Proof of optimality of minmax PIS algorithm

Jonathan Cohen

October 4, 2011

Assume we have an algorithm that operates on a graph $G = (V, E)$ and selects an independent set with the following properties:

- (selection) The algorithm decides at each node whether to include it in the set or not.
- (completeness) The algorithm selects at least one vertex in each connected component for inclusion.
- (independence) The selected set is independent.
- (locality) The decision to include vertex v depends only on information about edges incident at v, and vertices directly connected to v.

We can think of this algorithm as defining a selection operator which acts as the indicator function for the independent set. Call this operator Q_F, which is defined with respect to an arbitrary graph $G = (V, E)$ and subset of its edge $F \subseteq E$. For technical reasons, we always consider F to contains all self edges (u, u). Q_F has the following properties:

1. $Q_F(v) \mapsto \{true, false\}$ for $v \in V$
2. For any connected component under F, there exists at least one vertex v in that connected component such that $Q_F(v) = true$
3. If $Q_F(v)$ is true, then
   (a) $\forall u$ such that $(u, v) \in E$ and $u \neq v$, $Q_E(u) = false$ and
   (b) $Q_{F'}(v) = true$ where $F' \subseteq F$.

Condition 1 is a consequence of the selection property, and is equivalent to saying Q is the indicator function for this selected set. Condition 2 follows from completeness. Condition 3a guarantees that the set is independent.

Condition 3b says that if you drop edges from the graph, the independent set cannot shrink, and can be derived as follows. Consider a vertex v for which $Q_F(v)$ is true, but there exists an edge to (u, v) such that for $F' = F/\{(u, v)\}$, $Q_{F'}(v)$ is false. Now consider that F and F' contain no edges that are not incident at v. By locality, this cannot change the value of $Q_F(v)$ or $Q_{F'}(v)$. Further, since $Q_F(v)$ is true, that implies that for all vertices connected at v, Q_F is false (by independence). Again by locality, since (u, v) is only incident at u and v, the value of $Q_{F'}$ must be the same as the value of Q_F at all points besides v and u, which is false. Therefore, under $Q_{F'}$, we now have a connected component consisting of v and all points incident at v (discounting u) for which $Q_{F'}$ does not select any of the vertices in that component. This violates completeness, and therefore $Q_{F'}(v)$ must be true. This proves condition 3b.

Q induces a relationship on the nodes V, called $\trianglerighteq$, defined as follows:

Definition 1. *$u \trianglerighteq v$ if $\forall F \subseteq E$ such that $(u, v) \in F$, then either $Q_F(v) = false$ or $u = v$.*

We first state on useful property of $\trianglerighteq$

Lemma 1. *If $F = \{(a, b)\}$, $a \trianglerighteq b$, and $a \neq b$, then $Q_F(a) = true$ and $Q_F(b) = false$. This follows because $\{a, b\}$ form a connected component, so at least one of them must be true. Since $Q_F(b)$ must be false, it follows that $Q_F(a) = true$.*

We will now prove the $\trianglerighteq$ is a total ordering on V. This requires 3 conditions:

1. (antisymmetry) $a \trianglerighteq b$ and $b \trianglerighteq a \Rightarrow a = b$.
2. (transitivity) $a \trianglerighteq b$ and $b \trianglerighteq c \Rightarrow a \trianglerighteq c$.
3. (totality) Either $a \trianglerighteq b$ or $b \trianglerighteq a$ for all a and b.

Theorem 1. *$\trianglerighteq$ is a total order.*

Proof. Antisymmetry. Assume $a \trianglerighteq b$ and $b \trianglerighteq a$. Consider $F = \{(a, b)\}$. Since, under F, $\{a, b\}$ forms a connected component, then either $Q_F(a) = true$ or $Q_F(b) = true$. Assume $Q_F(a) = true$. Since $b \trianglerighteq a$, then $a = b$ by the definition of $\trianglerighteq$. If we assume $Q_F(b) = true$, since $a \trianglerighteq b$, then $a = b$ by the same argument.

Transitivity. Let $a \trianglerighteq b$ and $b \trianglerighteq c$ and assume $a \neq b \neq c \neq a$ (since otherwise the result is trivial). Consider $F = \{(a, b), (b, c), (a, c)\}$. Under this F, $\{a, b, c\}$ form a connected component, and therefore Q_F must be true for at least one of them. If $Q_F(c) = true$, then $Q_{F'}(c) = true$ for $F' = \{(b, c)\}$ by property 3b. That would violate Lemma 1, so this case cannot be true. Similarly, $Q_F(b) = false$ which can be seen by considering the set $\{(a, b)\}$. Therefore, $Q_F(a) = true$. Therefore, for $F'' = \{(a, c)\}$, $Q_{F''}(a) = true$ (by property 3b) and $Q_{F''}(c) = false$ (by property 3a). Now consider any set of edges H which contains $\{(a, c)\}$. If $Q_H(c) = true$, that would also imply that $Q_{F''}(c) = true$, which cannot be. Consequently, $a \trianglerighteq c$.

Totality. Assume a, b do not satisfy $a \trianglerighteq b$ and $a \neq b$. Then there exists an F containing (a, b) such that $Q_F(b) = true$. Let $F' = \{(a, b)\}$. By 3b, $Q_{F'}(b) = true$, and hence $Q_{F'}(a) = false$ (by 3a). Now consider any set H that contains (a, b). If $Q_H(a) = true$, it would imply that $Q_{F'}(a) = true$ by property 3b, which cannot be. Therefore, $Q_H(a) = false$, which implies that $b \trianglerighteq a$. $\square$

Next, we show that Q can be defined as the "max" operator under this total order.

Theorem 2. *Given a total order $\trianglerighteq$, define $Q'_F(a)$ to be true when a is the maximal element among its 1-ring under F, and false otherwise. Then Q' satisfies properties 1-3.*

Proof. Property 1 is by definition.

Property 2 is from the existence of a maximal element from any set under a total order.

Property 3a is from the uniqueness of a maximal element under a total ordering.

Property 3b is from monotonicity of $\trianglerighteq$.

$\square$

We have shown that $\trianglerighteq$ induces a total ordering on the vertices V, and that Q selects the maximum value of the 1-ring at a vertex under this ordering.

Now assume we have two operators, Q and P, both of which satisfy properties 1-3, which we use to select independent sets of G. Further, assume that these independent sets are disjoint (excluding any singleton connected components) so that if $P(a) \Rightarrow \neg Q(a)$ and $Q(a) \Rightarrow \neg P(a)$. Both induce total orders on the vertices, which we denote by $\trianglerighteq_P$ and $\trianglerighteq_Q$.

Let $\trianglelefteq$ be the inverse of $\trianglerighteq$ defined as $a \trianglelefteq b$ if and only if $b \trianglerighteq a$. This leads to the following result.

Theorem 3. $\trianglerighteq_P = \trianglelefteq_Q$,

Proof. Since we are excluding singletons, consider any pair of vertices a and b. Let $F = \{(a, b)\}$ and assume without loss of generality that $a \trianglerighteq_Q b$. Therefore, $Q_F(a) = true$, which implies that $P_F(a) = false$, which implies $P_F(b) = true$ by property 2. Because P_F selects the maximum under the total order $\trianglerighteq_P$, this means that $b \trianglerighteq_P a$. By the definition of $\trianglelefteq$, we have our result. $\square$

So if we have multiple algorithms that satisfy selection, independence, completeness, and locality, and the sets that they select are disjoint, the induced total orders must be inverses of each other. Since a total order and its inverse are uniquely defined, we can have at most two such total orders, which implies at most two selection operators, which implies at most two algorithms.

We now state our main result.

Theorem 4. *The minmax independent set algorithm which selects both the local min and the local max nodes simultaneously is optimal, in the sense that no parallel algortihm can select more than two disjoint independent sets at once.*