Graph-Based Bounds on *k*-Optimal Joint-Action Sets for Multiple Agents

Jonathan P. Pearce, Rajiv T. Maheswaran, and Milind Tambe

University of Southern California
{jppearce, maheswar, tambe}@usc.edu

Abstract. In multi-agent systems where sets of joint actions (JAs) are generated, tools are needed to evaluate these sets and efficiently allocate resources for the many JAs. To address evaluation, we introduce *k*-optimality as a metric that captures desirable properties of diversity and relative quality. Our main contribution is a method to utilize local interaction structure to obtain bounds on cardinalities of *k*-optimal JA sets. Bounds help choose the appropriate level of *k*-optimality for settings with fixed resources and help determine appropriate resource allocation for settings where a fixed level of *k*-optimality is desired. In addition, our bounds for 1-optimal JA sets also apply to the number of pure-strategy Nash equilibria in a graphical game of noncooperative agents.

1 Introduction

We consider a multi-agent system that generates a set of joint actions (JAs). The outcome for a single joint action (JA), a combination of individual actions, can be captured by a DCOP [1] or cost network [2] in cooperative domains and by graphical games [3] in noncooperative domains. These models decompose the system into a fixed interaction and reward structure. While we focus on the team setting where each JA in a set consumes resources, we also establish connections to settings with self-interested agents. Motivating cooperative domains include a team of troops that generate many sorties which consume supplies, a team of rescue units that generate many potential plans (for a disaster rescue commander) which consume human decision time, or a team of UAVs that generate many trajectories (for surveying or imaging) which consume film or fuel. JA sets can be (i) a sequence of JAs to execute or (ii) a set of choices, but in either case, they consume resources as a function of set size.

Most work in DCOPs and similar methods generates a single JA with high absolute reward, but reward alone is a poor metric for domains with JA sets, as it often yields clustered solutions. Clustering is undesirable, as diversity (the difference among JAs) is a key property for evaluating JA sets [4], e.g., commanders want varied options. However, diversity alone is undesirable, as we want to ensure a level of relative quality (each JA is best among a group of similar JAs and cannot be improved by simple changes). We define a metric, *k*-optimality, that naturally captures diversity and relative quality: a *k*-optimal JA has the highest reward within a neighborhood of JAs differing from it by at most *k* individual actions; *k*-optimality quantifies the neighborhood in which a local optimum is optimal. A *k*-optimal JA set (a collection of *k*-optimal JAs) then guarantees

M. Baldoni, U. Endriss, A. Omicini, and P. Torroni (eds.)
DALT 2005, *Third International Workshop*, Utrecht, July 25th, 2005
http://www.doc.ic.ac.uk/~ue/DALT-2005/

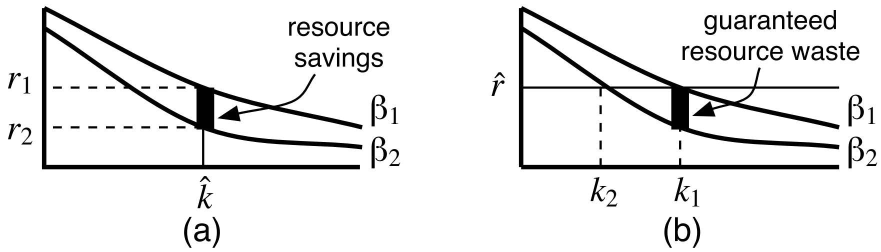


Fig. 1. A depiction of the advantages of tighter bounds

a level of relative quality (each JA is better than all JAs in a neighborhood of radius k) and diversity (JAs in the set must be separated by at least k individual actions).

Because each JA consumes resources, and the number of generated JAs may not be known *a priori*, resource allocation is a critical problem. Unfortunately, we cannot predict this number because the exact rewards in our domains are not known in advance. For example, supplies need to be allocated to troops executing multiple sorties before exact numbers and locations of adversaries are known. However, reward-independent bounds can be obtained on the size of k-optimal JA sets (i.e. to safely allocate enough supplies). First, we identify a mapping to coding theory yielding bounds independent of both reward and team interaction structure. We then provide, as our main contribution, a method to use the interaction structure (e.g., DCOP graph of arbitrary arity) to obtain significantly tighter bounds. We establish a connection to noncooperative settings by proving that our bounds for 1-optimal JAs also apply to the number of pure-strategy Nash equilibria in a graphical game.

Finding tighter bounds is useful in two ways: (i) If a particular level of k-optimality is desired, bounds indicate the maximum resource requirement for any k-optimal JA set. Thus, tighter bounds provide savings by allowing fewer resources to be allocated *a priori* while ensuring enough for all k-optimal JAs. Figure 1(a) shows, via a hypothetical example, how the tighter bound β_2 indicates that r_2 is sufficient for all $\hat{k}$-optimal JAs, yielding resource savings of $r_1 - r_2$ over using β_1. (ii) If resource availability is fixed, tighter bounds help us choose an appropriate level of k-optimality. If k is too low, we may exhaust our resources on bad solutions (similar JAs with poor relative quality). In contrast, (because fewer k-optimal JAs can exist as k increases), if k is too high, resources that could be spent on additional JAs are guaranteed to go unused. Tighter bounds provide a more accurate measure of guaranteed waste and thus, allow a more appropriate k to be chosen. In Figure 1(b), under hypothetical resource level $\hat{r}$, the looser bound β_1 hides the resource waste guaranteed when k_1 is used. This waste is revealed by β_2, indicating that k_2 reduces guaranteed resource waste.

2 *k*-optimality

We introduce the notion of a *k-optimal joint action* as a metric that captures both relative quality and diversity when selecting JA sets. The fitness of k-optimality for evaluating

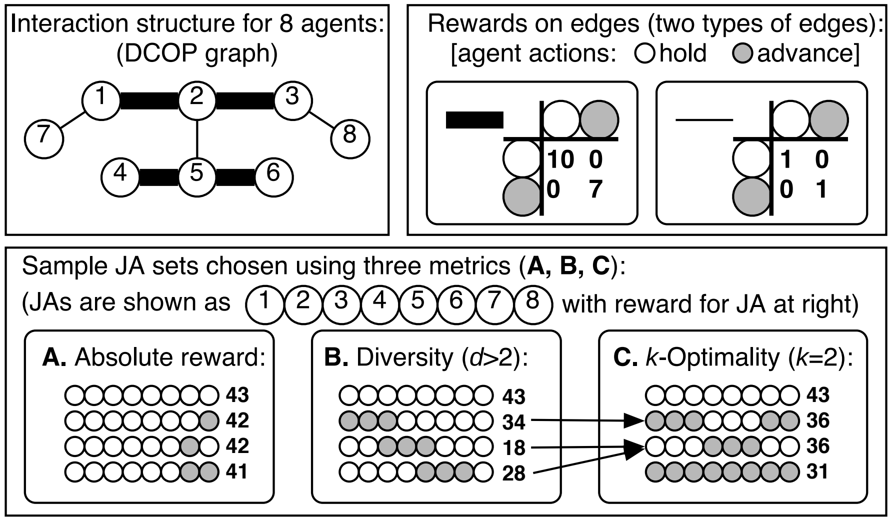


Fig. 2. Generating JA sets under various metrics

JA sets is illustrated in Figure 2, which shows a deployment of troops where each unit can advance or hold. Decision-support agents (assigned to each unit) coordinate to generate multiple sortie plans to be executed. Using reward alone as a metric to generate a JA set leads to a cluster of near-identical solutions (essentially, all troops hold). Using diversity alone (ensuring all JAs differ by more than two actions) leads to a JA set where many JAs can be improved with deviations of only two agents (shown by the arrows). With k-optimality (k=2), we generate a set of diverse JAs (all troops hold; front advances, rear holds; front holds, rear advances; all advance) where no JA can be improved with a two-agent deviation.

We begin with our model of the multi-agent team problem. For a set of agents $\mathcal{I} := \{1, \ldots, I\}$, the i^{th} agent takes action $a_i \in \mathcal{A}_i$. We denote the joint action of a subgroup of agents $S \subset \mathcal{I}$ by $a_S := \times_{i \in S} a_i \in \mathcal{A}_S$ where $\mathcal{A}_S := \bigtimes_{i \in S} \mathcal{A}_i$ and the joint actions (JAs) of the entire multi-agent team by $a = [a_1 \cdots a_I] \in \mathcal{A}$ where $\mathcal{A} := \bigtimes_{i \in \mathcal{I}} \mathcal{A}_i$. The team reward for taking a particular JA, a, is an aggregation of the rewards obtained by subgroups in the team:

$$R(a) = \sum_{S \in \mathcal{S}} R_S(a) = \sum_{S \in \mathcal{S}} R_S(a_S)$$

where S is a minimal subgroup that generates a reward (or incurs a cost) in an n-ary DCOP or cost network (i.e. a constraint), $\mathcal{S}$ is the collection of all such minimal subgroups for a given problem and $R_S(\cdot)$ denotes a function that maps $\mathcal{A}_S$ to $\mathbb{R}$. By minimality, we mean that the reward component R_S cannot be decomposed further: $\forall S \in \mathcal{S}$, $R_S(a_S) \neq R_{S_1}(a_{S_1}) + R_{S_2}(a_{S_2})$ for any $R_{S_1}(\cdot) : \mathcal{A}_{S_1} \to \mathbb{R}$, $R_{S_2}(\cdot) : \mathcal{A}_{S_2} \to \mathbb{R}$, $S_1, S_2 \subset \mathcal{I}$ s.t. $S_1 \cup S_2 = S$, $S_1, S_2 \neq \emptyset$. It is important to express the team reward in minimal form to accurately represent the dependencies and independencies among agents. Thus, $\mathcal{S} \subseteq \mathcal{P}(\mathcal{I})$ (where $\mathcal{P}(\cdot)$ denotes the power set) captures these local interactions.

To evaluate JA sets, specifically JAs with respect to each other, we need notions of neighborhood and distance among JAs. For two JAs, a and $\tilde{a}$, we define the following terms. The *deviating group* is

$$D(a, \tilde{a}) := \{i \in \mathcal{I} : a_i \neq \tilde{a}_i\},$$

the set of agents whose individual actions differ. The *distance* is

$$d(a, \tilde{a}) := |D(a, \tilde{a})|$$

where $|\cdot|$ denotes the cardinality of the set. The *relative reward* is

$$\Delta(a, \tilde{a}) := R(a) - R(\tilde{a}) = \sum_{S \in \mathcal{S}: S \cap D(a,\tilde{a}) \neq \emptyset} [R(a_S) - R(\tilde{a}_S)] .$$

We assume every subgroup G has a unique optimal (subgroup) joint action a^*_G for any context a_{G^C} (if $G \subset \mathcal{I}$ where $G \neq \emptyset$ and $G \neq \mathcal{I}$, then $\exists\ a^*_G \in \mathcal{A}_G$ s.t. $R(a^*_G; a_{G^C}) > R(a_G; a_{G^C})$ for all $a_G \neq a^*_G$; G^C denotes the complement of set G). This assumption is natural for any domain where rewards come from precise measurements, and is common in related work on bounds and estimates for numbers of local optima [5] and Nash equilibria [6, 7]. Given the above, we can now classify a as a *k-optimal joint action* if

$$\Delta(a, \tilde{a}) > 0\ \forall \tilde{a} \quad \text{s.t} \quad d(a, \tilde{a}) \leq k.$$

Every JA can be given a k, identifying the size of the neighborhood where it is locally optimal. A collection of k-optimal JAs will be mutually separated by a distance greater than k as they each have the highest reward within a radius of k. Thus, a higher k-optimality of a collection implies a greater level of relative reward and diversity. Let $A_k = \{a \in \mathcal{A} : \Delta(a, \tilde{a}) > 0\ \forall \tilde{a} \quad \text{s.t} \quad d(a, \tilde{a}) \leq k\}$ be the set of all k-optimal JAs. It is straightforward to show $A_{k+1} \subseteq A_k$.

Example 1. Figure 3 is a binary DCOP in which agents choose actions from $\{0, 1\}$, with rewards shown for the two constraints (minimal subgroups) $\mathcal{S} = \{\{1, 2\}, \{2, 3\}\}$. The JA $a = [1\ 1\ 1]$ is 1-optimal because any single agent who deviates reduces the team reward. However, $[1\ 1\ 1]$ is not 2-optimal because if the group $\{2, 3\}$ deviated, making the JA $\tilde{a} = [1\ 0\ 0]$, team reward would increase from 16 to 20. The optimal JA, $a^* = [0\ 0\ 0]$ is k-optimal for all $k \in \{0, 1, 2, 3\}$.□

3 Bounds on *k*-optimal joint actions

Bounds on $|A_k|$ can yield resource savings in domains where a particular level of k-optimality is desired, and can help determine the appropriate level of k-optimality to prevent guaranteed resource waste in fixed-resource settings. To find upper bounds on the number of k-optimal JAs, we discovered a correspondence to coding theory [8].

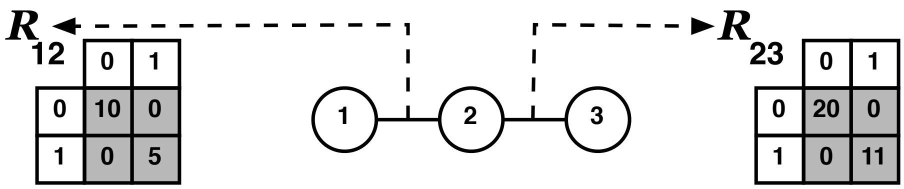


Fig. 3. DCOP example

Finding the maximum possible number of k-optimal JAs can be mapped to finding the maximum number of codewords in a space of q^I words where the minimum distance between any two codewords is $d = k + 1$. We can map words to JAs and codewords to k-optimal JAs as follows: A joint action a taken by I agents each with an action space of cardinality q is analogous to a word of length I from an alphabet of cardinality q. The distance $d(a, \tilde{a})$ can then be interpreted as a Hamming distance between two words. Then, if a is k-optimal, and $d(a, \tilde{a}) \leq k$, then $\tilde{a}$ cannot also be k-optimal because that implies the subgroup $D(a, \tilde{a})$ has two optimal (subgroup) joint actions to the context $D(a, \tilde{a})^C$, violating our assumption. Thus, any two k-optimal JAs must be separated by distance greater than k.

Three well-known bounds on codewords are Hamming[1]:

$$\beta_H = q^I / \left(\sum_{n=0}^{\lfloor k/2 \rfloor} \binom{I}{n} (q-1)^n \right),$$

Singleton:

$$\beta_S = q^{I-k},$$

and Plotkin:

$$\beta_P = \left\lfloor \frac{k+1}{k+1-(1-q^{-1})I} \right\rfloor$$

[8]. Thus, $|A_k|$, the number of k-optimal JAs for a given I and q, can be bounded by $\beta_{HSP} := \min\{\beta_H, \beta_S, \beta_P\}$. For example, to find a reward-independent bound on the number of 1-optimal JAs for three agents with $q = 2$, (e.g., the system in Figure 3), we obtain $\beta_{HSP} = 4$, without knowing R_{12} and R_{23} explicitly.

4 Graph-based exclusivity for multi-agent teams

The β_{HSP} bound is the same for a given (I, k, q), regardless of how the team reward is decomposed among subgroups of agents (i.e., the bound is the same for all $\mathcal{S}$). For

[1] For even k. For odd k, with $q = 2$, $\beta_H(I, k, q) = \beta_H(I-1, q, k-1)$ can be used to obtain a tighter bound. [8]

instance, the bound on 1-optimal JAs for Example 1 ($\beta_{HSP} = 4$ from the previous section) ignored that agent 1 does not interact directly with agent 3 and yields the same result independent of graph structure. However, taking local interactions (as captured by $\mathcal{S}$) into account can significantly tighten the bounds on $\{|A_k|\}_{k=1}^{I}$. In the previous analysis, pairs of JAs were mutually exclusive as k-optimal (only one of two could be k-optimal) if they were separated by a distance of k or less. We now show how some JAs separated by a distance of greater than k must be mutually exclusive as k-optimal.

We define $D_G(a, \tilde{a}) := \{i \in G : a_i \neq \tilde{a}_i\}$ and $V(G) := \cup_{S \in \mathcal{S}: G \cap S \neq \emptyset} S$. Intuitively, $D_G(a, \tilde{a})$ is the set of agents within the subgroup G who have chosen different actions between a and $\tilde{a}$, and $V(G)$ is the set of agents (including those in G) who are a member of some minimal subgroup $S \in \mathcal{S}$ that contains a member of G (e.g., G and the agents who share a constraint with some member of G). Then, $V(G)^C$ is the set of all agents whose contribution to the team reward is independent of the actions of G.

Proposition 1. *Let $a^* \in A_k$ and $\tilde{a} \in A$ be an assignment for which $d(a^*, \tilde{a}) > k$. If $\exists\, G \subset \mathcal{I}, G \neq \emptyset$ for which $|G| \leq k$ and $D_{V(G)}(a^*, \tilde{a}) = G$, then $\tilde{a} \notin A_k$.*

Proof. Given a^*, $\tilde{a}$, and G with the properties stated above, we have that $\forall a : d(a^*, a) \leq k$, $\Delta(a^*, a) > 0$. If a is defined such that $a_i = \tilde{a}_i$ for $i \in V(G)$ and $a_i = a_i^*$ for $i \notin V(G)$, then $D(a^*, a) = G$ and $d(a^*, a) \leq k$ which implies $\Delta(a^*, a) =$

$$\sum_{S \in \mathcal{S}: S \cap D(a^*, a) \neq \emptyset} R_S(a_S^*) - R_S(a_S) = \sum_{S \in \mathcal{S}: S \cap G \neq \emptyset} R_S(a_S^*) - R_S(a_S) =$$
$$\sum_{S \in \mathcal{S}: S \cap G \neq \emptyset} R_S(a_S^*) - R_S(\tilde{a}_S) > 0.$$

If $\hat{a}$ is defined such that $\hat{a}_i = a_i^*$ for $i \in V(G)$ and $a_i = \tilde{a}_i$ for $i \notin V(G)$, then $D(\tilde{a}, \hat{a}) = G$ and $d(\tilde{a}, \hat{a}) \leq k$, and $\Delta(\tilde{a}, \hat{a}) =$

$$\sum_{S \in \mathcal{S}: S \cap D(\tilde{a}, \hat{a}) \neq \emptyset} R_S(\tilde{a}_S) - R_S(\hat{a}_S) = \sum_{S \in \mathcal{S}: S \cap G} R_S(\tilde{a}_S) - R_S(\hat{a}_S) =$$
$$\sum_{S \in \mathcal{S}: S \cap G} R_S(\tilde{a}_S) - R_S(a_S^*) < 0,$$

thus, $\tilde{a} \notin A_k$. ■

Intuitively, if a JA a^* is k-optimal, then every subgroup of agents of size k or less has picked the best subgroup joint action for their context, so any other JA within a distance k of a^* contains a suboptimal subgroup joint action for their context. Since agents are typically not fully connected to all other agents, the *relevant context* a subgroup faces is not the entire set of other agents. Thus, the subgroup and its relevant context form a view (captured by $V(G)$) that is not the entire team. We consider the case where a JA $\tilde{a}$ has $d(a^*, \tilde{a}) > k$. We also have group G of size k within whose view $V(G)$, G are the only deviators between a^* and $\tilde{a}$ (although agents outside the view must also have deviated). We then show that $\tilde{a}$ contains a suboptimal subgroup joint action for a group G of size k or less and thus, cannot be k-optimal, i.e. if the group chose a_G^* instead of $\tilde{a}_G$ under its relevant context $V(G) \setminus G$ for $\tilde{a}$, then team reward would increase.

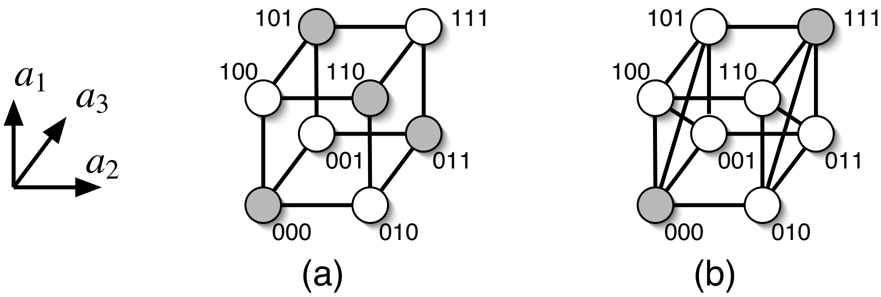


Fig. 4. Exclusivity graphs for 1-optimal JAs for Example 1, with sample maximum independent sets shaded.

To explain the significance of Proposition 1 to bounds, we introduce the notion of an *exclusivity relation* $E \subset \mathcal{I}$ which captures the restriction that if deviating group $D(a, \tilde{a}) = E$, then at most one of a and $\tilde{a}$ can be k-optimal. An *exclusivity relation set* for k-optimality, $\mathcal{E}_k \subset \mathcal{P}(\mathcal{I})$, is a collection of such relations that limits $|A_k|$, the number of JAs that can be k-optimal in a reward-independent setting (otherwise every JA could be k-optimal). The set $\mathcal{E}_k$ defines an *exclusivity graph* H_k where each node corresponds uniquely to one of all q^I JAs. Edges are defined between pairs of JAs, a and $\tilde{a}$, if $D(a, \tilde{a}) \in \mathcal{E}_k$. The size of the maximum independent set (MIS) of H_k, the largest subset of nodes such that no pair defines an edge, gives an upper bound on $|A_k|$. Naturally, an expanded $\mathcal{E}_k$ would imply a more connected exclusivity graph and thus a tighter bound on $|A_k|$.

Without introducing graph-based analysis, β_{HSP} for each k provides a bound on the MIS of H_k when $\mathcal{E}_k = \bigcup_{E \subset \mathcal{I}: 1 \leq |E| \leq k} E$. This set $\mathcal{E}_k$ captures only the restriction that no two JAs within a distance of k can both be k-optimal. The significance of Proposition 1 is that it provides additional exclusivity relations for JAs separated by distance greater than k, which arise only because we considered interaction structure (e.g., DCOP graph). This graph-based exclusivity relation set is

$$\widetilde{\mathcal{E}}_k = \bigcup_{E \subset \mathcal{I}: 1 \leq |E| \leq k} \; \bigcup_{F \in \mathcal{P}(V(E)^C)} [E \cup F]$$

which is a superset of $\mathcal{E}_k$. Additional relations exist because multiple exclusivity relations ($\bigcup_{F \in \mathcal{P}(V(E)^C)} [E \cup F]$) appear the same to the subgroup E because of its reduced view $V(E)$.

Consider Example 1, but with unknown rewards on the links. Here, the exclusivity relation set for 1-optimal JAs without considering interaction structure is $\mathcal{E}_1 = \{\{1\}, \{2\}, \{3\}\}$ which leads to the exclusivity graph in Figure 4(a) whose MIS implies a bound of 4. The exclusivity relation set for 1-optimal JAs when considering interaction structure is $\widetilde{\mathcal{E}}_1 = \{\{1\}\{2\}, \{3\}, \{1, 3\}\}$ which leads to the exclusivity graph in Figure 4(b) whose MIS implies a bound of 2. The set $\widetilde{\mathcal{E}}_1$ includes $\{1, 3\}$ due to the realization that agents 1 and 3 are not connected.

5 Graph-based exclusivity for Nash equilibria

Our graph-based bounds can be extended beyond agent teams to noncooperative settings. It is possible to employ the same exclusivity relations for 1-optimal JAs to bound the number of pure-strategy Nash equilibria in a graphical game using any of our bounds for $|A_1|$. Bounds on Nash equilibria [6] are useful both for design and analysis of mechanisms as they predict the maximum number of outcomes of a game.

We begin with a set of noncooperative agents $\mathcal{I} = \{1, \ldots I\}$, where the i^{th} agent's utility is

$$U^i(a_i; a_{\{\mathcal{I}\setminus i\}}) = \sum_{S_i \in \mathcal{S}_i} U^i_{S_i}(a_i; a_{\{S_i \setminus i\}})$$

which is a decomposition into an aggregation of component utilities generated from minimal subgroups (only those agents affecting a particular component utility are in the respective set S_i). The notation a_i and $a_{\{G\setminus i\}}$ refers to the i^{th} agent's action and the actions of the group G with i removed, respectively. We refer to a as a joint action (JA), with the understanding that it is composed of actions motivated by individual utilities. Let the *view* of the i^{th} agent in a noncooperative setting to be $V(i) = \cup_{S_i \in \mathcal{S}_i} S_i$. The deviating group with respect to G is: $D_G(a, \tilde{a}) := \{i \in G : a_i \neq \tilde{a}_i\}$. Assuming every player has a unique optimal response to its context, then if a^* is a pure-strategy Nash equilibrium, and $d(a^*, a) = 1$, $i = D(a^*, a)$, we know that

$$U^i(a^*_i; a^*_{\{\mathcal{I}\setminus i\}}) > U^i(a_i; a^*_{\{\mathcal{I}\setminus i\}})$$

and a is not a pure-strategy Nash equilibrium. However, applying the local interaction of the game, captured by the sets $\{\mathcal{S}_i\}$, we get exclusivity relations between JAs with distance greater than 1 as follows.

Proposition 2. *If a^* is a pure-strategy Nash equilibrium, $\tilde{a} \in A$ such that $d(a^*, \tilde{a}) > 1$, and $\exists i \in \mathcal{I}$ such that $D_{V(i)}(a^*, \tilde{a}) = i$, then $\tilde{a}$ is not a pure-strategy Nash equilibrium.*

Proof. We have

$$U^i(\tilde{a}_i; \tilde{a}_{\{\mathcal{I}\setminus i\}}) = \sum_{S_i \in \mathcal{S}_i} U^i_{S_i}(\tilde{a}_i; \tilde{a}_{\{S_i \setminus i\}}) = \sum_{S_i \in \mathcal{S}_i} U^i_{S_i}(\tilde{a}_i; a^*_{\{S_i \setminus i\}})$$
$$< \sum_{S_i \in \mathcal{S}_i} U^i_{S_i}(a^*_i; a^*_{\{S_i \setminus i\}}) = \sum_{S_i \in \mathcal{S}_i} U^i_{S_i}(a^*_i; \tilde{a}_{\{S_i \setminus i\}}) = U^i(a^*_i; \tilde{a}_{\{\mathcal{I}\setminus i\}}).$$

The first and last equalities are by definition. The second and third equalities are because $D_{V(i)}(a^*, \tilde{a}) = i$. The inequality is because a^* is a pure-strategy Nash equilibrium. The result is that $\tilde{a}_i$ is not an optimal response to $\tilde{a}_{\{\mathcal{I}\setminus i\}}$ and thus cannot be a pure-strategy Nash equilibrium. ■

Proposition 2 states that a^* and $\tilde{a}$ cannot both be Nash equilibria if $\exists i,\ D_{V(i)}(a^*, \tilde{a}) = i$, which is identical to the condition that prevents two JAs (in a team setting) from being 1-optimal. The commonality is that in both the cooperative and noncooperative settings, agents have optimal actions for any given context, and in both settings there is a notion of relevant context, $V(i) \setminus i$, which can be a subset of other agents $\{\mathcal{I} \setminus i\}$. The difference

Algorithm 1 for Symmetric Region Packing (SRP) bound

$\widetilde{\mathcal{E}}_k = \bigcup_{E \subset \mathcal{I}: 1 \leq |E| \leq k} \bigcup_{F \in \mathcal{P}(V(E)^C)} [E \cup F]$
a = [0 0 0]
$|A_k| = 1$
$B(a) = \cup_{E \in \widetilde{\mathcal{E}}_k} f(a, E)$
for all b $\in B(a)$ **do**
  $\overline{B}(b) = (\cup_{E \in \widetilde{\mathcal{E}}_k} f(b, E)) \setminus (a \cup B(a))$
  $\overline{H_k}(b)$.addNodes($\overline{B}(b)$)
  for all $\overline{b_1}, \overline{b_2} \in \overline{B}(b)$ **do**
    if $D(\overline{b_1}, \overline{b_2}) \in \widetilde{\mathcal{E}}_k$ **then**
      $\overline{H_k}(b)$.addEdge($\overline{b_1}, \overline{b_2}$)
    end if
  end for
  $M_b = |\text{cliquePartition}(\overline{H_k}(b))|$
  $|A_k| = |A_k| + 1/(1 + M_b)$
end for
$\beta_{SRP} = (q^I)/|A_k|$

is that the views are generated in different manners: $V(i) = \cup_{S \in \mathcal{S}: i \cap S \neq \emptyset} S$ in a cooperative setting, while $V(i) = \cup_{S_i \in \mathcal{S}_i} S_i$ in a noncooperative setting. Given the views, we can generate the exclusivity relation set in the same manner, $\mathcal{E}_1 = \bigcup_{i \in \mathcal{I}} \bigcup_{F \in \mathcal{P}(V(i)^C)} [i \cup F]$. If a noncooperative graphical game yields a particular exclusivity relation set, it defines the same exclusivity graph as a cooperative multi-agent domain with the same exclusivity relation set. Thus, the bound for the number of Nash equilibria for a noncooperative graphical game is identical to the bound for 1-optimal JAs for a cooperative multi-agent domain, if both share the same exclusivity relation set $\mathcal{E}_1$.

6 Algorithms for graph-based bounds

As seen earlier, the local interaction structure in both cooperative and noncooperative settings expands the exclusivity relation set for k-optimality. This set defines an exclusivity graph H_k whose maximum independent set (MIS) provides a bound for the number of JAs which are k-optimal (or alternatively, Nash equilibria). Finding the size of the MIS is NP-complete in the general case [9], so we investigated other techniques to obtain an upper bound on $|A_k|$. We observe that any fully-connected subset (clique) of H_k can contain at most one k-optimal JA. Therefore, the number of cliques in any clique partitioning of H_k also provides an upper bound on $|A_k|$. We used the polynomial-time F_{CLIQUE} algorithm, shown [10] to outperform several other clique-partitioning algorithms, to find clique partitionings with fewest cliques.

Another method that we developed (Algorithm 1) was the symmetric region packing bound, β_{SRP}, using a method analogous to sphere packing (the idea used to compute the Hamming bound [8]), where each k-optimal JA claims a region of the space of all JAs (the nodes of H_k). Because these regions are constructed to be disjoint and have identical volumes, dividing the space of all JAs by this volume yields a bound. Figure 5 shows β_{SRP} computed for 1-optimal JAs for Example 1. We choose an arbitrary JA

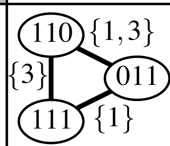
$\widetilde{\mathcal{E}}_1 = \{$	$\{1\},$	$\{2\},$	$\{3\},$	$\{1,3\}$ $\}$
$B([0\ 0\ 0]) = \{$	$[1\ 0\ 0],$	$[0\ 1\ 0],$	$[0\ 0\ 1],$	$[1\ 0\ 1]$ $\}$
$\widetilde{\mathcal{E}}_1 = \{$	$\overline{B}([1\ 0\ 0])$	$\overline{B}([0\ 1\ 0])$	$\overline{B}([0\ 0\ 1])$	$\overline{B}([1\ 0\ 1])$
$\{1\},$	$[0\ 0\ 0]$	$[1\ 1\ 0]$	$[1\ 0\ 1]$	$[0\ 0\ 1]$
$\{2\},$	$[1\ 1\ 0]$	$[0\ 0\ 0]$	$[0\ 1\ 1]$	$[1\ 1\ 1]$
$\{3\},$	$[1\ 0\ 1]$	$[0\ 1\ 1]$	$[0\ 0\ 0]$	$[1\ 0\ 0]$
$\{1,3\}\}$	$[0\ 0\ 1]$	$[1\ 1\ 1]$	$[1\ 0\ 0]$	$[1\ 1\ 1]$
$\overline{H}_1(b)$ (exclusivity subgraph)	110	 110 $\{1,3\}$ $\{3\}$ 011 111 $\{1\}$	111	011
$M_b =$	1	1	1	1
$1/(1+M_b) =$	$1/2$	$1/2$	$1/2$	$1/2$

Fig. 5. Computation of β_{SRP} for Example 1

$a \in A$ which we assume to be k-optimal ($a = [0\ 0\ 0]$ in Figure 5), around which we will construct a region claimed by a. Applying the exclusivity relations from $\widetilde{\mathcal{E}}_k$, we can generate a set $B(a) = \cup_{E \in \widetilde{\mathcal{E}}_k} f(a, E)$ where $f(a, E)$ yields the JA that is excluded from being k-optimal by a and E. The first two rows of Figure 5 show $\widetilde{\mathcal{E}}_1$ and the set $B([0\ 0\ 0])$. Applying the exclusivity relations again for each $b \in B(a)$, and discarding JAs already included in a or $B(a)$, we can generate a set $\overline{B}(b) = \cup_{E \in \widetilde{\mathcal{E}}_k} f(b, E)$ which contains all JAs that may also exclude b from being k-optimal. In Figure 5, we apply $\mathcal{E}_1$ to obtain $\overline{B}(b)$ for all $b \in B(a) = \{[1\ 0\ 0], [0\ 1\ 0], [0\ 0\ 1], [1\ 0\ 1]\}$ where the grayed out JAs are those discarded for being in $\{a\} \cup B(a)$. To ensure that the region that a claims is disjoint from the regions claimed by other k-optimal JAs, a should only claim a fraction of each $b \in B(a)$. This can be achieved if a shares each b equally with all other k-optimal JAs that might exclude b. These additional k-optimal JAs are contained within $\overline{B}(b)$. However, not all $\overline{b} \in \overline{B}(b)$ can actually be k-optimal as they might exclude each other. If we construct a graph $\overline{H}_k(b)$ with nodes for all $\overline{b} \in \overline{B}(b)$ and edges formed using $\widetilde{\mathcal{E}}_k$, and we find M_b, the size of the MIS, then a can safely claim $1/(1 + M_b)$ of b. We again use clique partitioning to safely estimate M_b. In Figure 5, $\overline{B}([0\ 1\ 0])$ leads to a three-node, three-edge exclusivity graph. By adding the values of $1/(1 + M_b)$ for all $b \in B(a)$ (plus one for itself), we obtain a can safely claim a region of size 3, which implies $\beta_{SRP} = \lfloor 2^3/3 \rfloor = 2$. Algorithm 1's runtime is polynomial in the number of possible JAs, which is a comparatively small cost for a bound that applies to every possible instantiation of rewards to actions.

7 Experimental results

To investigate the impact of graph-based analysis, we generated interaction graphs of varying size and density. We started with complete graphs (all pairs of agents are connected) where each node (agent) had a unique ID. Edges were removed one by one by choosing the lowest-ID node and removing the edge between it and its lowest-ID neighbor. Figure 6(a) shows the β_{HSP} and β_{SRP} bounds on the y-axis for k-optimal JAs for $k \in \{1, 2, 3, 4\}$ and $I \in \{7, 8, 9, 10\}$. The x-axis plots the number of links removed, δ, which identifies the interaction graph, given I. While $\beta_{HSP} < \beta_{SRP}$ for very dense graphs, β_{SRP} provides significant gains for the vast majority of cases.

We now provide a concrete demonstration of the gains due to tighter bounds, depicted in Figure 1. Each plot in Figure 6(a) is marked with a bar that indicates the resource savings from improved bounds for four interaction graphs with

$$(I, \delta) = \{(7, 12), (8, 15), (9, 20), (10, 24)\}.$$

For example, consider a troops/sorties problem similar to that in Figure 2, but where the interaction structure is defined by $(I, \delta) = (10, 24)$. For a fixed $k = 1$, β_{HSP} implies that we must equip the troops with 512 supplies in order to ensure that all supplies are not exhausted before all 1-optimal actions are executed. However, β_{SRP} indicates a 15-fold reduction to 34 supplies will suffice, yielding a savings of 478 supplies due to the use of graph structure when computing bounds.

For each of the four interaction graphs (I, δ), Figure 6(b) shows β_{HSP} and β_{SRP} (on a logarithmic scale) as a function of k. Each plot is marked with a bar displaying the guaranteed resource waste that is hidden by β_{HSP}, which indicates that k_1 should be chosen for a fixed resource level $\hat{r}$. Here, β_{SRP} shows that choosing k_2 will reduce the number of supplies guaranteed to be wasted when the supply level is fixed at $\hat{r}$, compared to k_1, experimentally validating our motivation in Figure 1. For instance, with interaction structure $(I, \delta) = (10, 24)$, with supply level fixed at 18, β_{HSP} suggests that choosing $k_1 = 4$ will avoid a guaranteed waste, because it believes that up to 18 sorties may be executed. However, β_{SRP} states that no more than 7 sortie plans can exist for $k_1 = 4$ revealing a guaranteed waste of at least 11 supplies if k_1 is chosen. It also suggests that $k_2 = 2$ will reduce this guaranteed waste, while at the same time safely ensuring that all supplies are not exhausted before all 2-optimal JAs (which include all 4-optimal JAs) are executed.

In Figure 7, we compared β_{HSP} and β_{SRP} to the bound obtained by applying F_{CLIQUE}, $\beta_{FCLIQUE}$. While $\beta_{FCLIQUE}$ is marginally better for $k = 1$, β_{SRP} has clear gains for $k = 4$. Identifying the relative effectiveness of various algorithms that exploit our exclusivity relation sets is clearly an area for future work.

8 Related work and conclusion

Our research on bounding k-optimal joint action sets in multi-agent domains is related to estimating numbers of local optima in centralized local search and evolutionary computing [5, 11]. The key difference is in the exploitation of constraint graph structure, not harnessed in previous work, to bound the number of optima.

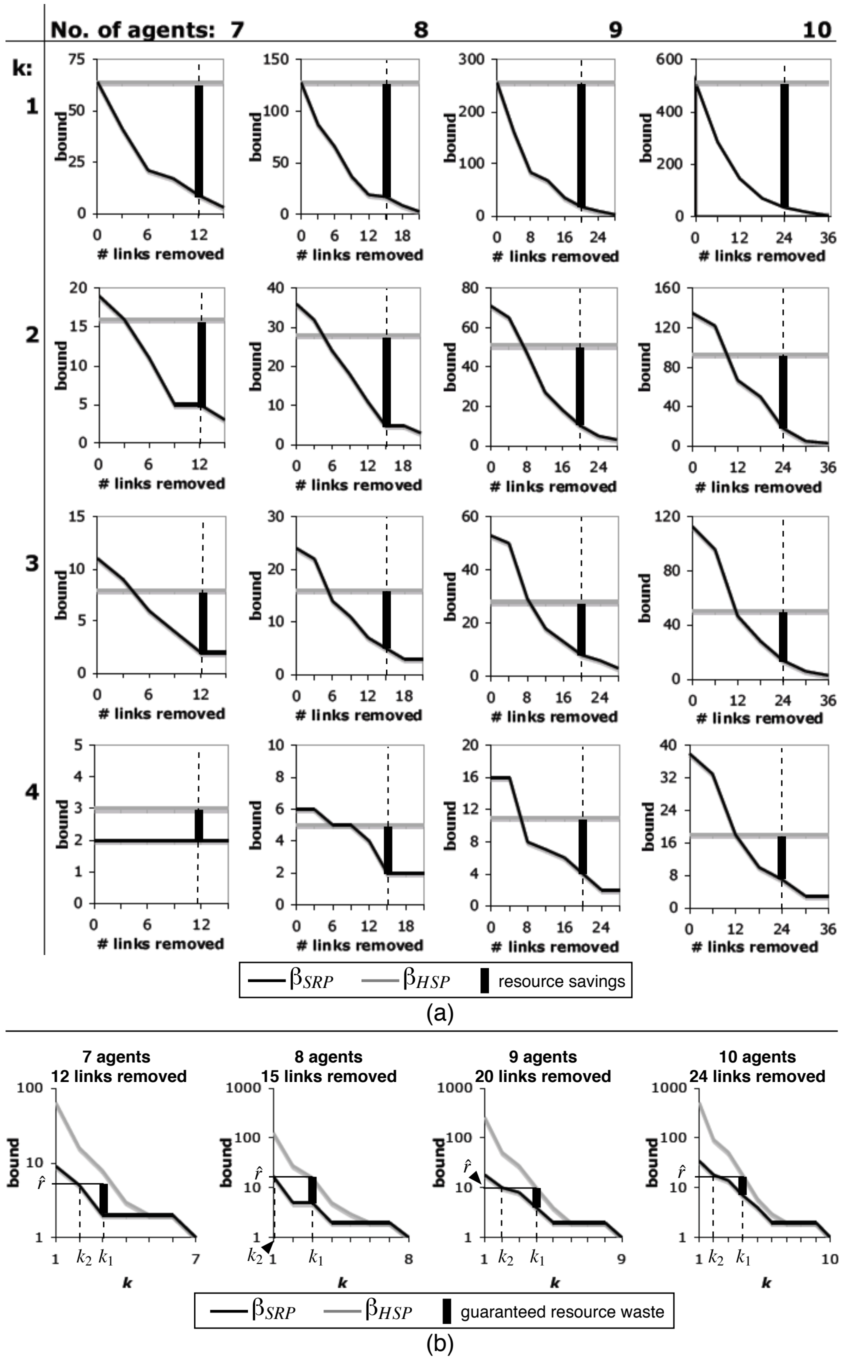


Fig. 6. Comparisons of β_{SRP} vs. β_{HSP}

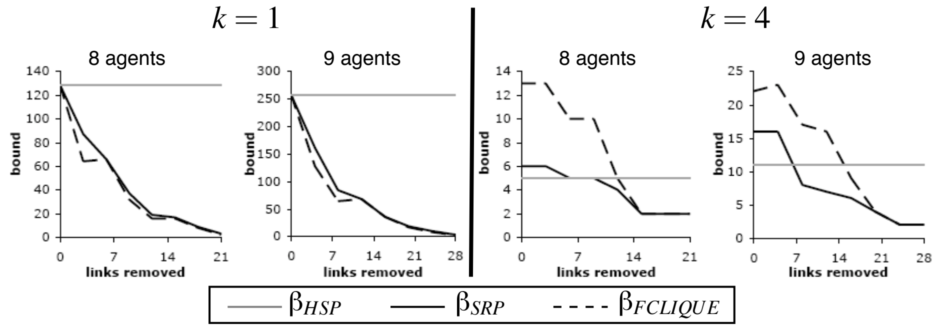


Fig. 7. Comparisons of β_{SRP}, β_{HSP} and $\beta_{FCLIQUE}$

Given that counting the number of Nash equilibria in a game with known payoffs is #$\mathcal{P}$-hard [12], bounds have been investigated for particular types of games [13, 6]. Graph structure is utilized in algorithms to expedite finding Nash equilibria for a given graphical game with known payoffs [14, 3]. However, finding tight bounds on Nash equilibria over all possible games on a given graph (i.e., reward-independent bounds) remained an open problem.

Finally, despite the seeming similarity of *k*-optimality to *k*-consistency [15] in centralized constraint satisfaction, the two concepts are entirely different, as *k*-consistency refers to reducing the domains of subsets of variables to maintain internal consistency in a satisfaction framework while *k*-optimality refers to comparing fixed joint actions where subsets of agents optimize with respect to an external context.

In this paper, (1) we have introduced *k*-optimality as a metric that captures the properties of diversity and relative quality which are desirable for evaluating JA sets. Finding bounds on *k*-optimal JA sets is useful for resource allocation problems associated with executing JA sets in sequence or presenting JA sets as options. (2) We discover a correspondence to coding theory that yields a bound (β_{HSP}) independent of reward and graph structure. Our main contribution is (3) a method to exploit interaction structure to obtain graph-based exclusivity relation sets which tighten reward-independent bounds. (4) We also show that our method extends to noncooperative settings, as exclusivity relation sets for 1-optimal JA sets can be used to find a reward-independent bound on the number of pure-strategy Nash equilibria in a graphical game. Finally, (5) we develop techniques for computing bounds ($\beta_{SRP}, \beta_{FCLIQUE}$) using the graph-based exclusivity relation sets and (6) illustrate their utility on a diverse collection of interaction structures.

References

1. Modi, P.J., Shen, W., Tambe, M., Yokoo, M.: Adopt: Asynchronous distributed constraint optimization with quality guarantees. Artificial Intelligence **161** (2005) 149–180
2. Dechter, R.: Bucket elimination: A unifying framework for reasoning. Artificial Intelligence **113** (1999) 41–85
3. Kearns, M., Littman, M., Singh, S.: Graphical models for game theory. In: Proc. UAI. (2001)

4. Tate, A., Dalton, J., Levine, J.: Generation of multiple qualitatively different plan options. In: Proc. AIPS. (1998)
5. Caruana, R., Mullin, M.: Estimating the number of local minima in complex search spaces. In: IJCAI Workshop on Optimization. (1999)
6. McLennan, A., Park, I.: Generic 4 x 4 two person games have at most 15 Nash equilibria. Games and Economic Behavior **26** (1999) 111–130
7. McLennan, A.: The expected number of Nash equilibria of a normal form game. Econometrica **73** (2005) 141–174
8. Ling, S., Xing, C.: Coding Theory: A First Course. Cambridge U. Press (2004)
9. Alon, N., Kahale, N.: Approximating the independence number via the theta-function. Mathematical Programming **80** (1998) 253–264
10. Kim, J.T., Shin, D.R.: New efficient clique partitioning algorithms for register-transfer synthesis of data paths. Journal of the Korean Phys. Soc. **40** (2002) 754–758
11. Whitley, D., Rana, S., Heckendorn, R.B.: Representation issues in neighborhood search and evolutionary algorithms. In D. Quagliarella, et al., ed.: Genetic Algorithms and Evolution Strategies in Engineering and Computer Science. Wiley (1998) 39–57
12. Conitzer, V., Sandholm, T.: Complexity results about Nash equilibria. In: IJCAI. (2003)
13. Keiding, H.: On the maximal number of Nash equilibria in an *n* x *n* bimatrix game. Games and Economic Behavior **21** (1995) 148–160
14. Blum, B., Shelton, C., Koller, D.: A continuation method for Nash equilibria in structured games. In: IJCAI. (2003)
15. Freuder, E.C.: Synthesizing constraint expressions. Communications of the ACM **21** (1978) 958–966