal{A}_m} P(g_{a,k} = a' \mid r_a) & \text{if } g_{a,k} \neq \emptyset \\ P(g_{a,k} \mid r_a) & , \text{ otherwise} \end{cases} \quad (11)$$

Note that the RHS of Eqn. (11) is defined in terms of non-localized distributions,  $P(g_{a,k} \mid r_a)$ .

$$\left( \prod_{b \in \mathcal{B}} P(\mathbf{q}_b \mid s_b) \right) \Rightarrow \left( \prod_{a \in \mathcal{A}_m} P(\mathbf{q}_a \mid s_a) \right) \quad (12)$$

$$P(I \mid \mathbf{q}) \Rightarrow P^{\mathcal{A}_m}(I \mid \mathbf{q}^{\mathcal{A}_m}) \quad (13)$$

where we evaluate Eqn. 13 in a similar fashion to  $P(I \mid \mathbf{q})$ . We note that  $P^{\mathcal{A}_m}(I \mid \mathbf{q}^{\mathcal{A}_m}) \neq P(I \mid \mathbf{q}^{\mathcal{A}_m})$  since we do not perform the correct marginalizations.

We define the overall localized distribution,  $P^{\mathcal{A}_m}(X^{\mathcal{A}_m}, I)$ , in terms of the previously defined localization distributions:

$$\begin{aligned} P^{\mathcal{A}_m}(X^{\mathcal{A}_m}, I) &= \left( \prod_{a \in \mathcal{A}_m} P^{\mathcal{A}_m}(s_a \mid \mathbf{g}_{V(b)}^{\mathcal{A}_m}) \right) \\ &\quad \left( \prod_{a \in \mathcal{A}_m} P(r_a \mid s_a) \prod_{k=1}^{c(b)} P^{\mathcal{A}_m}(g_{a,k} \mid r_a) \right) \\ &\quad \prod_{a \in \mathcal{A}_m} P(\mathbf{q}_a \mid s_a) \\ &\quad P^{\mathcal{A}_m}(I \mid \mathbf{q}^{\mathcal{A}_m}) \end{aligned} \quad (14)$$

## State representation

Representing the configuration of an image for a **non-localized** model requires specifying  $s_b$ ,  $g_b$ ,  $r_b$ ,  $q_b$ ,  $\forall b \in \mathcal{B}$ . For a **localized** model, the same random variables could be used as the state

representation. However, requiring specification of the  $r_b$ 's is unnecessary and too restrictive; we may not be able to well predict the rule a brick has used by examining the set of active bricks. This is especially true if the active set only contains one brick! Instead, we represent the state representation of the localized model by the random variables  $\mathbf{X}_{-r}^{\mathcal{A}_m} = \{s_a, g_a, q_a\}, \forall a \in \mathcal{A}$ , and we marginalize over  $r_a$  wherever it is used.

## Inference

Inference proceeds by iteratively adding one brick at a time to the active set. At any iteration,  $m$ , our goal is to maintain a representation of the localized posterior distribution  $P^{\mathcal{A}_m}(\mathbf{X}_{-r}^{\mathcal{A}_m} | I)$ . To maintain this representation, we employ a particle filtering approach. We treat samples from the previous localized posterior distribution  $P^{\mathcal{A}_{m-1}}(\mathbf{X}_{-r}^{\mathcal{A}_{m-1}} | I)$  as draws from a proposal distribution, and re-weight these particles accordingly (discuss re-weighting and evaluating marginals later).

Suppose we have sampled a previous state  $\mathbf{X}_{-r}^{\mathcal{A}_{m-1}}$ . Our goal is draw a sample from the conditional  $P(\mathbf{X}_{-r}^{\mathcal{A}_m} | \mathbf{X}_{-r}^{\mathcal{A}_{m-1}}, I)$ . It will turn out that we cannot directly draw from this distribution exactly, so we sample from an approximation of  $P(\mathbf{X}_{-r}^{\mathcal{A}_m} | \mathbf{X}_{-r}^{\mathcal{A}_{m-1}}, I)$ . Suppose  $\mathcal{A}_m = \mathcal{A}_{m-1} \cup a_b$ . That is, we obtain  $\mathcal{A}_m$  from  $\mathcal{A}_{m-1}$  by adding brick  $b$  to  $\mathcal{A}_{m-1}$ .

Drawing a(n approximate) sample requires us to sample values for:  $s_b, q_b, \mathbf{g}_b$ , and possible "completions" for  $\text{comp}(g^{\mathcal{A}_{m-1}})$ , which we denoted previously as  $\text{comp}(g^{\mathcal{A}_{m-1}})$ . We note that for  $g \in \text{comp}(\mathbf{g}^{\mathcal{A}_m})$ , the only possible values that can be sampled for  $g$  are  $g = \emptyset, g = b$ . That is, each not-yet-assigned brick slot choose either to point to the newly-added active brick  $b$ , or remain empty.

We first note that for any setting of the random variables outlined above, we may classify it as one of three events: 1) brick  $b$  is off, 2) brick  $b$  is on, and has no parents, and 3) brick  $b$  is on, and has at least one parent. We compute the (approximate) mass associated with each of these three events.

### Event 1: brick $b$ is off

For this event, there is only one setting of the random variables that has non-zero probability:  $s_b = 0, q_b = \perp, g_{b,k} = \perp \forall k, g = \emptyset \forall g \in \text{comp}(\mathbf{g}^{\mathcal{A}_m})$ . Note that  $g^{\mathcal{A}_m} \neq g^{\mathcal{A}_{m-1}}$  only because  $g^{\mathcal{A}_{m-1}}$  specifies the (lack of) children for the brick  $b$ , which is not specified in  $g^{\mathcal{A}_{m-1}}$ . The other

corresponding entries of  $g^{\mathcal{A}_m}$  and  $g^{\mathcal{A}_{m-1}}$  are indeed equal, however. It is trivial to evaluate Eqn. 14 with these settings of the random variables.