Using Annotated Guidelines to Influence the Behavior of Organizationally Adept Agents

Daniel D. Corkill, Chongjie Zhang, Bruno da Silva, Yoonheui Kim, Xiaoqin Zhang[2], and Victor R. Lesser

[1] University of Massachusetts Amherst, Amherst, MA 01003, USA
[2] University of Massachusetts Dartmouth, North Dartmouth, MA 02747, USA

Abstract. An *organizationally adept* agent (OAA) adjusts its behavior when given annotated organizational guidelines. More importantly, it can also determine when such guidelines become ineffective and proactively adapt its behavior to better achieve organizational objectives. A central OAA tenet is a clear separation between *operational decision making* (the detailed moment-to-moment behavior decisions made by an agent) and *organizational control* (longer-term directives designed using estimates of environment and agent characteristics and expressed to agents as annotated guidelines that bias and inform their operational decision making). This separation enables the OAA to stop following guidelines when the estimates used in their design were incorrect or when the environment changes over time and to propose and negotiate agreements with other OAAs to replace such guidelines.
We describe a fully operational OAA architecture that: 1) allows agents to operate reasonably without organizational guidelines; 2) uses belief values in operational decision making that are updated by experience and can be seeded by expectations conveyed in guideline annotations; 3) assesses the appropriateness of guidelines based on deviations from annotated estimates developed during their design; and 4) can make agreements to replace inappropriate guidelines. We present details of this approach to agent organization and analyze its effectiveness using call-center OAAs striving to extinguish fires in RoboCup Rescue scenarios.

1 Introduction

Coordinating many individual agents' activities to achieve collective benefit in complex, dynamic environments is hard. As the number of agents working together on tasks increases (and the precedence relationships and potential interference among tasks and resources grows), it becomes increasing difficult for an agent to decide what it should be doing, when it should be doing it, and with what other agents it should coordinate its actions). Designed agent organizations [3, 14, 6] provide agents with organizational directives that, when followed, reduce the complexity and uncertainty of each agent's activity decisions, lower the cost of distributed resource allocation and agent coordination, help limit inappropriate agent behavior, and reduce unnecessary communication and agent activities.

In order to participate in an agent organization, an agent must be able to adjust its normal behavior when it is given organizational directives. These directives contain general, long-term guidelines, in the form of parametrized role assignments and priorities (e.g., prefer extinguishing fires in region A over fires in Region B), that are subject to ongoing elaboration into precise, moment-to-moment activity decisions by the agents [11, 4]. Following organizational directives is beneficial when agent directives can be designed that perform well over a range of potential long-term environment and agent characteristics. On the other hand, following directives when the estimates used in their design are incorrect or have changed over time can be worse than not having directives at all.

We have developed a novel paradigm for an *organizationally adept* agent (OAA) [2] that can operate in an uncertain organizational setting. Creating an OAA involves a number of essential features, the most important of which is providing each OAA with information about the assumptions made by the organization designer (whether designed by a human or by an automated designer process [14, 7]) and the range of conditions the OAA should expect if the organization is operating as intended (e.g., the environment is not within the expected range of characteristics or that task performance differs from expectations). This richer coupling of the designer's intent with the OAAs is accomplished by augmenting the organizational guidelines given to each OAA with designer-expectation *annotations*. Annotations include the assumed range of environmental characteristics (e.g., expected task-arrival rates) and performance estimates (e.g., task-completion time and agent-interaction amounts). These annotations help an OAA determine when the expectations that were used when designing the supplied guidelines do not hold and when the OAA should abandon following guidelines in favor of proactively adapting its behavior. A central tenet in the OAA approach is a clear separation between *operational decision making*, the detailed moment-to-moment behavior decisions made by agents, and *organizational control*, expressed through annotated guidelines that bias and inform operational decision making. This separation allows an agent to distinguish decisions influenced by the guidelines from choices that would have been made without them.

The decentralized OAA approach to detecting and adapting inappropriate organizational directives is most closely related to early work by Horling [5]. More recent work in the ALIVE system [15, 12] also deals with adapting organizational structures based on failure events rather than on deviation from the designer's expectations as conveyed using guideline annotations. ALIVE also takes a centralized approach to the adaptation process. Work on MOISE+ [8] includes a nice conceptualization of the need for organizational change but takes a top-down approach to reorganization. Our agent-centric OAA approach monitors local deviations from design assumptions (conveyed in the annotations), decides when to abandon guidelines, and performs localized (not top-down) organizational-behavior adaptation; perhaps in advance of a more informed redesign by a human or automated designer.

We next describe the important characteristics of the OAA architecture and how it can: 1) operate even without organizational guidelines; 2) adjust its operational decisions to conform with organizational guidelines, if supplied; 3) assess the appropriateness of the organizational guidelines based on deviation from annotations describing the task and environmental assumptions used when the guidelines were designed; 4) stop following guidelines deemed to be inappropriate; and 5) propose and negotiate agreements with other agents to use in place of inappropriate guidelines. As we will discuss, organizational control requires deep ties into an agent's operational and domain reasoning, and our generic OAA architecture makes this organizational-control connection clear and manageable. We then present details of an instantiation of the OAA architecture in a domain composed of call-center OAAs operating in RoboCup Rescue [10] and analyze their performance and ability to perform these five OAA abilities in different RoboCup Rescue scenarios.

2 OAA Architecture

The heart of the OAA architecture is an event-driven, BDI-like [13] operational decision-making engine that adjusts its decisions when it is provided with parametrized role priority assignments specified in organizational guidelines, and its decisions are informed by the belief values contained in guideline annotations. The OAA receives *percepts* both from the external environment (e.g., sensor reports or messages from other agents) and from its internal decision-making process and task-execution performance (e.g., plan failure or inability to achieve a goal). These percepts cause changes in the OAA's *beliefs*, and those changes can trigger the creation and modification of *goals*. Goals that pertain to normal operational activity decisions (e.g., to extinguish a specific fire), to operational adaptation (e.g., to borrow a fire-brigade resource), and to organization adaptation (e.g., to negotiate an agreement to replace inappropriate guidelines) can be instantiated from external and internal percepts. Each created goal is instantiated from a *goal class* that has a set of *plan templates*, each of which can potentially achieve the goal when fully instantiated and executed. Each plan template consists of a partial order of primitive *actions* and includes a specification of the number and types of resources R the plan template requires. A *plan* p is instantiated by assigning (binding) specific resources R_p to a plan template (e.g., use specific fire brigades to fight a specific fire).

The OAA operational decision-making engine determines the specific resource assignments that should be made to create $\{\text{goal, plan}\}$ pairs for execution. It uses a greedy, utility-based, non-preemptive scheduling process that maximizes the total estimated utility based on the current availability of resources. Before we detail how organizational guidelines and annotations bias $\{\text{goal, plan}\}$ instantiation and scheduling, we describe how unbiased utility-based OAA scheduling operates.

The estimated utility $U(g, p)$ of an *intention* (goal g achieved by plan p with resources R_p) is defined as:

$$DB(g,p) = B(g) \times S(g,p) \times P(g,p) \tag{1a}$$

$$U(g,p) = DB(g,p) - O(R_p, D(g,p)) \tag{1b}$$

$DB(g, p)$ is the estimated *discounted benefit* of using plan p to achieve goal g and is computed using: $B(g)$, the *benefit* of fully achieving goal g; $S(g, p)$, the estimated *degree of satisfaction* of g by plan p (a value between 0 and 1); and $P(g, p)$, the expected *success probability* of g using plan p (the probability that plan p will complete and achieve any degree of satisfaction of g). The $S(g, p)$, $P(g, p)$, and $D(g, p)$ components use the agent's domain-specific knowledge. For example, the degree of satisfaction estimate $S(g, p)$ takes into account the effects that the performance of the resources executing plan p have on its outcome, the timeliness of goal achievement (e.g., when travel delays reduce the benefit or percentage of goal achievement), etc. We will discuss specifics of these estimators for the RoboCup Rescue domain in Section 2.2.

A key feature of the OAA approach is including as beliefs, values that act as parameters in operational decision making. For example, such beliefs are used in computing $S(g, p)$, $P(g, p)$, and $D(g, p)$ values. These beliefs start out as initial value settings that reflect the general (unsituated) expertise of competent agents, and they are repeatedly updated by the OAA based on experience. The evolving values allow the OAA to make reasonable decisions (that potentially improve with experience in the current environment) in the absence of organizational directives. The annotations to organizational guidelines include designer-estimated values that are used to seed the OAA's beliefs to values that are close to what the designer assumes the agent should experience when agents are following the parametrized role assignments and priorities contained in their guidelines. Such seeded belief values are also updated by the OAA based on experience, but at a slower rate than unsituated (unseeded) values.

Note that the belief values that are seeded using guideline annotations *can be very different* from what each OAA would experience over time without the use of designed guidelines, as the prioritized role assignments can bias agents to behave differently from the behavior that would emerge from local OAA adaptations. On the other hand, significant deviation of belief values from annotation-seeded settings can help an OAA identify what aspects of designed guidelines are not working as planned and, based on discussions with nearby OAAs, suggest potential local modifications (called *agreements*) that can be negotiated to replace ineffective guidelines.

$U(g, p)$ is simply the discounted benefit less the estimated *opportunity cost* O of using resources R_p for $D(g, p)$, the estimated duration of plan p. The estimated opportunity cost is defined as the sum of the estimated discounted benefit ($DB(g', p')$) values of the best alternative intentions that cannot be chosen due to the use of R_p during the non-preemptive execution of intention (g, p) if intention (g, p) is selected. Two factors make estimating opportunity costs difficult for an OAA: 1) uncertainty of future intention choices that may arise during $D(g, p)$ and 2) resource and goal exchanges (asking for help) that requires an OAA to estimate the non-local effects of such possibilities. Estimating opportunity cost reasonably when agents can exchange resources and goals requires a potential-benefit model that includes, in addition to information about the agent itself, information about the potential activities of other agents that could exchange resources or goals with the agent (and, transitively, the estimates of potential activities of those agents must include further exchange possibilities, and so on). We will discuss these issues shortly (in Section 2.2) and our approach to maintaining an OAA's potential-benefit model without significant communication.

Multiple intentions can be executed concurrently. An action fails when it cannot be completed, such as through the loss of a resource, in which case the plan containing that action is terminated. If sufficient utility can still be achieved, the plan template for the goal with a terminated plan can be re-instantiated with different resource bindings or an alternate plan template can be considered. A plan fails when, after it is finished, the goal is not achieved. Achievement of a goal fails when all reasonable plans to achieve it have failed.

2.1 Using Guidelines to Bias Operational Decisions

When organizational guidelines $\mathcal{G}$ have been provided (an example RoboCup Rescue guideline is shown in Figure 1), the OAA's operational decision engine incorporates the parametrized role priorities rp contained in the guidelines by using the following *biased utility* $BU(g,p)$ calculation in place of $U(g,p)$ (Equation 1b). Formally, each guideline e in $\mathcal{G}$ consists of the following:[3]

$$e = \langle \mathcal{PR}^*, \mathcal{A}^* \rangle \tag{2a}$$

$$\mathcal{PR} = \langle r, rp, \langle pn, pv \rangle^* \rangle \tag{2b}$$

$$\mathcal{A} = \langle a, \langle an, av \rangle^* \rangle \tag{2c}$$

Each $\mathcal{PR}$ is a parametrized role assignment, consisting of a role name r, the organizational priority rp (a positive or negative number) associated with performing the parametrized role, and parameters pv that delimit the role. $\mathcal{A}$ is an annotation, consisting of an expectation name a and attribute name an and value av pairs conveying the expectation.

The biased utility of an intention $BU(g,p)$ given guidelines $\mathcal{G}$ is then:

$$OB(g) = \sum_{e \in \mathcal{AG}(g)} rp(e) \tag{3a}$$

where $\mathcal{AG}(g) \equiv \{e \in \mathcal{G} \text{ s.t. } e \models g\} \subset \mathcal{G}$ and $(i \models j)$ means that guideline i applies to goal g

$$BDB(g,p) = (1-w) \times \big(DB(g,p)\big) + \big(w \times \big(DB(g,p) \times OB(g)\big) \tag{3b}$$

$$BU(g,p) = BDB(g,p) - O(R_p, D(g,p)) \tag{3c}$$

$OB(g)$ (Equation 3a) is the total *organizational bias* for achieving goal g, computed by summing the parametrized-role priorities, rp, from all guidelines that pertain to g. The organizationally biased discounted-benefit value $BDB(g,p)$ (Equation 3b) blends the original discounted benefit value $DB(g,p)$ (from Equation 1a) with an $OB(g)$-biased value according to the *organizational proclivity weight* w (a value between 0 and 1) that controls the influence that organizational guideline biases have over the OAA's unbiased (self) decisions. The biased utility value $BU(g,p)$ (Equation 3c) is simply the biased discounted-benefit value less the estimated opportunity cost O (the same value as was used in Equation 1b).

In large organizations, a distribution of different degrees of organization acceptance (w weights) may have advantages, as some agents will more aggressively explore activities deemed important from their own (skeptical) perspective than others [1]. Furthermore, adjusting the level of an agent's organizational acceptance dynamically, based on

[3] We use S^* to denote a (possibly empty) set $S^* \equiv \{s_1, s_2, \ldots\}$ of elements of type S. For instance, $\mathbb{N}^*$ is a set composed of zero or more natural numbers.

its estimate of how well the organization is functioning, can also be beneficial. In our original OAA paper [2], we showed experimental results of varying w. In the experimental results shown in this paper, a constant w value of 0.8 (fairly strong organizational bias) was maintained in order to show the effects of appropriate and inappropriate organizational guidelines.

2.2 Operational Decisions in RoboCup Rescue

We have evaluated our OAA architecture and approach using the fire-extinguishing portion of RoboCup Rescue, a commonly used multi-agent environment for distributed resource-allocation problems [10]. Our focus has been on *call center* agents that direct the *fire brigade* resources under their control to extinguish as many fires to important buildings as quickly as possible. The objective is to minimize the total importance-weighted damage to buildings. In the evaluations discussed in this paper, four OAA call centers fight fires in the city of Kobe using 24 fire brigades. A call center can use its fire brigades to execute plans to achieve its own goals of extinguishing a building fire, and it can request temporary use of fire brigades from other call centers to use in its plans. Unless directed otherwise by its organizational guidelines, a call center coordinates with other centers to avoid sending redundant fire brigades to a fire (by recalling its brigades using a highest estimated utility protocol). A call center can agree to loan fire brigades temporarily to another call center. Finally a call center can agree to grant open-ended authority to direct some of its fire brigades to another call center. This more permanent (open-ended) resource transfer remains in effect until revoked by the granting call center.

Estimators Specific estimator functions, $S(g, p)$, $P(g, p)$, and $D(g, p)$, for call center intentions in the RoboCup Rescue domain had to be developed. We describe them here.

The plan duration estimator $D(g, p)$ for an ExtinguishFire goal considers the current state of the building fire associated with goal g and the specific fire brigades being considered for plan p. This estimation uses experience-updated *estimated ExtinguishFire plan duration* belief values (Table 1) to estimate the time required to actually fight the fire, plus travel-time estimates for getting the fire brigades to the building. If the brigades are under the direct control of the call center, the current location of brigades is used in estimating travel time. If the call center has to request fire brigades from a nearby call center, the travel-time estimates are based on experience-updated averages for brigades that have been borrowed from that call center in the past (since the current location is unknown by the requesting center). As we have noted, experience-updated belief values start out as initial values when guidelines have not been provided or are seeded with annotated values when guidelines have been given.

The degree of satisfaction estimator $S(g, p)$ for an ExtinguishFire goal must take into account the current size of the fire, the rate at which the fire brigades allocated to p can diminish the fire, the delay in getting the fire brigades to the fire's location, and so on. Similarly, the success probability estimator, $P(g, p)$ for an an ExtinguishFire goal must estimate the likelihood that the plan p, when fully executed, does not extinguish the fire (e.g., some fires are particularly difficult to extinguish) or that one of the plan's actions fails (e.g., no water is available near the fire).

estimated ExtinguishFire duration the number of time steps required to extinguish a fire of a given size in a building of a given volume using n fire brigades (a table of values)

fire brigade quantity the minimum and maximum number of fire brigades that are effective in fighting a fire of a given size in a building with a given volume (a table of values)

attainable benefit the attainable benefit per time step in the OAA's neighborhood when n fire brigades are used (a vector of values)

time window the length of time an OAA needs to look into the past in order to obtain good statistics

attainable benefit initial-value weight controls how aggressively perceived data is combined with initial values when the OAA does not have a full time window of historical data.

deviations confidence level the confidence level that the OAA should consider when performing the non-parametric hypothesis test (i.e., the confidence that the collected observations allow the OAA to state that most likely the current dynamics of the environment are indeed significantly different from expectations)

Table 1. Example belief values used as parameters in RoboCup Rescue operational decision-making

Opportunity cost In order to estimate opportunity costs when resources and goals can be exchanged among call centers, each OAA first needs to estimate, based on its observation history, the average discounted benefit per time step that can be obtained in its neighborhood (i.e., call centers that are close enough to make a resource or goal exchange useful) when using k fire brigades. Let $\mathrm{ODB}_n(k)$ be such an estimate; specifically, the expected discounted benefit that can be obtained by call center n when using k fire brigades. Such estimates are transmitted to nearby call centers (and updated whenever they change significantly) along with the current number of uncommitted fire brigades it controls. Using its own estimates of the achievable benefit per time step and the estimates provided by nearby call centers, every call center n then computes the *maximum attainable* benefit per time step, $\mathrm{ODB}^*(k)$, can that be obtained whenever k fire brigades are used. For instance, let us suppose we are interested in optimally distributing $k \in [1 \ldots R$ fire brigades across N call centers, where the optimality criteria is defined in terms of the maximum attainable expected benefit per time step. The naive approach for computing such a quantity would require the evaluation of all combinations of how to allocate r resources over N agents, for all values of $r \in [1 \ldots R]$. This would thus demand the evaluation of

$$\sum_{r=[1 \ldots R]} \binom{r}{N} = \binom{R+1}{N+1}$$

possibilities, which is factorial in its complexity. In Algorithm 1 we present a dynamic programming algorithm that computes such a quantity much more efficiently. Algorithm 1 runs in $\Theta(NR^2)$, where N is the number of known neighbors and R is the maximum number of resources (i.e., fire brigades) being considered, and returns the maximum attainable benefit per time step when using any given number of resources.

The opportunity cost of executing a plan that utilizes k resources for an estimated t time steps is then $t \times (\mathrm{ODB}^*(\hat{f}) - \mathrm{ODB}^*(\hat{f} - k))$, where $\hat{f}$ is the agent's estimate of

Algorithm 1 Maximum attainable benefit per time step when optimally allocating n resources across agents in the organization.

Let $A = [1 \dots N]$ be the (arbitrarily ordered) list of N agents (e.g., call centers) in the system;
Let R be the total number of resources (e.g., fire brigades) in the system;
Let $\mathrm{ODB}_n(k)$ be the expected benefit per time step of agent $n \in A$ when using k resources;
Let $\mathrm{ODB}^*_n(k)$ be the maximum attainable benefit per time step when optimally allocating k resources over agents $[n \dots N] \subset A$;
Output: $\mathrm{ODB}^*(\cdot)$, the maximum attainable benefit per time step when allocating any given number $k \in [1 \dots R]$ of resources across agents $A = [1 \dots N]$.

```
for k = 0 to R do
  ODB*_N(k) = ODB_N(k)
end for
for n = N − 1 downto 1 do
  for k = 0 to R do
    ODB*_n(k) = max_{r=0...R} ( ODB_n(r) + ODB*_{n+1}(k − r) )
  end for
end for
Return ODB*(·) ≡ ODB*_1(·)
```

the total number of unutilized resources available (within practical range) for its use. This $\hat{f}$ estimate can be computed by adding the OAA's own amount of unutilized fire brigades and the quantities provided by nearby call centers.

Annotated guidelines Call centers perform two roles. The first role is extinguishing fires by directing fire brigades to fight them. The second role is granting control of their fire brigades to other call centers, either temporarily or more permanently. We designed organizational guidelines for the four call centers, where each call center is assigned responsibility for a non-overlapping region of the city. These regional-responsibility guidelines do not change the fire observability range of call centers.

An example of an annotated guideline given to a call center is shown in Figure 1. The guideline specifies that the agent is responsible for extinguishing fires in the given region with an organizational priority of 10. When the biased utility, $BU(g, p)$, is calculated for an ExtinguishFire goal g in this region, the priority 10 is included in the organizational priority calculation, $OB(g)$, for achieving goal g,

Two different organization designs were developed, each intended to operate over a different range of environmental conditions. One of the hypotheses we wanted to evaluate was that call centers with annotated guidelines appropriate to their environment can outperform call centers operating without guidelines in that same environment. Another hypothesis was that call centers with guidelines designed for a different range of conditions could perform worse than call centers without guidelines, and also that OAA mechanisms can recognize when that is the case and adapt away from the ineffective guidelines.

```
<guideline role="ExtinguishFires" description="region">
  <region priority="10" shape="rectangle">
    <start-point x="0.0" y="0.0" />
    <end-point x="0.5" y="0.6" />
  </region>
  <annotation>
    <task-expectations>
      <mean max="3" />
    </task-expectations>
    <performance-expectations>
      <success rate="0.8" />
    </performance-expectations>
            ...
  </annotation>
</guideline>
```

Fig. 1. An annotated guideline specifying a prioritized ExtinguishFires role responsibility

3 Detecting Expectation Deviations

In order to detect deviations from annotations, it is necessary for call centers to compute statistics over observable data as they interact with the environment. Such statistics are used to test whether the predictions made by the annotations are valid in the current context, and thus whether the guidelines are still applicable. In the case of region guidelines, task annotations describe the expected total number F of active fires within the region at any given moment in time. Testing whether the *actual* average observed number of active fires is generally coherent with this expectation is not easy, since the random variable F depends on a series of factors: the arrival rate of fires; the way in which fires spread to neighboring buildings [9]; and the rate at which fires are put out by the fire brigades, which itself depends on their individual strategies and priorities. Although some of these factors, like the arrival rate of fires, can reasonably be assumed to be Poisson-distributed, the same cannot be said for F itself. Since we do not know how F is distributed, most of the classic statistical hypothesis tests cannot be applied. In order to allow for a general statistical test of relevance for guidelines, we have decided to apply the Wilcoxon signed-rank non-parametric hypothesis test. This test requires that region annotations specify only the expected (mean) number of active fires per time step.[4] We note that although it is trivial to obtain samples of this quantity, such samples are not independent, since the number of fires at any given time step is directly correlated with the number of fires in the previous time step. Therefore, this specific assumption made by the hypothesis test is not strictly obeyed; nonetheless, we have noticed that in practice the independence assumption can be reasonably made and does not seem to alter the capability of the test to detect significant deviations from the annotations. Also, for a variety of possible generating distributions of F, the test seems to perform well, even with as few as 20 samples. Finally, we note that because the actual

[4] More complex expectation annotations, requiring a correspondingly complex detection model, could be used (e.g., annotation values that specify the expected exponential growth of spreading fires as the arrival rate varies).

average number of active fires per time step might be non-stationary and vary due to a variety of factors, we consider only a finite window of past samples when computing the required statistics. Also, we note that although we currently execute the hypothesis test under a 95% confidence level, this quantity can be made to depend on how rigorous each agent decides to be when judging whether a guideline is still applicable.

4 Forming and Negotiating Agreements

Whenever a call center detects a deviation from the guideline annotations, a process is invoked that considers forming long-term agreements between OAAs. In this paper, we consider pair-wise agreements that increase the expected long-term social benefit by having some call centers grant sole control of n fire brigades to another call center. By eliminating repeated requests for assistance, a chronically busy call center can increase goal satisfaction and success probability values by eliminating the delay and uncertainty associated with asking for help.

When a call center detects that it is performing a higher than expected rate of assistance requests, it generates goals for forming open-ended fire brigade control-transfer agreements with nearby call centers. It asks a set of call centers $\{A_i|i = 1, ...m\}$, each contributing n_i fire brigades, such that $\sum_{i=1}^{m} n_i = n$. Whenever a call center needs to evaluate the utility of accepting or rejecting such a goal, it needs to take into account the previously estimated long-term benefit of directly controlling n extra fire brigades and the expected long-term loss incurred by the call centers providing the fire brigades.

The estimated long-term marginal utility increase of directly controlling n extra fire brigades is obtained using the estimated weighted benefit increase of achieving the same goals, but now using n additional, directly controlled fire brigades (using the OAA's observation history):

$$U^{+}(n) = \sum_{g with p_R} B(g) * (S(g, p_{R+n}) * P(g, p_{R+n}) - S(g, p_R) * P(g, p_R)) \quad (4)$$

B, S, and P were introduced earlier in Equation 1a. Plan p_{R+n}, is a modified version of p_R with n additional directly-controlled fire brigades. If p_R requests k fire brigades from neighbors, then plan p_{R+n} only requires $max(k - n, 0)$ fire brigades from neighbors. Directly controlling additional fire brigades tends to increase both the satisfaction degree S and the success probability P, resulting in a marginal utility increase $U^{+}(n)$ reflecting the benefit of an agreement of controlling n additional fire brigades. Since the impact of each additional fire brigade may not be the same, we must calculate $U^{+}(n)$ for all reasonable values, $1..n_{max}$, where n_{max} is determined based on the deviation from expectation annotations. The greater the deviation, the larger max should be. In a similar way, each neighboring call center, upon receiving a request for donating n_i fire brigades, estimates the long-term marginal utility decrease caused by having n_i fewer directly controlled fire brigades $U^{-}(n_i)$.

An optimal solution is a set of m agreements, each requesting n_i fire brigades from neighbor A_i, $\sum_{i=1}^{m} n_i = n$, that maximizes the social welfare:

$$U^{+}(n) - \sum_{i=1}^{m} U^{-}(n_i) \quad (5)$$

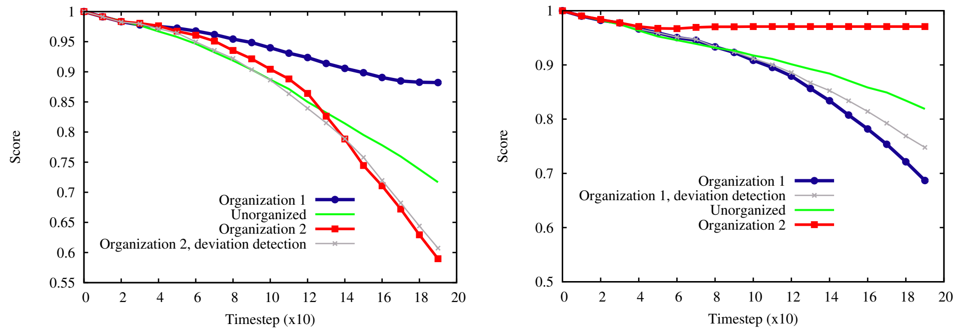
In general, finding the optimal solution is similar to the winner determination problem in combinatorial auctions: the complexity is NP-hard. However, given the small number of nearby call centers and n_{max}, finding a solution is computationally feasible. Once the value of n_{max} is determined, the call center calculates $U^{+}(n)$ for $n = 1..n_{max}$. It also sends out requests to nearby call centers requesting their marginal utility decrease for providing $1..n$ fire brigades. Upon receiving this information, the requesting call center conducts a search to find the value of n, a set of neighbors A_i and n_i, $\sum_{i=1}^{m} n_i = n$ that maximizes the social welfare as described in Equation 5. When a solution is found, the call center asks each call center A_i to accept the agreement providing n_i fire brigades. Call center A_i would accept this request if its fire brigade status has not changed in a way that worsens the marginal utility decrease estimates $U^{-}(n_i)$ it provided. Otherwise A_i may reject this request and provide revised estimates to the proposing call center for consideration.

5 Experimental Analysis

We evaluated the performance of our OAA call centers by assessing two hypotheses: 1) following guidelines whose annotations are coherent with the environment improves performance and 2) following guidelines whose annotations are not coherent with the environment decreases the performance to levels that can be even lower than when the agent operates without guidelines. This latter hypothesis, if shown to be true, attests to the importance of being able to detect deviations from the expectations contained in guideline annotations and to stop following those inappropriate guidelines as soon as possible. In order to demonstrate that a poor-performing design is not deficient in every setting, it is very important to show that the appropriate design performs better in its designed setting but worse in settings for which other organizations were designed.

To be fair, we developed highly competent call-center agents that make skillful operational decisions to extinguish fires without organization. We wanted to ensure that supplied guidelines only inform and limit an OAA's operational decisions; they do not extend the agent's abilities or purpose. We believed that appropriately organized agents would function better than unorganized agents, which must perform unguided consideration of potential agent activities and explicitly coordinate them. From a research perspective, this high bar on operational competency highlights the effect of organization (good or bad), without any hidden transfer of capabilities or expertise by organizational directives.

In order to test the above-mentioned hypotheses we developed annotated guidelines for two different organization designs, each intended to operate over a different range of environmental conditions. We then evaluated the joint performance of the four OAA call centers in two scenarios, A and B, each one corresponding to different patterns of fire arrival. In Scenario A, fires arrive uniformly throughout the city, while in Scenario B the same number of fires arrive, but they are more heavily concentrated in two relatively small regions of the city. In the following experiments, performance is measured using a score that represents the fire damage in the city; scores closer to 1 indicate that the city is mostly intact and scores closer to 0 indicate that the city is almost completely burned out.



(a) Scenario A (b) Scenario B

Fig. 2. RoboCup Rescue performance

Organization 1 contains guidelines that inform each call center of the region it should preferentially cover, and does so in a way that the city is evenly divided into four non-overlapping responsibility regions. A total of 24 fire brigades is evenly allocated to the four call centers. Organization 2, on the other hand, contains guidelines that assigns responsibility regions of different sizes so that number of fire brigades controlled by each center corresponds to the non-uniform fire arrival rate of Scenario B. Both of these organization designs bias agents so that they have a higher preference for allocating fire brigades to fight fires in their own regions, although borrowing is also allowed.

Our experiments show that if an appropriate organization design is used and the guidelines bias agents to act preferentially in their own regions, avoiding fighting fires in other regions, the performance of call centers is superior to what is attainable when call centers act solely according to an unbiased operational strategy. This is true, for instance, in the case when pattern of fire arrivals is uniform in space (Scenario A) and Organization 1 is used, since the guidelines inhibit call centers from greedily moving their fire brigades around the city and fighting large far-away fires without considering the possibility that new fires might ignite in their own regions of responsibility. These results are shown in Figure 2(a). Figure 2(a) also shows how call centers which use inappropriate guidelines (in this case, from Organization 2) have lower performance than centers that act based solely on their unbiased operational strategy.

Figure 2(b) presents the performance of centers again following guidelines from Organization 1 or 2, or acting without any guidelines, but this time under a non-uniform fire arrival pattern (Scenario B). In this case, it can be observed that because Organization 1 was designed under the assumption of uniform spatial distribution, the performance of call centers following its guidelines is even worse than if it was operating without them. Organization 2, on the other hand, is coherent with the environmental dynamics of Scenario B and thus successfully biases agents in a way that their fire brigades can be efficiently allocated, and kept within, the regions of the city in which a higher number of fires is occurring.

In the experiments discussed thus far, we disabled deviation detection so that the call centers would continue following inappropriate guidelines. When we re-enabled

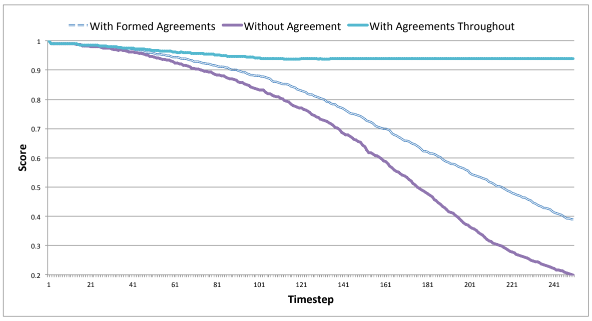


Fig. 3. Effect of agreements on performance when guidelines are inappropriate

Table 2. Fire brigades transferred under the negotiated agreements

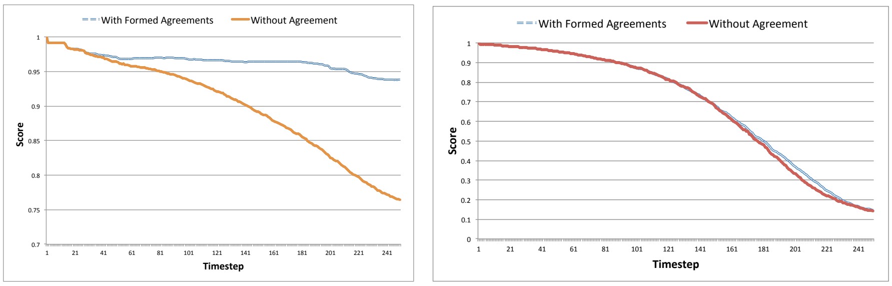
Time	Transferred fire brigade	From call center	To call center
83	FB-17	CC-SW	CC-NE
83	FB-18	CC-SW	CC-NW

deviation detection, call centers quickly recognized (in time step 20) when their annotated guidelines were inappropriate and stopped following them. This performance is also shown in the two figures.

5.1 Agreement Experiments

To analyze the benefit of agreement formation on performance when inappropriate guidelines are abandoned, we used Organization 1 (appropriate for a uniform distribution of fires) with the non-uniform fire setting, Scenario B. When agreement formation is enabled (shown in Figure 3), once a call center detects a deviation from guidelines, it looks for nearby call centers with which to form beneficial long-term agreements. Unless revoked, these negotiated open-ended agreements transfer fire brigades from call centers that are experiencing fewer fires in their guideline regions to centers with more fires. The negotiated agreements offset the effect of the inappropriate organization guidelines and enable the call centers achieve a better overall performance score by reducing the operational delay associated with repeated requests for help. The specific agreements that were negotiated in our experiment is shown in Table 2.

To understand how the call centers would perform if they were using the negotiated agreements from the outset, we ran the same scenario with those agreements in place at the start, referred to as ”With Agreements Throughout.” The results show that using the agreements from the beginning improves the performance significantly over the experiments where the call-center agents are started with inappropriate organizational guidelines, as shown in Figure 3. Forming agreements improves the performance over proceeding without them, but the call centers can never recover from the fires that spread before they recognize inappropriate guidelines and form agreements to replace them.



(a) Moderate spreading fires (rate: 0.8) (b) Fast spreading fires (rate: 1.0)

Fig. 4. Using agreements under different fire-spreading rates

The effect of agreement formation on the performance also depends on the system load (the number of fires and how fast they spread). Figure 4 shows the results in the same scenario as described above, but with more fires. As shown in Figure 4(b), when fires are spreading very fast (at a rate of 1.0), agreements only improve the organization performance slightly from time 150 to time 220. This is because fast-spreading fires that arise due to following inappropriate guidelines cannot be controlled by the limited number of fire-brigade resources no matter how carefully they are transferred and used. In contrast, when the fire-spread is more moderate (at a rate of 0.8), forming agreements enables the centers maintain a very good performance score, as shown in Figure 4(a).

6 Summary and Future Work

We presented a decision-making architecture for an organizationally adept agent (OAA) and its instantiation and use in RoboCup Rescue scenarios. We showed how an OAA can adjust its operational decisions to conform with organizational guidelines, if they are made available. We also showed how an OAA uses expectations contained in guideline annotations to inform its operational decision making and to identify deviations from the task and environmental assumptions that were used when the guidelines were designed (and to stop following guidelines deemed to be inappropriate). Finally, we showed how OAAs can form local agreements to use in place of inappropriate guidelines.

The RoboCup Rescue results indicate that the OAA call centers are fairly competent in their ability to operate without organizational directives, but that appropriate organizational guidelines can improve performance further. We do not know of prior work where two organizational designs developed for different environmental settings have been analyzed comparatively, where the appropriate design performs better than unorganized agents in its designed setting but worse in the setting for which the other organization was designed (and vice versa). Yet, we have barely scratched the surface of complex organizationally adept behavior. More complex, larger scale, heterogeneous organizations are a next step, as is exploration of more complex operational tasks and correspondingly more advanced organizational guidelines. Generating agreements from scratch is another future-research direction. The OAA operational decision-making mechanisms and interaction protocols that we developed in order to have a highly competent agent when operating without organizational directives are appropriate for agents operating in uncertain environments—even if they must operate without organization.

Acknowledgment This material is based in part upon work supported by the National Science Foundation under Award No. IIS-0964590. Any opinions, findings, conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

1. D. D. Corkill. *A Framework for Organizational Self-Design in Distributed Problem-Solving Networks*. PhD thesis, University of Massachusetts Amherst, Amherst, Massachusetts 01003, Feb. 1983.
2. D. D. Corkill, E. Durfee, V. R. Lesser, H. Zafar, and C. Zhang. Organizationally adept agents. In *12th International Workshop on Coordination, Organizations, Institutions and Norms in Agent Systems* (COIN@AMASS-2011), pages 15–30, Taipei, Taiwan, May 2011.
3. D. D. Corkill and V. R. Lesser. The use of meta-level control for coordination in a distributed problem-solving network. In *Proceedings of the Eighth International Joint Conference on Artificial Intelligence*, pages 748–756, Karlsruhe, Federal Republic of Germany, Aug. 1983.
4. E. H. Durfee and Y. pa So. The effects of runtime coordination strategies within static organizations. In *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence*, pages 612–618, Nagoya, Japan, Aug. 1997.
5. B. Horling, B. Benyo, and V. Lesser. Using self-diagnosis to adapt organizational structures. In *Proceedings of the Fifth International Conference on Autonomous Agents*, pages 529–536, Montreal, Canada, June 2001.
6. B. Horling and V. Lesser. A survey of multi-agent organizational paradigms. *Knowledge Engineering Review*, 2005.
7. B. Horling and V. Lesser. Using quantitative models to search for appropriate organizational designs. *Autonomous Agents and Multi-Agent Systems*, 16(2):95–149, 2008.
8. J. F. Hübner, J. S. Sichman, and O. Boissier. Developing organised multi-agent systems using the MOISE+ model: Programming issues at the system and agent levels. *International Journal of Agent-Oriented Software Engineering*, 1(3/4):370–395, 2009.
9. T. Jess. Addressing problem spreading in agent planning and control. Technical Report UM-CS-2012-004, Department of Computer Science, University of Massachusetts Amherst, Amherst, Massachusetts 01003, Mar. 2012.
10. H. Kitano and S. Tadokoro. RoboCup-Rescue: A grand challenge for multi-agent and intelligent systems. *AI Magazine*, 22(1):39–52, 2001.
11. J. G. March and H. A. Simon. *Organizations*. John Wiley & Sons, 1958.
12. T. B. Quillinan, F. Brazier, H. A. Frank, L. Penserini, and N. Wijngaards. Developing agent-based organizational models for crisis management. In *Proceedings of the Industry Track of the Eighth International Joint Conference on Autonomous Agents and Multi-Agent Systems* (AAMAS 2009), pages 45–51, Budapest, Hungary, May 2009.
13. A. S. Rao and M. P. Georgeff. BDI agents: From theory to practice. In *Proceedings of the First International Conference on Multi-Agent Systems* (ICMAS'95), pages 312–319, San Francisco, California, June 1995.
14. M. Sims, D. Corkill, and V. Lesser. Automated organization design for multi-agent systems. *Autonomous Agents and Multi-Agent Systems*, 16(2):151–185, Apr. 2008.
15. A. Staikopoulos, S. Saudrais, S. Clarke, J. Padget, O. Cliffe, and M. D. Vos. Mutual dynamic adaptation of models and service enactment in ALIVE. In *Proceedings of the Third International Models@Runtime Workshop*, pages 26–35, Toulouse, France, Sept. 2008.