Proof of lifting (Lemma 1): Let $\mathcal{V} := value(\mathbf{F}\flat)$. We claim that $value(\mathbf{F}) = \mathcal{V}$ if there are no novel discordant atoms. To prove this suppose the negation: there are no novel discordant atoms yet the value of $\mathbf{F}$ in every model $\mathbf{M}$ is $<\mathcal{V}$. We define an partial valuation $\mathbf{I}$ that assigns to each ground instance of active atoms the value of their MSG in $\mathbf{M}_\flat$. *E.g.* we are considering ground atom P, it takes value $\mathbf{I}(P) = value(Q\flat, \mathbf{M}_\flat)$ where the $Q = \mathrm{msg}(P, \mathcal{A}(\mathbf{F}))$. The partial valuation can be turned into a total valuation by arbitrarily assigning values to the other atoms.

The value of $\mathbf{F}$ decreases only if an atom in the covering set S changes. If values of atoms in the covering set are fixed, then changing the value of a non-active atom can only increase the value of $\vee$-clauses—the values of the Σ-clauses in the covering set are preserved and so are a lower-bound for the $\vee$-clause. Hence we can only increase the value of the main $\wedge$-clause. Therefore the partial interpretation that assigns values to all active atoms lower-bounds the formula's value for all total valuations that extend $\mathbf{I}$.

Since in every model $value(\mathbf{F}, \mathbf{M}) < \mathcal{V}$, so there is at least one instance of a $\vee$-clause $C^<$ in every model such that all its Σ-clauses—including the clause $\Sigma^<$ in the covering set—are less than $\mathcal{V}$ when instantiated. In particular let $C^<$ be the minimal clause (in $\succ$-order) in an arbitrary model that has such an instance and let the appropriate instantiator be θ (Def. 2).

If $value(\Sigma^<\theta) \leq value(C^<\theta) < \mathcal{V}$ then at least one active atom $P \in \Sigma^< \in C^<$ must have a different value than $P\flat$ when instantiated: $value(P\theta, \mathbf{I}) \neq value(P\flat, \mathbf{M}_\flat)$. Since $P\theta$ takes on a different value, there must another way of generating it—there must be some Σ-clause $\Sigma^G \in C^G$ in the covering set that generates ground instance $P\theta$. So there must be an atom $Q \in \Sigma^G$ such that $Q\sigma = P\theta$ for some σ (it may be helpful to refer to Figure 1 for clarification). P must be greater than Q in the $\succ$-ordering since $P\theta$ is generated from Q and not P. The ordering must additionally be strict since if P and Q were mere renamings they would have the same value in $\mathbf{M}_\flat$.

$$\begin{tikzcd}
 & P \arrow[d, "\rho"'] \arrow[dd, bend left=60, "\theta"] \\
Q \arrow[r, "\rho"'] \arrow[dr, bend right, "\sigma"'] & P\rho \arrow[d] \\
 & P\theta
\end{tikzcd}$$

Figure 1: A diagram summarizing the structure of $\mathbf{I}$ for Q and P. Edges indicate substitution.

$C^<$ and C^G are potential clauses to apply SIG to with witness $\rho = \mathrm{mgu}(P, Q)$. Such a most general unifier exists since they have a common instance and can be standardized apart. Since F has no new discordant atoms then both conclusions must already be in F. Since P is strictly greater than Q it must be the case that ρ is not renaming for P. Therefore $P \succ P\rho$ and $C^< \succ C^<\rho$.

If there exists a ground instance of $C^<\rho$ such that all Σ-clauses are less in $\mathbf{I}$ than $\mathcal{V}$ then this contradicts the minimality (in terms of $\succ$) of $C^<$ as a counterexample. Therefore, every ground instance of $C^<\rho$ has value no less than $\mathcal{V}$. In particular, this implies that the value of $C^<\theta$ is no less than $\mathcal{V}$ which contradicts our supposition that $C^<\theta$ was a counterexample for the lifted value of $\mathbf{F}$ being bounded from below by $\mathcal{V}$. Therefore, $\mathcal{V}$ is a lower bound of $\mathbf{F}$. □

Proof of subproblem discord locality (Lemma 2): Suppose that the $value(F\flat) \geq 0$. For any set of any ground instances like S, any model that generalizes it must have no less value. The MSG S' is such a set, therefore $value(S') \leq value(S)$. By our supposition and Proposition 2 $value(S) < 0 \leq value(F\flat) \leq value(S'\flat)$, so it must be the case that the ground instance $S'\flat$ lacks some of the constraints on the values that are present in the ground instance $S'\sigma = S$. Here σ is the substitution that explains why S' is a generalization of S.

Since $value(S'\sigma) < value(S'\flat)$ there must be additional constraints in $S'\sigma$ that are not present in $S'\flat$. Since they differ only by how they were instantiated this means that there are atoms P and Q that are distinct in $S'\flat$, yet are unified—and therefore constrained to take on the same value—in $S'\sigma$. Therefore $value(P\flat) \neq value(Q\flat)$ yet $P\sigma = Q\sigma$ so they satisfy our definition of a discordant atomic pair and can generate an instance.

If no such instance is novel, then by Lemma 1 $S'\flat$ is an lower-bound on the value of S'—this contradicts our supposition that $value(S) < value(S'\flat)$. Therefore, there has to be at least one new instance generated by clauses C^P and C^Q with some witness θ.

Both conclusions are in S'', the MSG of S in $\mathbf{F}' = \mathbf{F} \wedge C^P\theta \wedge C^Q\theta$. By construction P and Q are unified in S and by definition of SIG the witness θ is the most general unifier of P and Q. Therefore $C^P\theta$ is a generalization of $C^P\sigma \in S$ (similarly for C^Q). They are more specific than elements of the old set S' by Proposition 3 and therefore are members of the new MSG of S and thus SIG is applicable. □

Proof of finiteness (Lemma 3): In order to prove that the S_i—the sequence of MSGs—eventually terminate we show they descend along a partial ordering. We then show that there are only finitely many sets that could be S_i, and that cycling is impossible. This proves that the sequence must have finite length.

Define a partial ordering $\succ^S$ over generalizations S_i in the following way: $S' \succ^S S''$ if for all C in the original subproblem S, $C' \succeq C''$, where $C' = \mathrm{msg}(C, S')$ and $C'' = \mathrm{msg}(C, S'')$. Also, at least one of these must be strict: there must exist a C such that $C' \succ C''$.

When we apply SIG to S_i (and we always can as long as $value(S_i\flat) \geq 0$) then both conclusions are members of S_{i+1} by Lemma 2. By Proposition 3 at least one of these conclusions is strictly smaller in $\succ$-order than its premise so $S_i \succ^S S_{i+1}$.

S is a finite subproblem—say that it has N elements—therefore each S_i must have at most N elements since each element $C \in S_i$ uniquely generalizes an element of S (if this is not true then C is redundant and can be discarded from S_i). Additionally, since S is finite it must have a finite parse-tree depth (*e.g.* the depth of $P(x)$ is one, the depth of $P(f(x))$ is two). Let this depth be d. Let n be the maximum arity of all functions and predicates.

There is only a finite number of sets that could generalize S. Since instantiation can never decrease the size of a formula (in terms of the parse-tree), only formulas that are no larger than the largest $\vee$-clause in S could be a member of any S_i. There is a finite number of formula with maximum depth d with maximum arity n, and each S_i is set of at most N of these formula, so there is a finite number of sets that could be S_i.

Therefore the S_i either terminate or cycle. If they cycle then we must have an element such that $S_j \succ^S S_j$, which is a contradiction. Therefore, the sequence of S_i is finite and terminates in either S or some non-ground generalization of S that exhibits negative value. □

Extended planning example: Consider the following planning problem: a vehicle with a bad battery is trying to get to the top of a hill. It is trying to reach this particular goal in a fixed amount of time (eight time-steps), but due to its battery it decelerates as it moves. Because of noisy sensor information, we do not know the exact structure of the hill. We do know that most terrain takes between two to three time-steps for the vehicle to traverse. We also have better information about the start and goal: the start takes less time-steps to get out of (the bottom of the hill is flat) and the goal takes more time-steps to enter (the top of the hill is unusually craggy).

Due to its failing battery the vehicle is slowing down: each subsequent edge on the path takes an additional time-step. So if a particular edge E takes between two and three time-steps with a fresh battery, after visiting a previous locations E it takes between three and four.

There are two features of this problem that make it interesting. Firstly, since the graph contains an unknown number of nodes it is truly a first-order problem. We cannot pass the description of this problem to a proposition solver and need some way—like our solver—of instantiating the unnamed objects.

Secondly, this problem would be longer to express in FOL (without LIA theory) since the idea of 'length' or 'time' requires linear integer arithmetic which is easily captured in FOP formula that follows. FOL would require a verbose 'one-of-k' representation—or additional axioms for arithmetic—which is less concise.

We can model this situation as follows (omitting some details). First, we define a path:

$$\begin{aligned}&\texttt{Start}(j) - 1 \vee \texttt{Path}(i,j) &(11)\\ &\quad - \texttt{Path}(i,p(j)) - \texttt{Edge}(p(j),j) - \texttt{Decl}(p(j)) &\\ &\texttt{Start}(j) - 1 \vee \texttt{Path}(i,p(j)) &(12)\\ &\quad + \texttt{Edge}(p(j),j) + \texttt{Decl}(p(j)) - \texttt{Path}(i,j) &\end{aligned}$$

These clauses insist that either a location is the start, or there is a previous location (denoted by the Skolem function $p(i)$) that explains how it got there with. This lets us break down the total time into some prior path time $\texttt{Path}(i,p(j))$ plus the base time of the edge $\texttt{Edge}(p(j),j)$ plus the deceleration penalty $\texttt{Decl}(p(j))$. Line 11 and 12 bound on $\texttt{Path}(i,j)$ from above and below, and together they enforce equality. The deceleration penalty grows with the $p(i)$ function.

We describe the edge structure of this graph of unknown size as follows (see Figure 2 for a diagram). All edges out of a take between one or two time-steps. All edges into b require three or four steps. General edges, between neither a nor b take two or three time-steps. All other edges take at least twelve time-steps, which is too costly to be useful (since our plan must be eight time-steps long in total).

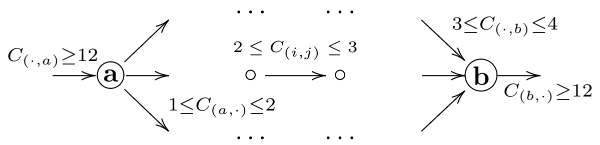
$C_{(\cdot,a)}\geq 12$
a
$1\leq C_{(a,\cdot)}\leq 2$
. . .
$2 \leq C_{(i,j)} \leq 3$
$3\leq C_{(\cdot,b)}\leq 4$
b
$C_{(b,\cdot)}\geq 12$

Figure 2: An description of a graph of unknown size. C_e represents the cost of the edge that it is decorating.

For example, the constraint on general edges can be expressed as:

$$-\texttt{A}(i) - \texttt{A}(j) - \texttt{B}(i) - \texttt{B}(j) \vee \texttt{Link}(i,j) - 2 \qquad (13)$$

$$-\texttt{A}(i) - \texttt{A}(j) - \texttt{B}(i) - \texttt{B}(j) \vee 3 - \texttt{Link}(i,j). \qquad (14)$$

These constraints say that either one of the endpoints conincides with the start or end, or the link cost must be greater than two or less than three. The predicate $\texttt{A}(i)$ represents whether alocation conincides with the fixed start **a**. $\texttt{B}(i)$ is similarly defined for the goal **b**

Finally, we add the negation of the statement that asserts that the vehicle can go from the start **a** to the finish **b** in exactly eight time-steps:

$$\texttt{Path}(\mathbf{a},\mathbf{b}) - 8 \qquad (15)$$

$$8 - \texttt{Path}(\mathbf{a},\mathbf{b}) \qquad (16)$$

This is impossible: with one previous $p(b)$, the longest path is 7 (take 2 to get out of **a**, take $4+1$ to get into **b** after decelerating by 1). With two previous locations the shortest path is 9 (1 out of **a**, $2+1$ between $p(p(b))$ and $p(b)$, and $3+2$ into **b**). Our proof procedure proves the infeasiblity of this formula after adding 76 refutations.