A. Proofs

Proposition 1 Suppose $dsep_G(\mathbf{X}, \mathbf{Z}, \mathbf{Y})$ holds but $dsep_{G^\star}(\mathbf{X}^\star, \mathbf{Z}^\star, \mathbf{Y}^\star)$ does not hold. There must exist a path in $G^\star$ that connects $\mathbf{X}^\star$ and $\mathbf{Y}^\star$ but is not blocked by $\mathbf{Z}^\star$. This path also connects $\mathbf{X}$ and $\mathbf{Y}$ in G. Moreover, it is not blocked by $\mathbf{Z}$ since no variable in $\mathbf{Z} \setminus \mathbf{Z}^\star$ can be on this path. This is a contradiction with $dsep_G(\mathbf{X}, \mathbf{Z}, \mathbf{Y})$. Hence $dsep_{G^\star}(\mathbf{X}^\star, \mathbf{Z}^\star, \mathbf{Y}^\star)$ holds. □

Proposition 2 Suppose $dsep_G(X, \mathbf{Z}, \mathbf{Y})$ holds but $dsep_G(\mathbf{U} \setminus \mathbf{Z}, \mathbf{Z}, \mathbf{Y})$ does not hold. We have $\mathbf{U} \cap \mathbf{Y} = \emptyset$ by $dsep_G(X, \mathbf{Z}, \mathbf{Y})$. Moreover, there must exist a path connecting some $U \in \mathbf{U} \setminus \mathbf{Z}$ and $\mathbf{Y}$ that is not blocked by $\mathbf{Z}$. If X is on this path, then we have a path connecting X and $\mathbf{Y}$ that is not blocked by $\mathbf{Z}$. Otherwise, augmenting this path with the edge $X \leftarrow U$ leads to a path with the same properties. Either case contradicts $dsep_G(X, \mathbf{Z}, \mathbf{Y})$. □

Proposition 3 Suppose $dsep_G(\mathbf{X}, \mathbf{Z}, \mathbf{Y})$. A leaf node outside $\mathbf{X} \cup \mathbf{Y} \cup \mathbf{Z}$ is irrelevant to $P(\mathbf{x}|\mathbf{zy})$, so repeatedly remove all such leaf nodes. Prune edges outgoing from nodes $\mathbf{Z}$ as this does not change the value of $P(\mathbf{x}|\mathbf{zy})$ either. Nodes $\mathbf{X}$ are now disconnected from $\mathbf{Y}$. If $dsep_G(T, \mathbf{Z}, \mathbf{Y})$ does not hold, then node T is connected to $\mathbf{Y}$ and disconnected from $\mathbf{X}$ so $P(\mathbf{x}|\mathbf{zy})$ cannot depend on the CPT of T. Hence, if $P(\mathbf{x}|\mathbf{zy})$ depends on the CPT of T, then $dsep_G(T, \mathbf{Z}, \mathbf{Y})$. □

The next proof uses the following corollary of Proposition 1.

Corollary 1 *If $dsep_G(\mathbf{X}, \mathbf{Z}, \mathbf{Y})$, G is a proper superset of G_i and G_i is a proper superset of G_j, then $dsep_{G_i}(\mathbf{X}_j, \mathbf{Z}_j, \mathbf{Y}_j)$, where $\mathbf{X}_j$, $\mathbf{Y}_j$, $\mathbf{Z}_j$ are the subsets of $\mathbf{X}$, $\mathbf{Y}$, $\mathbf{Z}$ in DAG G_j.*

Theorem 1 The proof is by induction on DAGs $G_1, \ldots, G_{n+1}$ in Definition 1. We will show: if $dsep_{G_i}(X_k, \mathbf{Z}_i, \mathbf{Y}_i)$, the CPT for node X_k is independent of $\mathbf{Y}_i$ given $\mathbf{Z}_i$. Here, $\mathbf{Y}_i$ and $\mathbf{Z}_i$ are projections of $\mathbf{Y}$ and $\mathbf{Z}$ on some proper subset of G_i. The theorem statement is for $i = n + 1$.

This holds trivially for G_1 (empty). Consider G_i for $i > 1$ and assume this holds for G_j, where $j < i$. Suppose $dsep_{G_i}(X_k, \mathbf{Z}_i, \mathbf{Y}_i)$ for $k < i$. Then G_k is a proper subset of G_i and contains parents $\mathbf{U}_k$ of node X_k. We have $dsep_{G_i}(\mathbf{U}_k \setminus \mathbf{Z}_i, \mathbf{Z}_i, \mathbf{Y}_i)$ by Proposition 2 and $dsep_{G_k}(\mathbf{U}_k \setminus \mathbf{Z}_k, \mathbf{Z}_k, \mathbf{Y}_k)$ by Corollary 1. The CPT of node X_k is selected in G_k based on $P_k(\mathbf{U}_k|\mathbf{z}_k\mathbf{y}_k)$. Consider node X_m in G_k ($m < k$). By Proposition 3, if the CPT of node X_m is relevant to $P_k(\mathbf{U}_k|\mathbf{z}_k\mathbf{y}_k)$, then $dsep_{G_k}(X_m, \mathbf{Z}_k, \mathbf{Y}_k)$. By the induction hypothesis, this CPT is independent of $\mathbf{Y}_k$ given $\mathbf{Z}_k$. That is, every CPT in G_k that is relevant to $P_k(\mathbf{U}_k|\mathbf{z}_k\mathbf{y}_k)$ is independent of $\mathbf{Y}_k$ given $\mathbf{Z}_k$. Moreover, $P_k(\mathbf{U}_k|\mathbf{z}_k\mathbf{y}_k) = P_k(\mathbf{U}_k|\mathbf{z}_k)$ since $dsep_{G_k}(\mathbf{U}_k \setminus \mathbf{Z}_k, \mathbf{Z}_k, \mathbf{Y}_k)$. Hence, the CPT of node X_k is independent of $\mathbf{Y}_k$ given $\mathbf{Z}_k$ and of $\mathbf{Y}_i$ given $\mathbf{Z}_i$. □

Theorem 2 Assume $dsep_G(\mathbf{X}, \mathbf{Z}, \mathbf{Y})$. We show $Q(\mathbf{x}||\mathbf{yz}) = Q(\mathbf{x}||\mathbf{z})$, which reduces to $P^{\mathbf{zy}}(\mathbf{x}|\mathbf{zy}) = P^{\mathbf{z}}(\mathbf{x}|\mathbf{z})$. By Proposition 3, if $P^{\mathbf{zy}}(\mathbf{x}|\mathbf{zy})$ depends on the CPT of some node T, then $dsep_G(T, \mathbf{Z}, \mathbf{Y})$. By Theorem 1, this CPT is independent of $\mathbf{Y}$ given $\mathbf{Z}$. Hence, $P^{\mathbf{zy}}(\mathbf{x}|\mathbf{zy}) = P^{\mathbf{z}}(\mathbf{x}|\mathbf{zy})$. Moreover, $P^{\mathbf{z}}(\mathbf{x}|\mathbf{zy}) = P^{\mathbf{z}}(\mathbf{x}|\mathbf{z})$ since $dsep_G(\mathbf{X}, \mathbf{Z}, \mathbf{Y})$. Therefore, $P^{\mathbf{zy}}(\mathbf{x}|\mathbf{zy}) = P^{\mathbf{z}}(\mathbf{x}|\mathbf{z})$. □

Theorem 3 The TAC can simulate the AC by setting its parameters $\theta^+_{x|\mathbf{u}}$ and $\theta^-_{x|\mathbf{u}}$ to the corresponding parameter $\theta_{x|\mathbf{u}}$ in the AC. Hence, the TAC is no less expressive than the AC. Proposition 4 identifies a class of functions that cannot be represented by an AC. A function in this class is given in Section 5.3 (kidney stones), which can be represented by a corresponding TAC. Hence, the TAC is more expressive than the AC. □

Proposition 4 $P(c|a, b) > P(c|\bar{a}, b)$ and $P(c|a, \bar{b}) > P(c|\bar{a}, \bar{b})$ imply $\theta_{c|ab} > \theta_{c|\bar{a}b}$ and $\theta_{c|a\bar{b}} > \theta_{c|\bar{a}\bar{b}}$. Moreover, $P(c|a) = \theta_b\theta_{c|ab} + \theta_{\bar{b}}\theta_{c|a\bar{b}}$ and $P(c|\bar{a}) = \theta_b\theta_{c|\bar{a}b} + \theta_{\bar{b}}\theta_{c|\bar{a}\bar{b}}$. Hence, $P(c|a) - P(c|\bar{a}) = \theta_b(\theta_{c|ab} - \theta_{c|\bar{a}b}) + \theta_{\bar{b}}(\theta_{c|a\bar{b}} - \theta_{c|\bar{a}\bar{b}}) > 0$. □