# A Non-local Coarsening Result in Granular Probabilistic Networks

Cory J. Butz, Hong Yao, and Howard J. Hamilton

Department of Computer Science, University of Regina,
Regina, Canada, S4S 0A2, {butz,yao2hong,hamilton}@cs.uregina.ca

**Abstract.** In our earlier works, we coined the phrase *granular probabilistic reasoning* and showed a *local* coarsening result. In this paper, we present a *non-local* method for coarsening variables (i.e., the variables are spread throughout the network) and establish its correctness.

## 1 Introduction

In [3], we coined the phrase *granular probabilistic reasoning* to mean the ability to coarsen and refine parts of a probabilistic network depending on whether they are of interest or not. Granular probabilistic reasoning is of importance as it not only leads to more efficient probabilistic inference, but it also facilitates the design of large probabilistic networks [4]. It is then not surprising that Xiang [4] explicitly states that our granular probabilistic reasoning [3] demands further attention.

We proposed two operators called *nest* and *unnest* for coarsening and refining variables in a network, respectively [3]. In [1], we showed that the nest operator can be applied *locally* to a marginal distribution with the same effect as if it were applied directly to the joint distribution. However, no study has ever addressed how to coarsen variables spread throughout a network.

In this paper, we present a method, called *Non-local Nest*, for coarsening *non-local variables*, that is, variables spread throughout a network. This method gathers all variables to be coarsened into one marginal distribution, and then applies the nest operator. We also prove our method is correct.

This paper is organized as follows. Section 2 reviews a local nest method. We present a non-local method for nesting variables in Section 3. The conclusion is presented in Section 4.

## 2 A Local Method for Nesting

Consider the joint distribution $p(R)$ represented as a probabilistic relation $\mathbf{r}(R)$ in Fig. 1, where $R = \{A, B, C, D, E, F\} = ABCDEF$ is a set of variables. Configurations with zero probability are not shown. The *nest* operator $\phi$ is used to *coarsen* a relation $\mathbf{r}(XY)$. Intuitively, $\phi_{A=Y}(\mathbf{r})$ groups together all the Y-values into a nested distribution for coarse variable $A$ given the same X-value.

G. Wang et al. (Eds.): RSFDGrC 2003, LNAI 2639, pp. 686–689, 2003.

More formally,

$$\phi_{A=Y}(\mathbf{r})$$
$$= \{t \mid t(X) = u(X),\ t(A) = \{u(Yp(R))\},\ t(p(XA)) = \sum_u u(p(R)),\ \text{and } u \in \mathbf{r}\}.$$

Attribute $p(R)$ in the A-value is relabeled $p(Y)$ and the values are normalized.

$\mathbf{r}(R) =$

| $A$ | $B$ | $C$ | $D$ | $E$ | $F$ | $p(R)$ |
|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 | 0 | 0.05 |
| 0 | 0 | 0 | 0 | 1 | 0 | 0.05 |
| 0 | 0 | 1 | 0 | 1 | 1 | 0.20 |
| 0 | 1 | 0 | 1 | 0 | 2 | 0.15 |
| 0 | 1 | 0 | 1 | 1 | 2 | 0.15 |
| 1 | 0 | 0 | 2 | 2 | 0 | 0.40 |

**Fig. 1.** A probabilistic relation $\mathbf{r}(R)$ representing a joint distribution $p(R)$.

*Example 1.* Recall the relation $r(ABCDEF)$ in Fig. 1. Nesting variables $Y = DE$ as the single variable $G$ gives the nested relation $\phi_{G=\{D,E\}}(\mathbf{r})$ in Fig. 2. For instance, given the fixed X-value $\langle A : 0, B : 1, C : 0, F : 2\rangle$, the Y-values $\langle D : 1, E : 0, p(R) : 0.15\rangle$ and $\langle D : 1, E : 1, p(R) : 0.15\rangle$ are grouped into a nested distribution. Here the attribute $p(R)$ is relabeled as $p(DE)$, and the probability values 0.15 and 0.15 are normalized as 0.50 and 0.50.

In practice, a joint distribution $\mathbf{r}(R)$ is represented as a *Markov network* (MN) [2]. The dependency structure of a MN is an *acyclic hypergraph* (a *join-tree*) [2]. The acyclic hypergraph encodes *conditional independencies* [2] satisfied by $\mathbf{r}(R)$. For example, the joint distribution $p(R)$ in Fig. 1 can be expressed as the MN:

$$p(R) = \frac{p(ABD) \cdot p(ABC) \cdot p(ACE) \cdot p(BCF)}{p(AB) \cdot p(AC) \cdot p(BC)}, \tag{1}$$

where $\mathcal{R} = \{R_1 = \{A, B, D\}, R_2 = \{A, B, C\},\ R_3 = \{A, C, E\}, R_4 = \{B, C, F\}\}$ is an acyclic hypergraph, and the marginal distributions $\mathbf{r}_1(ABD)$, $\mathbf{r}_2(ABC)$, $\mathbf{r}_3(ACE)$, and $\mathbf{r}_4(BCF)$ of $\mathbf{r}(R)$ are shown in Fig. 3.

In our probabilistic relational model [2], the MN in Eq. (1) is expressed as

$$\mathbf{r}(R) = ((\mathbf{r}_1(ABD) \otimes \mathbf{r}_2(ABC)) \otimes \mathbf{r}_3(ACE)) \otimes \mathbf{r}_4(BCF),$$

where the *Markov join operator* $\otimes$ means

$$\mathbf{r}(XY) \otimes \mathbf{r}(YZ) = \frac{p(XY) \cdot p(YZ)}{p(Y)}.$$

$\phi_{G=\{D,E\}}(\mathbf{r}) =$

| A | B | C | G | F | $p(ABCGF)$ |
|---|---|---|---|---|---|
| 0 | 0 | 0 | D E $p(DE)$<br>0 0 0.5<br>0 1 0.5 | 0 | 0.1 |
| 0 | 0 | 1 | D E $p(DE)$<br>0 1 1.0 | 1 | 0.2 |
| 0 | 1 | 0 | D E $p(DE)$<br>1 0 0.5<br>1 1 0.5 | 2 | 0.3 |
| 1 | 0 | 0 | D E $p(DE)$<br>2 2 1.0 | 0 | 0.4 |

**Fig. 2.** The nested relation $\phi_{G=\{D,E\}}(\mathbf{r})$, where $\mathbf{r}$ is the relation in Fig. 1.

$\mathbf{r}_1 =$

| A | B | D | $p(R_1)$ |
|---|---|---|---|
| 0 | 0 | 0 | 0.3 |
| 0 | 1 | 1 | 0.3 |
| 1 | 0 | 2 | 0.4 |

$\mathbf{r}_2 =$

| A | B | C | $p(R_2)$ |
|---|---|---|---|
| 0 | 0 | 0 | 0.1 |
| 0 | 0 | 1 | 0.2 |
| 0 | 1 | 0 | 0.3 |
| 1 | 0 | 0 | 0.4 |

$\mathbf{r}_3 =$

| A | C | E | $p(R_3)$ |
|---|---|---|---|
| 0 | 0 | 0 | 0.2 |
| 0 | 0 | 1 | 0.2 |
| 0 | 1 | 1 | 0.2 |
| 1 | 0 | 2 | 0.4 |

$\mathbf{r}_4 =$

| B | C | F | $p(R_4)$ |
|---|---|---|---|
| 0 | 0 | 0 | 0.5 |
| 0 | 1 | 1 | 0.2 |
| 1 | 0 | 2 | 0.3 |

**Fig. 3.** The marginals $\mathbf{r}_1(ABD)$, $\mathbf{r}_2(ABC)$, $\mathbf{r}_3(ACE)$, and $\mathbf{r}_4(BCF)$ of relation $\mathbf{r}$.

We may omit the parentheses for simplified notation. For example, the Markov join $\mathbf{r}_1(ABD) \otimes \mathbf{r}_2(ABC) \otimes \mathbf{r}_3(ACE)$ is shown in Fig. 4.

The main result in [1] was that the nest operator can be applied *locally* to one marginal distribution in a Markov network with the same effect as if applied directly to the joint distribution itself. For example, $\phi_{G=\{E\}}(\mathbf{r})$ is the same nested distribution as $\mathbf{r}_1(ABD) \otimes \mathbf{r}_2(ABC) \otimes \phi_{G=\{E\}}(\mathbf{r}_3(ACE)) \otimes \mathbf{r}_4(BCF)$. We next study how to coarsen variables spread throughout a network.

## 3 A Non-local Method for Nesting

We call $Y \subseteq R$ a *nestable set* with respect to a MN on $\mathcal{R} = \{R_1, \ldots, R_n\}$, if $Y$ does not intersect any *separating set* [2] of $\mathcal{R}$. Since the nest operator is unary, the first task is to combine the nestable set $Y$ of variables into a single table. The well-known relational database *selective reduction algorithm* (SRA) is applied for this purpose. The nest operator can then be applied to coarsen $Y$ as attribute $A$. We now present the formal algorithm *Non-local Nest* (NLN) to coarsen a nestable set $Y$ as attribute $A$ in a given MN on acyclic hypergraph $\mathcal{R}$.

$\mathbf{r}_1(ABD) \otimes \mathbf{r}_2(ABC) \otimes \mathbf{r}_3(ACE) =$

| $A$ | $B$ | $C$ | $D$ | $E$ | $p(ABCDE)$ |
|---|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 | 0.05 |
| 0 | 0 | 0 | 0 | 1 | 0.05 |
| 0 | 0 | 1 | 0 | 1 | 0.20 |
| 0 | 1 | 0 | 1 | 0 | 0.15 |
| 0 | 1 | 0 | 1 | 1 | 0.15 |
| 1 | 0 | 0 | 2 | 2 | 0.40 |

**Fig. 4.** The Markov join $\mathbf{r}_1(ABD) \otimes \mathbf{r}_2(ABC) \otimes \mathbf{r}_3(ACE)$.

**Algorithm 1** NLN$(Y, A, \mathcal{R})$
1. Let $R_j, \ldots, R_n$ be those elements of $\mathcal{R}$ *not* deleted by the call SRA$(Y, \mathcal{R})$.
2. Return $\phi_{A=Y}(\mathbf{r}_{jn})$, where $\mathbf{r}_{jn} = \mathbf{r}_j(R_j) \otimes \ldots \otimes \mathbf{r}_n(R_n)$.

**Theorem 1.** Let $\mathbf{r}(R) = \mathbf{r}_1(R_1) \otimes \ldots \otimes \mathbf{r}_i(R_i) \otimes \mathbf{r}_j(R_j) \ldots \otimes \mathbf{r}_n(R_n)$ be represented as a MN, and let $Y$ be a nestable set. Then $\phi_{A=Y}(\mathbf{r})$ is the same as $\mathbf{r}_1(R_1) \otimes \ldots \otimes \mathbf{r}_i(R_i) \otimes \mathbf{r}'$, where $\mathbf{r}'$ is the nested relation returned by the call NLN$(Y, A, \{R_1, R_2, \ldots, R_n\})$.

*Example 2.* $DE$ is a nestable set with respect to $\mathcal{R} = \{R_1, R_2, R_3, R_4\}$ as used in our running example. Suppose we wish to compute NLN$(DE, G, \mathcal{R})$. Then $SRA(DE, \mathcal{R}) = \{R_1, R_2, R_3\}$. Again, the Markov join $\mathbf{r}_1 \otimes \mathbf{r}_2 \otimes \mathbf{r}_3$ is shown in Fig. 4. The reader can verify that the nested relation $\phi_{G=\{D,E\}}(\mathbf{r})$ in Fig. 2 is the same as $\mathbf{r}_4(BCF) \otimes \phi_{G=\{D,E\}}(\mathbf{r}_1(ABD) \otimes \mathbf{r}_2(ABC) \otimes \mathbf{r}_3(ACE))$.

## 4 Conclusion

Xiang [4] explicitly states that more work needs to be done on our *granular* probabilistic networks [3]. In this paper, we have extended the work in [1] by coarsening a *non-local* set of variables (i.e., variables spread throughout a network). Theorem 1 establishes the *correctness* of our approach.

## References

1. Butz, C.J., Wong, S.K.M.: A local nest property in granular probabilistic networks, *Proc. of the Fifth Joint Conf. on Information Sciences*, **1**, 158–161, 2000.
2. Wong, S.K.M., Butz, C.J.: Constructing the dependency structure of a multi-agent probabilistic network. IEEE Trans. Knowl. Data Eng. **13**(3) (2001) 395–415.
3. Wong, S.K.M., Butz, C.J.: Contextual weak independence in Bayesian networks, *Proc. of the Fifteenth Conf. on Uncertainty in Artificial Intelligence*, 670–679, 1999.
4. Xiang, Y.: *Probabilistic Reasoning in Multi-Agent Systems: A Graphical Models Approach.* Cambridge Publishers, 2002.