Supplementary Material

Proof of Lemma 1

Proof. Consider a partial assignment x and let $j \in x^-$ s.t. $\texttt{cl}_{inc}(x^+, j, M')$ holds, meaning that $j \in \texttt{clos}_{M'}(x^+)$. As x^+ cannot be extended to a closed itemset without adding j (j being in x^-), the result follows. □

Proof of Proposition 2

Proof. Let x be a consistent partial assignment and i be a free item.
1) Assume that $\texttt{cl}_{inc}(x^+, i, M')$ holds. It follows that $i \in \texttt{clos}_{M'}(x^+)$. From Lemma 1, x^+ cannot be extended to a closed itemset w.r.t. M' with $i \in x^-$. Thus $\{0\} \notin dom(x_i)$.
2) (**proof by contradiction**) Suppose $j \in x^-$ s.t. $\texttt{cl}_{inc}(x^+ \cup \{i\}, j, M')$, which means that $j \in \texttt{clos}_{M'}(x^+ \cup \{i\})$. Since $j \in x^-$, which means that $j \notin \texttt{clos}_{M'}(x^+)$, it follows that $j \notin \texttt{clos}_{M'}(x^+ \cup \{i\})$. □

Algorithm 1 ADEQUATECLOSURE

In: $\mathcal{D}$: dataset, M' : set of measures
InOut: x : vector of Boolean variables
1: **if** $\exists i \in x^-$: $\mathrm{cl}_{inc}(x^+, i, M')$ **then return** false
2: **for** $i \in x^*$ **do**
3:     **if** $\mathrm{cl}_{inc}(x^+, i, M')$ **then** $dom(x_i) \leftarrow dom(x_i) \backslash \{0\}$
4: **end for**
5: **for** $j \in x^-$ **do**
6:     **for** $i \in x^*$ **do**
7:         **if** $\mathrm{cl}_{inc}(x^+ \cup \{i\}, j, M')$ **then** $dom(x_i) \leftarrow dom(x_i) \backslash \{1\}$
8:     **end for**
9: **end for**
10: **return** true

Algorithm 1 shows the propagator for ADEQUATECLOSURE. It takes as input the transactional dataset $\mathcal{D}$, the item variables x, and the set of measures M'. It starts by checking the consistency of the current partial assignment (Lemma 1), that is, if x^+ cannot be extended to a closed itemset w.r.t. M' without adding i ($i \in x^-$) (line 1), if so, the constraint is violated and a fail is returned. Algorithm 1 must then check items $i \in x^*$ that are required to extend x^+ to a closed pattern w.r.t. M'. Thus, we remove 0 from each free item variable $i \in x^*$ such that $\texttt{cl}_{inc}(x^+, i, M')$ holds by using rule ($\mathbf{R_1}$). Second, using rule ($\mathbf{R_2}$), we remove 1 from each free item variable $i \in x^*$ such that $x^+ \cup \{i\}$ cannot be extended to a closed pattern w.r.t. M' without adding the absent item $j \in x^-$ (lines 5-9).

Computing $m(x^+)$ depends on the type of m. For instance, computing $sup(x^+)$ requires a reversible BitSet to store the cover of x^+. Other measures such as min or max only require a reversible integer.

Proposition (Consistency, time complexity)**.** *Let $\mathcal{D}$ be a transaction database of n items and m transactions and a set of measures M', which contains c sup-based measures. Algorithm 1 enforces domain consistency in time $\mathcal{O}(n^2mc)$.*

Proof of the domain consistency of Algorithm 1

Proof. (**proof by contradiction**) Let i be an item, M' a set of preserving measures and $v \in dom(x_i)$ such that x_i is not domain consistent:

- if $v = 1$, then it means that x^+ can't be extended to a closed pattern w.r.t. M'. Such case is not possible since rule 2 is applied.
- if $v = 0$, then it means that x^+ can't be extended to a closed pattern w.r.t. M' without adding i. This will never happen because rule 1 is applied.

□

Proof of the time complexity of Algorithm 1

Proof. Let n be the number of items and m the number of transactions in the database. Computing $m(x^+)$ requires at most $O(nm)$ for each sup-based measures (such as sup or $sup_\vee$) and $O(n)$ for each attribute-based measures (such as min or max). Checking rule 1 can be done in $O(nm)$ for sup-based measures and $O(n)$ for attribute measures. Finally, checking rule 2 requires $O(n^2m)$ for sup-based measures and $O(n^2)$ for attribute measures. Therefore, if c is the number of sup-based measures in M', the worst case complexity is $O(n^2mc)$. □

Proof of Proposition 3

Proof. We have $P = X \cup Y$, X is a generator. Since $Y' = Y \cup clos^-(P)$, we have $X \cup Y' = clos_{sup}(P)$. Thus, r' is an MNR. By definition, $sup(P) = sup(clos_{sup}(P))$ and thus $sup(r) = sup(r')$. As $conf(r) = \frac{sup(r)}{sup(X)}$ and $conf(r') = \frac{sup(r')}{sup(X)}$, we get $conf(r) = conf(r')$. □

CP model for computing MNRs (solving step)

$$
\textsc{Sky-MNR}_{\mathcal{D}}(x,y,z) \equiv \begin{cases}
\forall s \in \mathcal{S}ky: \\
sky_s \Leftrightarrow (\bigwedge\limits_{i \in s} z_i) \wedge (\bigwedge\limits_{i \notin \mathtt{clos}_{sup}(s)} \neg z_i) \wedge (\bigwedge\limits_{i \in clos^-(s)} y_i) & (1) \\
\sum_{s \in \mathcal{S}ky} sky_s \geq 1 & (2) \\
\forall i \in \mathcal{I}: \neg x_i \vee \neg y_i & (3) \\
\sum_{i \in \mathcal{I}} x_i \times \sum_{i \in \mathcal{I}} y_i \geq 1 & (4) \\
\forall i \in \mathcal{I}: z_i \Leftrightarrow x_i \vee y_i & (5) \\
\textsc{Generator}_{\mathcal{D}}(x) & (6)
\end{cases}
$$

(a) Variables. There are two groups of decision variables:
- *rule variables*: x, y and z are three vectors of Boolean variables, where x_i, y_i and z_i represent the presence of item $i \in \mathcal{I}$ in the antecedent x, the consequence y and in the rule as a whole z.
- *skypattern variables*: sky is a vector of Boolean variables, where sky_s holds iff the MNR is related to the skypattern s.

(b) Constraints. Our model involves the following set of constraints:
- **Cons.(1)** is a channeling constraint ensuring that the rule associated to skypattern s satisfies $X \cup Y = \mathtt{clos}_{sup}(s)$ and $clos^-(s) \subseteq Y$.
- **Cons.(2)** associates a rule to at least one skypattern.
- **Cons.(3)** ensures that an item cannot appear in the antecedent and in the consequence of the rule simultaneously.
- **Cons.(4)** ensures that the antecedent and the consequence of the rule are not empty.
- **Cons.(5)** is a channeling constraint between z and x, y.
- **Cons.(6)** ensures that x is a generator.

Rules quality for SKY4MNR vs CP4MNR

Figures 1 and 2 plot the values of support (noted freq) vs. confidence and confidence vs. lift for both approaches for different datasets.

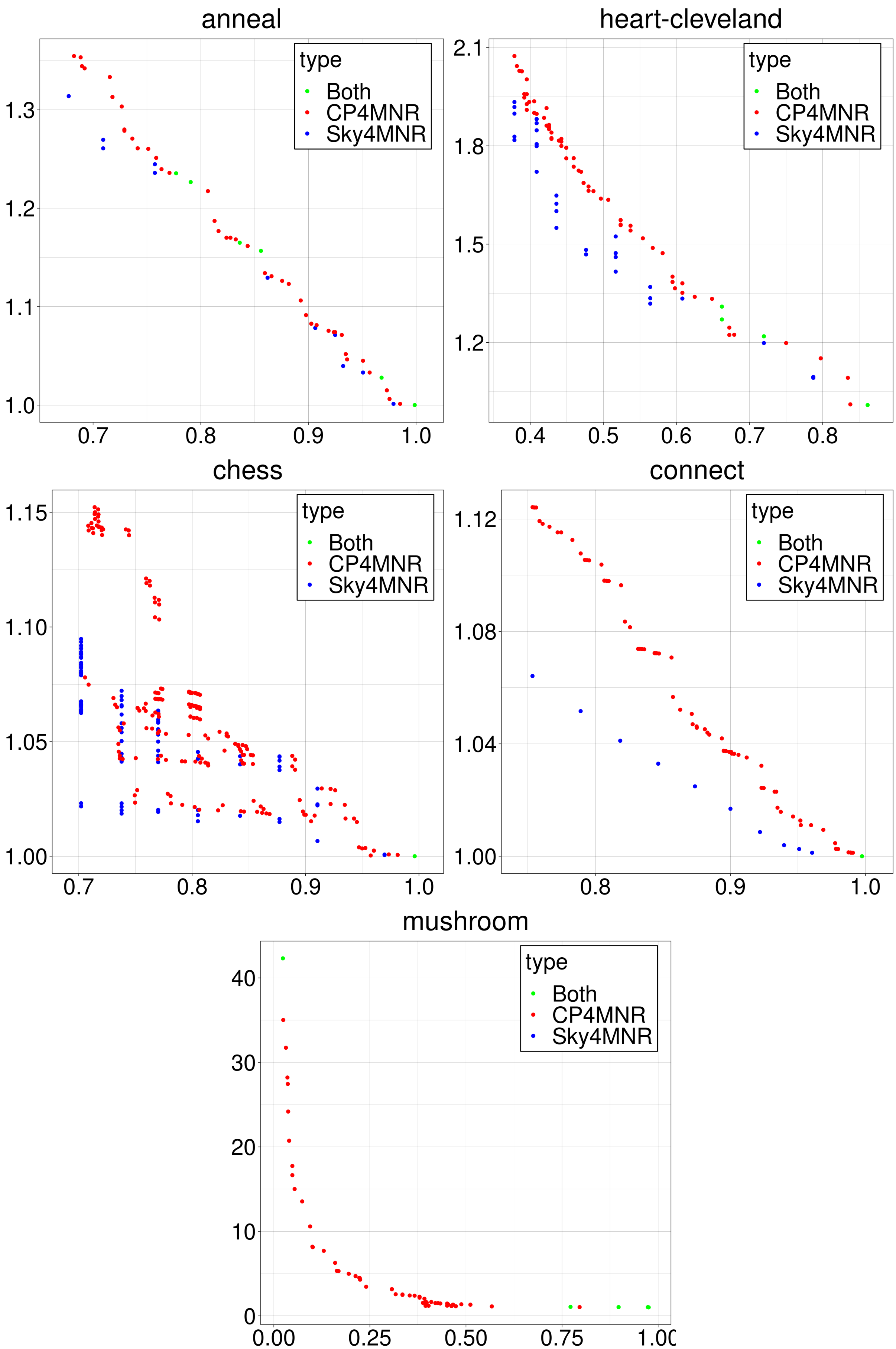

Figure 1: Qualitative analysis of MNRs (all datasets): support (x-axis) vs. lift (y-axis).

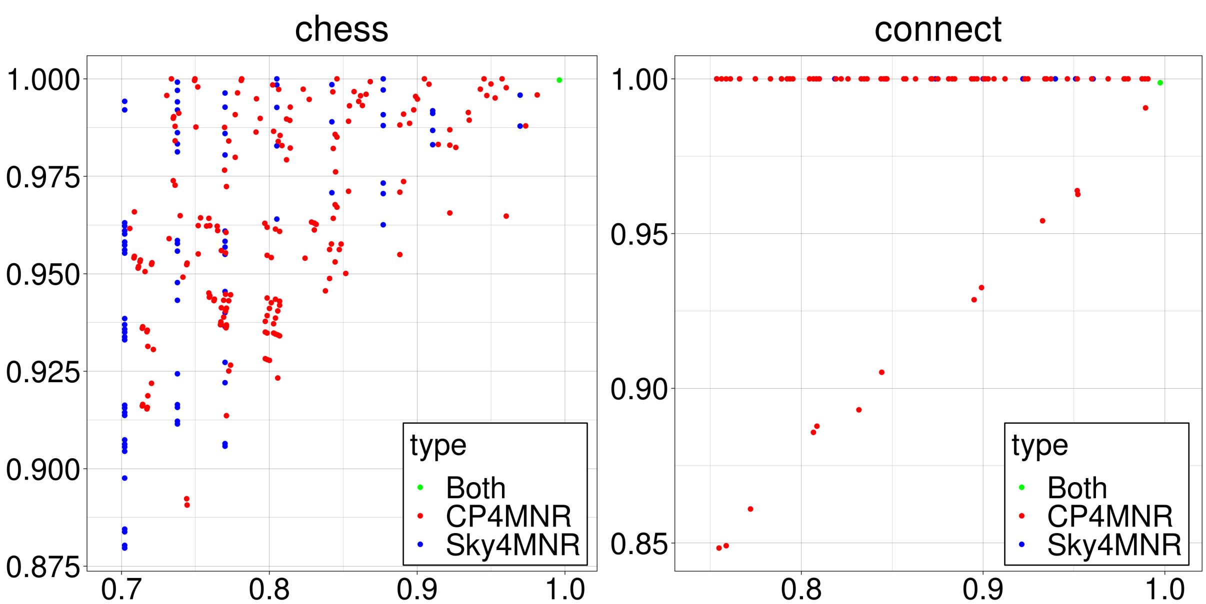

Figure 2: Qualitative analysis of MNRs (Chess and Connect datasets): support (x-axis) vs. confidence (y-axis)