

Distributed Search Systems with Self-Adaptive Organizational Setups

Friederike Wall

Universitaet Klagenfurt, Department of Controlling and Strategic Management, Austria

Abstract — This paper studies the effects of learning-induced alterations of distributed search systems' organizations. In particular, scenarios where alterations of the search-systems' organizational setup are based on a form of reinforcement learning are compared to scenarios where the organizational setup is kept constant and to scenarios where the setup is changed randomly. The results indicate that learning-induced alterations may lead to high levels of performance combined with high levels of efficiency in terms of reorganization-effort. However, the results also suggest that the complexity of the underlying search problem together with the aspiration level (which drives positive or negative reinforcement) considerably shapes the effects of learning.

Keywords — Agent-based Simulation, Complexity, NK Fitness Landscapes, Reinforcement Learning.

I. INTRODUCTION

THE organizational setup of distributed search systems is a topic that is investigated in many disciplines, such as control theory, complex systems science or computational organization theory (for extensive reviews cf. [1], [2], [3]). The *coherence* of and the *coordination* within distributed search systems are among the predominant issues in this line of research, where the former is defined in terms of some of the systems' properties (e.g., solution quality) and the latter is concerned with actions and interactions of agents collaborating in a distributed search system [4], [5]. Thus, the key topics of the organizational setup of distributed search systems addressed refer to the appropriate segmentation of the overall search problem into sub-tasks, the way sub-tasks are assigned to agents, and the mechanisms to consolidate the (partial) solutions to sub-tasks into an overall solution. The overall solution should be as satisfactory as possible where its quality is determined on the basis of coherence metrics (e.g., [4], [5], [6]).

Hence, feasible consensus mechanisms, performant algorithms for search and optimization, and the appropriate assignment of tasks are of particular interest in this line of research [1] – in order to contribute to improving results with respect to coherence metrics of relevance. However, this line of research (mostly implicitly) assumes that the designer of a distributed search system decides which of these mechanisms, algorithms and ways of assignment are employed in the organizational setup of the search system. This paper follows an approach that, in a way, can be regarded as complement to the aforementioned line of research: not the designer of a search system is allowed to (exogenously) decide on the systems' organizational setup but the search system's organizational setup evolves endogenously. In particular, we allow for self-adaptation of the search systems' organizational setup, i.e., while searching for better solutions for the overall search problem (during “run-time”) the search system is allowed to change its organization, where changes are based on feedback [7].

The idea of self-adaptive distributed search systems builds on prior

studies which provide evidence that distributed search processes could remarkably benefit with respect to solution quality obtained from inducing organizational dynamics while searching for better solutions – may it be in the organizational setup of collaborating robots or “swarms” of unmanned aerial vehicles or in the organizational design of a firm where managers search for higher levels of firm performance [8], [9], [10], [11]. Apparently, organizational change per se tends to enhance the performance of a search system by inducing a shift towards more exploration, i.e., discovery of new solutions, and less exploitation, i.e., stepwise improvement. However, it is worth emphasizing that these studies employ *merely random-driven* organizational change in the sense that the search systems do not learn which organizational setups are more successful than others.

By investigating the effects of learning-based organizational dynamics, this paper goes a step beyond research studies that employ random-driven organizational changes. *In particular, this paper studies the effects of endowing distributed search systems with some capabilities to learn about their organization's performance and to adapt the organizational setup according to the search systems' performance.* This paper is an extended version of [12] which was presented at the 13th International Conference on Distributed Computing and Artificial Intelligence (DCAI). The extensions predominantly relate to the dimensionality of the search problems under investigation, to the time horizon of simulations, and to a sensitivity analysis with respect to the number of search agents.

It appears to be of particular interest to investigate whether search systems which employ learning-based organizational change outperform systems which make use of random changes in their organizational setup or systems which do not change their setup changes at all. This study intends to provide findings on the relative potential benefits of learning-based organizational dynamics. Since, it is well known that the task environment (in terms of the task complexity) tends to affect the performance of search, this paper particularly controls for the complexity of the search problems by employing an agent-based simulation model which is based on the framework of fitness landscapes [13], [14]. The next section introduces the key elements of the simulation model. Section III gives an overview of the performed simulation experiments. The results are presented in Section IV where, first, an in-depth analysis of some baseline scenarios of organizational change modes for different levels of complexity are provided. Second, a sensitivity analysis is presented which puts particular emphasis on the need for coordination within the search system where this need is considerably affected by the number of search agents who carry out sub-tasks.

II. OUTLINE OF THE SIMULATION MODEL

The study employs an agent-based simulation model which captures two intertwined adaptive processes: In (1) the short-term, search agents seek to find superior solutions for the search problem. The quality of a solution is measured on the basis of system's overall performance level

achieved. We model search agents to operate on NK fitness landscapes [13], [14]. In (2) the mid-term, the search systems are allowed to adapt major features of their organizational setup. Changes are driven by reinforcement-learning, which is based on performance enhancements achieved. A schematic flow-chart of key features of the simulation model is displayed in Figure 1.

A. Short-Term Adaptive Search for Higher Levels of Performance

The study employs the framework of NK-fitness landscapes, which were originally introduced in the domain of evolutionary biology [13]. An advantage of NK fitness landscapes is that they easily allow for controlling the complexity of the underlying search problem [15].

1) Search Problem

In each time step t of the observation period T , the search systems face an N -dimensional binary search problem, i.e., they seek for a superior configuration $\mathbf{d}_t = (d_{1t}, \dots, d_{Nt})$ (with $d_{it} \in \{0, 1\}$, $i = 1, \dots, N$) out of a set of possible solutions, which is given by 2^N different binary vectors. Each of the two states $d_{it} \in \{0, 1\}$ contributes with C_{it} to fitness $V(\mathbf{d}_t)$ of the search system or – in other words – to the performance achieved by the search system. According to the NK framework, C_{it} is randomly drawn from a uniform distribution with $0 \leq C_{it} \leq 1$.

The parameter K captures the complexity of the underlying search problem [15]: In particular, fitness contribution C_{it} might not only depend on the single choice d_{it} but also on a number of other choices where K indicates the number of d_{jt} , $j \neq i$ that affect C_{it} in addition to d_{it} . In case of no interactions K is 0, and K equals $N - 1$ for the case of maximum interdependence. Hence, contribution C_{it} results from

$$C_{it} = f_i(d_{it}; d_{i_1t}, \dots, d_{i_Kt}) \quad (1)$$

with $\{i_1, \dots, i_K\} \subset \{1, \dots, i-1, i+1, \dots, N\}$.

The overall performance V_t achieved by the search system in period t is computed as the normalized sum of contributions C_{it} :

$$V_t = V(\mathbf{d}_t) = \frac{1}{N} \sum_{i=1}^N C_{it} \quad (2)$$

2) Agents and their Choices

The search for higher levels of performance V_t is collaboratively performed by M search agents. In particular, the N -dimensional search problem is partitioned into M disjoint partial problems, and each of these sub-problems is exclusively delegated to one search agent r , $r = 1, \dots, M$. The partial search problems are of equal size with

$N_r = \frac{N}{M} \forall r = (1, \dots, M)$. From the perspective of search agent r , the search problem is segmented into a partial search vector \mathbf{d}_t^r which contains the choices which are under the search agent r 's primary control and into a partial vector $\mathbf{d}_t^{r, res}$, which captures the residual choices that the other search agents $q \neq r$ are in charge of. However, with cross-segment interactions among the sub-problems, choices of agent r might affect the contribution of the other agents' choices to overall performance, and vice versa.

In each time step t , a search agent seeks to identify the best configuration for the “own” choices \mathbf{d}_t^r assuming that the other agents do not alter their choices made in $t-1$. Each agent r randomly discovers

two alternatives in addition to the status quo choice $\mathbf{d}_{t-1}^{r,*}$, where, as compared to the status quo, one alternative (named $a1$) differs in one and the other one (labelled $a2$) differs in two single choices d_{it} . In consequence, in time step t , agent r has three options to choose from,

i.e., keeping the status quo or switching to $\mathbf{d}_t^{r,a1}$ or $\mathbf{d}_t^{r,a2}$. Which of these options appears to be favorable from the search agent's perspective depends on the agent's “objective” P_t^r . This objective is shaped by parameter α^r which defines the extent to which the residual part of the decision problem is considered in addition to the “own” partial search problem. The objective function can be formalized by

$$P_t^r(\mathbf{d}_t^r) = P_t^{r, own}(\mathbf{d}_t^r) + \alpha^r \cdot P_t^{r, res} \quad (3)$$

with $V_t = V(\mathbf{d}_t) = \frac{1}{N} \sum_{i=1}^N C_{it}$ and $P_t^{r, res} = \sum_{q=1, q \neq r}^M P_t^{q, own}$ where $p = \sum_{s=1}^{r-1} N_s$ for $r > 1$ and $p = 0$ for $r = 1$.

However, the agents' ex ante evaluations of alternatives do not necessarily need to be perfect. In particular, agents might misjudge the options' contributions to objective $P_t^r(\mathbf{d}_t^r)$. This may not only be an unintentional shortcoming of, e.g., agents' information processing capabilities but may also be intentionally induced: Some evidence suggests that imperfect information on the fitness (performance) of options could increase the effectiveness of search processes (e.g., [16], [17]). Previous research shows that false-positive evaluations of options increase the diversity of search. As a consequence, there is a chance to end situations of inertia induced by sticking to a local peak and to reach basins of attractions for higher levels of fitness. Hence, intentionally or not, our agents may eventually be endowed with slightly distorted information about the performance of options. Distortions are captured by adding error terms as exemplarily shown in Eq. (4):

$$\tilde{P}_t^{r, own}(\mathbf{d}_t^r) = P_t^{r, own}(\mathbf{d}_t^r) + e^{r, own}(\mathbf{d}_t^r) \quad (4)$$

For the sake of simplicity, distortions are modeled as relative errors added to the true performance (for other functions see [16]). The error terms follow a Gaussian distribution $N(0; \sigma^2)$ with expected value 0 and standard deviations $\sigma^{r, own}$ and $\sigma^{r, res}$. Variances are assumed to be the equal for all search agents $r = 1, \dots, M$ and stable over time (if not altered by self-adaptation as described subsequently); all error terms are assumed to be independent from each other.

Apart from the search agents, the model captures a kind of “central agent” whose role is a twofold: (1) In the short-termed adaptive search, the central agent could – depending on the particular mode of coordination – intervene in the selection of choices. (2) In the mid-term, the central agent assesses performance enhancements and “learns” about successful organizational setups by reinforcement. The next section provides more details on the central agent's roles.

B. Mid-Term Adaptation of the Organizational Setup based on Reinforcement Learning

The very core of this study is related to learning on the performance contributions of a search system's organization and, eventually, altering the organizational setup accordingly. The following two subsections describe the modelled mode of reinforcement learning as well as

the dimensions of the organizational setup which may be subject to organizational change.

1) Mode of Reinforcement Learning.

In each T^* -th time step, the central agent faces an L -dimensional decision problem related to the L dimensions of the organizational setup which can be altered. In particular, the central agent chooses a setup

$\mathbf{\ddot{O}}_t = (a_1(t), \dots, a_L(t))$ of alternatives $a_l \in A_l$ for all $l = 1, \dots, L$ and with $|A_l|$ giving the number of alternatives a_l in set A_l .

The model employs a simple mode of reinforcement learning (for overviews see [18], [19]) based on statistical learning, i.e., a generalized form of the Bush-Mosteller model [20], [21]: In every T^* -th period, the propensities of choices are updated based on the stimuli resulting from the evaluation of the outcome (performance effects) achieved under the regime of prior choices of the organizational setup.

The outcome ω of configuration $\mathbf{\ddot{O}}_t$ is given by the maximal relative performance enhancement which is achieved within the last T^* periods of the adaptive walk, i.e.,

$$\varpi(\mathbf{\ddot{O}}_t) = \max \left[\frac{V_{t\tilde{t}} - V_{t-T^*}}{V_{t-T^*}}, \tilde{t} = 1, \dots, (T^* - 1) \right] \quad (5)$$

The evaluation of the outcome can be regarded positive (1) or negative (-1), where the assessment of ω depends on whether or not it, at least, equals an aspiration level v . Hence, the stimulus $\tau(t)$ results as

$$\tau(t) = \begin{cases} 1 & \text{if } \varpi(\mathbf{\Phi}_t) \geq v \\ -1 & \text{if } \varpi(\mathbf{\Phi}_t) < v \end{cases} \quad (6)$$

Let $p(a_l, t)$ denote the probability of an alternative within dimension l of organizational setup to be chosen at time t (where $0 \leq p(a_l, t) \leq 1$ and $\sum_{a_l \in A_l} p(a_l, t) = 1$); $a_l(t)$ denotes the option out of set A_l which is implemented at time-step t . The probabilities of options $a_l \in A_l$ are updated based to the following rule, where 1 ($0 \leq \lambda \leq 1$) reflects the reinforcement strength [21]:

$$p(a_l, t+1) = p(a_l, t) + \begin{cases} \lambda \cdot \tau(t) \cdot (1 - p(a_l, t)) & \text{if } a_l = a_l(t) \wedge \tau(t) = 1 \\ \lambda \cdot \tau(t) \cdot p(a_l, t) & \text{if } a_l = a_l(t) \wedge \tau(t) = -1 \\ -\lambda \cdot \tau(t) \cdot p(a_l, t) & \text{if } a_l \neq a_l(t) \wedge \tau(t) = 1 \\ -\lambda \cdot \tau(t) \cdot \frac{p(a_l, t) \cdot p(a_l(t), t)}{1 - p(a_l(t), t)} & \text{if } a_l \neq a_l(t) \wedge \tau(t) = -1 \end{cases} \quad (7)$$

After the probabilities are updated according to Eq. 7, the organizational setup $\mathbf{\ddot{O}}_t$ to be implemented from time steps $t+1$ to $t+T^*$ is determined randomly based on the updated probabilities.

2) Organizational Setup.

The vector of the organizational setup $\mathbf{\ddot{O}}$ is modelled to be three-dimensional, i.e., $L = 3$. Within each dimension, three options are given (i.e., $|A_l| = 3 \forall l \in \{1, 2, 3\}$). These dimensions relate to (see also Table I):

1. The objective of the search agents as controlled by parameter α^r in Eq. 3. For the sake of simplicity, α^r is modelled to be the same

- for all search agents $r = 1, \dots, M$. In the following we skip index r .
2. The precisions of ex ante-evaluations made by search agents and by the central agent as given by $\sigma^{r,own}$ and $\sigma^{r,res}$, and σ^{cent} , respectively.
3. The mode of coordination as selected out of three alternatives: (a) “decentralized”: without intervention from the central agent or any coordination which other search agents, each search agent autonomously decides on the “own” partial choices \mathbf{d}_t^r ; (b) “lateral veto”: the search agents inform each other about their preferences and are endowed with mutual veto power; (c) “centralized”: each search agent informs the central agent about the two most preferred options from \mathbf{d}_{t-1}^{r*} , $\mathbf{d}_t^{r,a1}$ and $\mathbf{d}_t^{r,a2}$; the central agent chooses that combination of preferences which promises the highest overall performance V .

III. SIMULATION EXPERIMENTS AND PARAMETER SETTINGS

In the simulation experiments, after a performance landscape is generated, the initial organizational setup (i.e., vector $\mathbf{\ddot{O}}_{t=0}$) of a search system is determined randomly with uniform probabilities $p(a_l, t = 0)$ out of the options in each dimension l as introduced above and summarized in Table I. Next, the search systems are placed randomly in the performance landscape. Then, over an observation time T of 500 periods, the search systems are observed while searching for higher levels of performance. In each T^* -th period, probabilities are updated and organizational configurations are (eventually) altered (cf. Sec. B). Fig. 1 displays the key events during a simulation experiment capturing learning-based adaptation of the organizational setup.

TABLE I
PARAMETER SETTINGS

Parameter	Values / Types
Observation period	$T = 500$
Number of choices	$N = 12$
Interaction structures	block-diagonal ($K = 2$); full interdependent ($K = 11$)
Number and scope of search agents	<i>Baseline scenarios:</i> $M = 4$, with $d1 = (d1, \dots, d3)$, $d2 = (d4, \dots, d6)$, $d3 = (d7, \dots, d9)$, $d4 = (d10, \dots, d12)$ <i>Sensitivity analysis:</i> $M = 2$ with $d1 = (d1, \dots, d6)$, $d2 = (d7, \dots, d12)$ $M = 6$ with $d1 = (d1, \dots, d3)$ to $d6 = (d10, \dots, d12)$
Number of organizational dimensions	$L = 3$
Dimension $l=1$: Agents' objective	$\alpha \hat{=} \{0, 0.5, 1\}$
Dimension $l=2$: Precision of evaluation	$(sr, own, sr, res, scent) \hat{=} \{(0, 0, 0), (0.1, 0.15, 0.125), (0.05, 0.25, 0.125)\}$
Dimension $l=3$: Coordination mode	decentralized; lateral veto; centralized
Interval of change	$T^* = 25$ and for contrasting to “no change”: $T^* > T$
Reinforcement strength	$\lambda \hat{=} \{0, 0.5\}$
Aspiration level	$v \hat{=} \{0, 0.01\}$

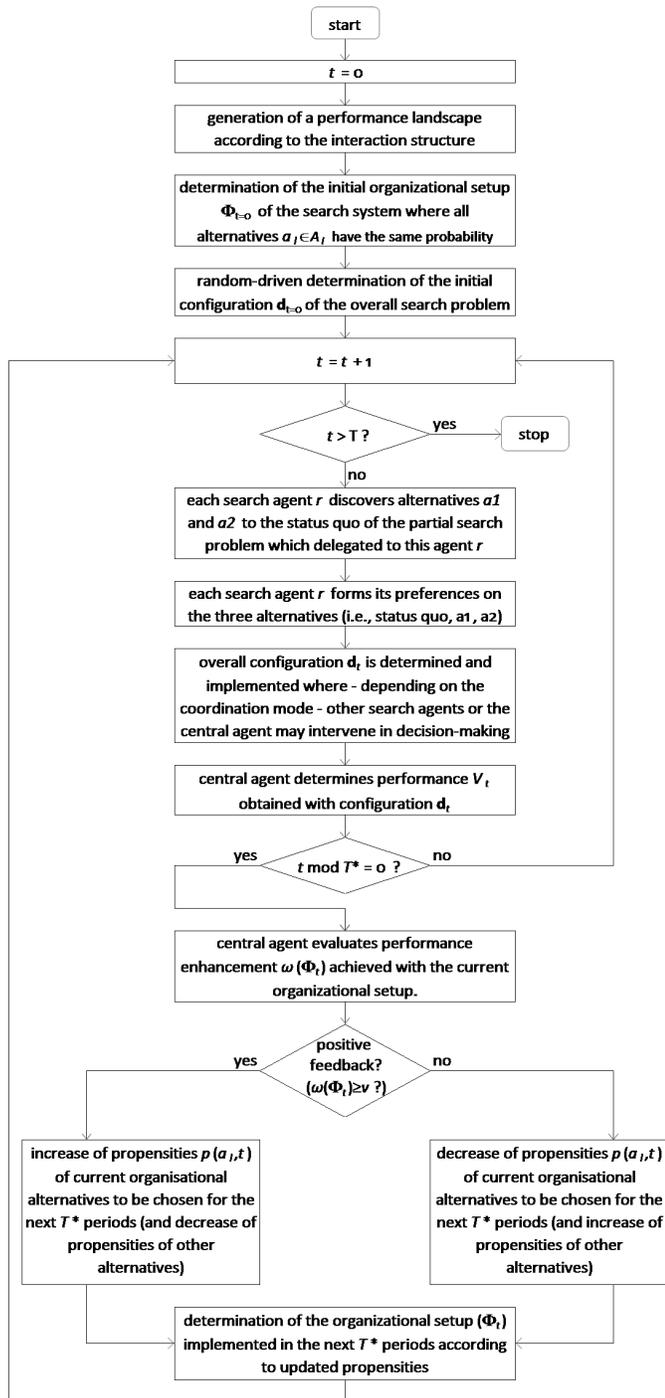


Fig. 1. Schematic representation of one simulation run over T periods including potential changes of the search system's organizational setup.

In order to oppose search systems with learning capabilities (i.e., with $\lambda > 0$) to non-learning systems employing organizational change, simulations for $\lambda = 0$ are conducted. Moreover, search systems which do not alter their organization within the observation time T (i.e., with $T^* > T$) are simulated.

In order to capture the complexity of the underlying search problem, simulations for two interaction structures are performed which, in a way, represent two extreme scenarios [22]: in the *block-diagonal* structure the overall search problem can be segmented into disjoint parts with maximal intense intra-sub-problem interactions but no cross-sub-problem interactions (K^*). An example is given in Figure 2.a

with $K = 2$ and $K^* = 0$ where each of the four sub-problems is assigned to one search agent. In this setup, one agent's decisions do not affect the performance contributions of the other agents' choices.

The second case is characterized by full interdependence, i.e., all single options d_i affect the performance contributions of all other choices $d_{j \neq i}$ and the search problem's complexity is raised to its maximum, i.e., intensity of interactions K as well as the cross-sub-problem interactions K^* are maximal (see Figure 2.b for an example with $K = 11$ and $K^* = 9$).

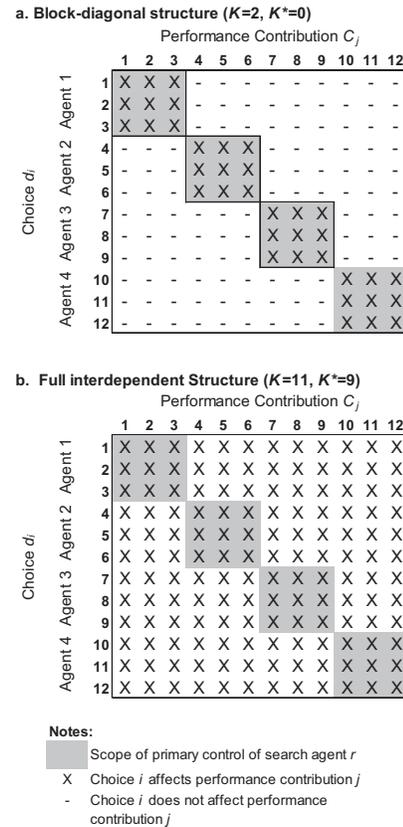


Fig. 2. Interaction structures and assignment of choices to search agents for the a. block-diagonal and b. full interdependent structure.

IV. RESULTS

The simulation experiments are conducted for two baseline scenarios of complexity (see Figures 2.a and 2.b) and for four modes of organizational adaption: (I) no change, (II) change without learning, (III) learning-based change with low aspiration level and with (IV) high aspiration level. These baseline scenarios are, then, modified in the number of search agents. In the modified scenarios, two or six search agents are employed instead of four.

A. Baseline Scenarios

Table II displays condensed results of the simulated scenarios. The final performance achieved in the end of the observation period ($V_{t=500}$) and the performance achieved on average in each of the 500 periods ($\bar{V}_{[0;500]}$) may serve as indicators for the effectiveness of the search process. The same applies to the relative frequency of how often the global maximum is found in the final period. The ratio of alterations of d informs about the diversity of the short termed search processes, while the average number of alterations of organizational dimensions informs about the diversity of organizational setups that are implemented during adaptive walks and, eventually, modified within in the (mid-

termed) search for appropriate organizational setups. Figure 3 depicts the averaged adaptive walks for the different modes of change in the block-diagonal structure of interactions, and Figure 4 reports on the full interdependent structure correspondingly. In particular, Figures 3 and 4 show the performance differences of the “change, no learning” mode and the two modes employing learning, against the “no change” mode.

TABLE II
CONDENSED RESULTS OF THE BASELINE SCENARIOS

Scenario of learning and change	Final Perform. $V_{t=500}$ ^a	Average Perform. $\bar{V}_{[0,500]}$ ^a	Frequ. glob. max at $t=500$	Ratio of altered configs. of \mathbf{d}	Average no. of altered organiz. dimens. over T
<i>Block diagonal interaction structure (K=5)^b</i>					
I. No change	0.954 ±0.0042	0.9519 ±0.0035	36.08%	20.00%	n/a
II. Change, no learning	0.9599 ±0.0044	0.9582 ±0.001	45.40%	22.97%	38.2
III. Learning low asp.lvl	0.9696 ±0.0035	0.9609 ±0.0025	47.40%	19.85%	6.9
IV. Learning high asp.lvl	0.914 ±0.0055	0.9296 ±0.0022	17.76%	43.19%	20.0
<i>Full interdependent interaction structure (K=11)^b</i>					
I. No change	0.8372 ±0.0081	0.811 ±0.0075	4.52%	5.63%	n/a
II. Change, no learning	0.8843 ±0.0075	0.8421 ±0.0048	7.84%	10.24%	37.9
III. Learning low asp.lvl	0.8684 ±0.007	0.8317 ±0.0059	4.96%	8.74%	6.2
IV. Learning high asp.lvl	0.8669 ±0.0093	0.8261 ±0.0058	7.76%	15.98%	36.2

^a Confidence intervals, at a confidence level of 99.9%, for final performance range between 0.002 and 0.005 and for average performance between 0.001 and 0.003.

^b Scenarios: “no change”: $T^* > 500$; “change, no learning”: $\lambda = 0$; $\nu = 0.01$; “learning, low aspiration level”: $\lambda = 0.5$; $\nu = 0$; “learning, high aspiration level”: $\lambda = 0.5$; $\nu = 0.01$. For further parameter settings see Table I.

Each data row shows the results of 2,500 adaptive walks: 10 walks on 250 distinct landscapes.

In the following, three aspects of the presented results are discussed in detail: (1) performance differences of scenarios in which the organizational setup is changed against scenarios in which the organizational setup is modelled to be constant, (2) the effects of learning-based adaptation compared to purely random adaptations of the organizational setup, (3) the intensity of organizational change (which is captured by the average number of altered dimensions).

Concerning the first aspect, Table II as well as Figures 3 and 4 indicate that – with one exception – performance levels of scenarios employing change persistently go beyond the level of performance achieved without change. This behavior can be observed after approximately 40 periods. These results confirm findings of research which indicate that altering the organizational setup in the course of distributed search processes may be favorable [8], [10], [11]: It has been argued that this is driven by the increased diversity of search which reduces the peril of sticking to local peaks. This is broadly confirmed by the ratio of alterations of configurations \mathbf{d} and the frequency of how often the global maximum is found in $t = 500$ (cf. Table II).

However, results also suggest that learning by reinforcement with high aspiration levels is not universally beneficial. Apparently, the complexity of the search problem together with the aspiration level subtly affects the benefits of learning. In case of the block-diagonal

structure, employing learning-based change with a high aspiration level leads to performance levels that are remarkably below the performances achieved without change throughout the adaptive walk from about time-step 75 to 500 and the final performance $V_{t=500}$ is about 4 points of percentage below the “no change” case. An explanation why, in case of the block-diagonal interaction structure, a high aspiration level apparently induces such a rather poor performance, may be based in the specific selective effects induced in this scenario:

With increasing aspiration level it becomes more unlikely that a positive stimulus $\tau(t)$ is achieved under the regime of a certain organizational setup – even if the setup had brought some (lower than ν) performance enhancements in the last T^* periods. Hence, even potentially appropriate organizational setups are likely to receive low probabilities to be re-chosen for the next T^* periods. In the block-diagonal structure with its fairly low level of interactions ($K = 2$), it is rather likely that the global maximum is found [22]: of course, no further performance enhancement is possible in these cases and the aspiration level is not reached. Whenever the global maximum is found (with aspiration level $\nu > 0$) the organizational setup is likely to be modified. An altered organizational setup also induces a modified evaluation of the current configuration \mathbf{d} [11]. As a result, a move away from the global maximum in the performance landscape becomes likely.

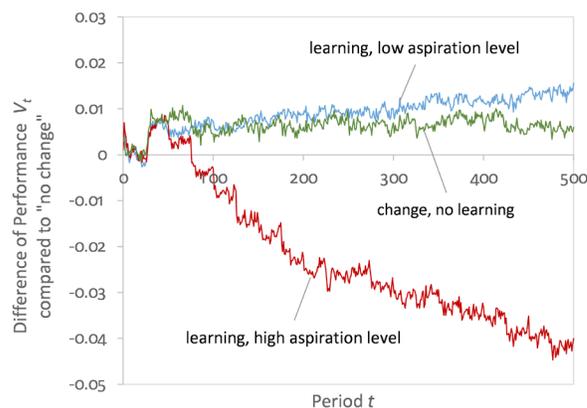


Fig. 3. Performance differences of adaptive search processes employing organizational change against search processes without alterations of the organizational setup in case of the block-diagonal interaction structure. Each curve represents the difference of means of the average of 2,500 adaptive walks, i.e., 250 distinct performance landscapes with 10 adaptive walks on each. For parameter settings see Table I.

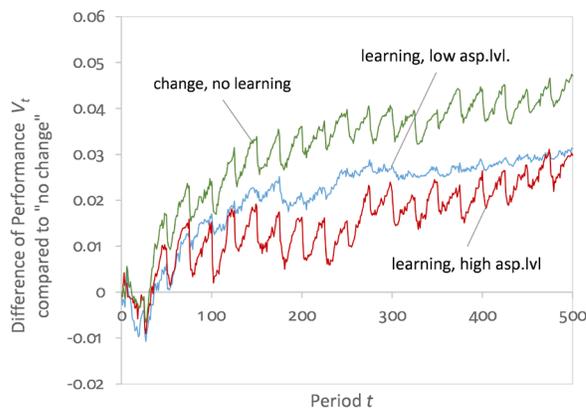


Fig. 4. Performance differences of adaptive search processes employing organizational change against search processes without alterations of the organizational setup in case of the full interdependent interaction structure. Each curve represents the difference of means of the average of 2,500 adaptive walks, i.e., 250 distinct performance landscapes with 10 adaptive walks on each. For parameter settings see Table I.

The second aspect to be discussed in detail is related to the performance effects of learning-based adaptation compared to the purely random-driven alterations. The results suggest that learning-based change is not universally more beneficial than purely random-driven organizational change. Rather, it appears that the aspiration level ν is of remarkable relevance: in both interaction structures, learning-based adaptation employing a high aspiration level leads to a level of final performance that is inferior to the performances achieved under purely random-driven change. Employing a low aspiration level performs best in the block-diagonal structure it leads to a medium performance in the case of high complexity.

As argued above, a high aspiration level induces more organizational alterations which leads to more diversity of search, i.e., more alterations of d , as compared to the low aspiration level. For highly complex interaction structures, a particular peril is that the search processes may stick to a local optimum, and, hence, increasing diversity of search “per se” may be beneficial. This might explain the good performance of the “change, no learning” mode. However, a high aspiration level, in a way, “penalizes” particularly those search processes which have reached a good solution from which further improvements are hard to achieve: as argued above, the block-diagonal interaction structure is particularly prone to this effect; however, the rather low performance in the full-interdependent structure (Figure 4) might also be caused by this effect.

With the third aspect to be discussed more into detail the intensity of organizational change (right-most column in Table II), and, thus, the *efficiency* of the mode of change and learning comes into play. The average number of organizational dimensions in which alterations occur during the adaptive search may be regarded as an indicator for the effort (“costs”), if any, of organizational dynamics.

Obviously, the context of the search organization is relevant for whether, or not, and, if so, in which shape costs of organizational change occur. For example, in case of a network of unmanned aerial vehicles, collaboratively serving a certain service area, the switch from one coordination mode to another might not cause any costs (apart from activating another of already available coordination mechanisms); however, in case of firm managers, collaboratively searching for better configurations of key performance drivers, reorganizations are rather costly, including, for example, learning costs of new organizational procedures or the adjustment of incentive schemes. Hence, the average number of dimensions changed may be rather critical for the efficiency of inducing organizational dynamics of search.

Results suggest that, in both interaction structures under investigation, learning with a low aspiration level yields good performance and a high level of efficiency as compared to the other scenarios:

In case of the block-diagonal interaction structure the final performance achieved with a low aspiration level exceeds the performance reached via purely random-driven change by about 7 points of percentage while the average number of organizational alterations is remarkably lower (i.e., 6.9 altered dimensions on average in case of learning with low aspiration level versus 38.2 in case of purely randomized change). If the complexity of the search problem is high the performance of the “change, no learning” scenario exceeds the performance of learning-based adaptation with low aspiration levels; however, this comes along with, on average, 37.9 organizational alterations compared to 6.2 alterations in the latter case.

In sum, it appears that learning with low aspiration level may provide rather high performance levels combined with few organizational alterations. Thus, whenever organizational alterations do not come along without any cost, learning with low aspiration level appears to be particularly interesting with respect to the efficiency of search.

B. Sensitivity Analysis

In the sensitivity analysis, the baseline scenarios are modified with respect to the number of search agents: the simulations additionally are conducted for systems with two and with six search agents. In particular, the interactions among decisions remain unchanged, but the *assignment of decisions* is modified. Figures 5.a and 5.b show the assignment for the case of two agents and six agents, respectively, in the block-diagonal structure as compared to Figure 2.a.

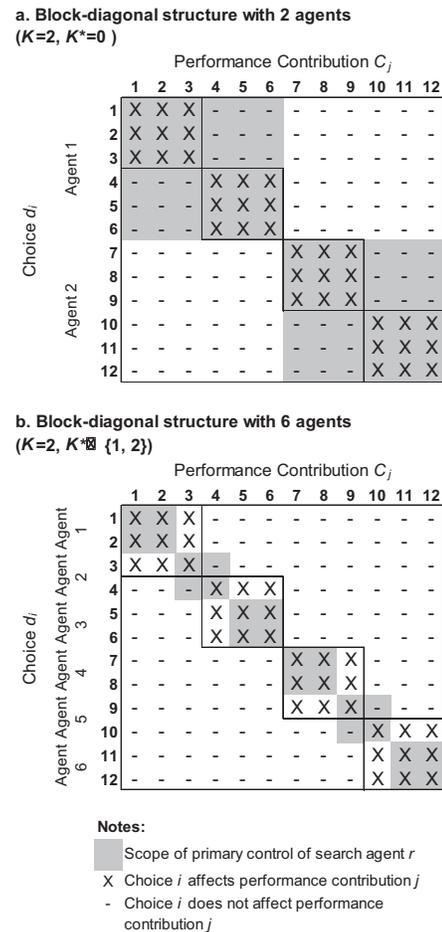


Fig. 5. Assignment of four independent sub-problems to a. two and b. six search agents.

In the simulation model, with increasing (decreasing) the number of search agents the diversity of search is increased (decreased): in each time step, each search agent discovers two alternatives to the status quo of the own partial sub-problem (Section II.A) – one alternative where one bit is flipped and another with two bits flipped. Thus, in case of two search agents, at maximum four bits of the entire configuration d could be flipped in time step t ; in contrast, with six agents at maximum 12 bits could be flipped. Thus, with increasing number of agents the need for coordination is increased too, and viceversa.

Figures 6 and 7 plot the final performance $V_{t=500}$ achieved in the block-diagonal and the full-interdependent interaction structure, respectively, with two, four and six search agents.

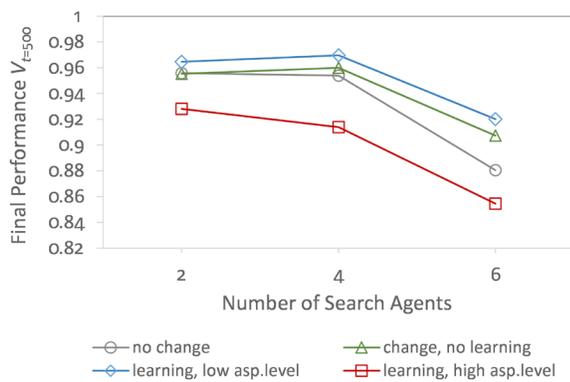


Fig. 6. Sensitivity of final performance to number of search agents in the block-diagonal interaction structure. Each mark represents the average of 2,500 adaptive walks, i.e., 250 distinct performance landscapes with 10 adaptive walks on each. For parameter settings see Table I.

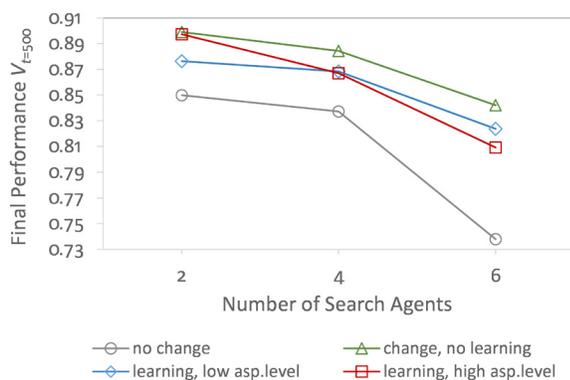


Fig. 7. Sensitivity of final performance to number of search agents in the full interdependent interaction structure. Each mark represents the average of 2,500 adaptive walks, i.e., 250 distinct performance landscapes with 10 adaptive walks on each. For parameter settings see Table I.

The results suggest that the “change, no learning” mode and learning with low aspiration level are least sensitive to the number of agents. In contrast, the final performance obtained by learning with high aspiration level varies considerably with the number of agents. However, the “no change” mode appears most sensitive to an increase in the number M of search agents compared to the modes employing organizational change.

With the transition from two to four agents, the final performance shows rather slight de- or increases – depending on the mode of change and the interaction structure. However, with the transition from four to six agents the final performance obtained decreases remarkably in both interaction structures.

An interesting question is what might cause these effects. A reason might be given by the relation between assignment of decisions to search agents and the interactions among agents’ decisions. For example, with six search agents in the block-diagonal structure (Figure 5.b), cross-agent interactions among search agents’ choices occur whereas for two and four agents no cross-agent interactions show up (Figure 5.a). Hence, in this interaction structure the need for coordination among agents’ choices ranges from no need at all (i.e., $K^* = 0$ for $M = 2$ and $M = 4$) to some coordination need as captured by $K^* = 1$ or $K^* = 2$ (see Figure 5.b).

V. CONCLUSION

The major finding of this study is that employing self-adaptation for the organizational setup of distributed search systems via reinforcement-

based learning potentially leads to high levels of performance and this, in particular, with a rather high level of efficiency, as given by the extent of reorganization. These findings are particularly interesting when reorganizing the search system causes marginal costs – may it be due to learning of new organizational procedures on the agents’ site or adjustments required in institutional arrangements.

However, the results also suggest that the complexity of the search problem together with the aspiration level considerably shapes the effects of reinforcement learning – which, at worst, may even be harmful if compared to refraining from any organizational alterations. These findings may sensitize the designer of a distributed search system to employing learning by reinforcement as the level of the aspired performance enhancements should not be overstretched in order to avoid “hyper-actively and ineffectively” alternating search systems. Moreover, the sensitivity analysis suggests that learning with high aspiration level is particularly sensitive to the need for coordination among search agents.

These findings emphasize the need for further research efforts. An obvious next step is to test the key idea of inducing learning-based organizational change in more practical settings than the one presented here. Though some preliminary results obtained for learning-based selection of the coordination mode in terms of the job scheduling policy employed by a swarm of unmanned aerial vehicles [23] provide some support for the ideas presented in this paper, further applications are definitely of interest.

Moreover, further studies should perform in-depth analyses of the role of the aspiration level and other parameters like the interval between of organizational alterations or the learning strength which were fixed in the simulation experiments presented in this paper. Furthermore, the basic search problem captured in this study is rather unstructured in terms of randomized performance contributions (apart from the structure of interactions); hence, in further research studies learning-based organizational adjustments of the search system may turn out to be even more beneficial in case of more structured search problems.

REFERENCES

- [1] C. Yongcan, Y. Wenwu, R. Wei and C. Guanrong, “An overview of recent progress in the study of distributed multi-agent coordination,” *IEEE Trans. on Industrial Informatics*, vol. 9, pp. 427–438, Jan. 2013.
- [2] T. Gross and B. Blasius, “Adaptive coevolutionary networks: a review,” *Journal of the Royal Society Interface*, vol. 20, pp. 259–271, Mar. 2008.
- [3] K. M. Carley and L. Gasser, “Computational organization theory,” in: *Multiagent systems: A modern approach to distributed artificial intelligence*, G. Weiss, Ed., Cambridge: MIT Press, 1999, pp. 299–330.
- [4] A. H. Bond and L. Gasser, “Chapter 1 – Orientation,” in *Readings in Distributed Artificial Intelligence*, A. H. Bond and L. Gasser (Ed.), San Mateo: Morgan Kaufmann, 1988, pp. 1–56.
- [5] M. Wooldridge, *An Introduction to Multiagent Systems* (2nd ed.), Chichester: Wiley, 2009.
- [6] F. von Martial, *Coordinating Plans of Autonomous Agents*, Lecture Notes in Artificial Intelligence, vol. 610. Berlin/Heidelberg/New York: Springer, reprint 2008 (1992).
- [7] Y. Brun, G. Di Marzo Serugendo, C. Gacek, H. Giese, H. Kienle, M. Litoiu, H. Müller, M. Pezzè, and M. Shaw, “Engineering Self-Adaptive Systems through Feedback Loops”, in *Software Engineering for Self-Adaptive Systems*, BHC Cheng, R. de Lemos, H. Giese, P. Inverardi, and J. Magee, Ed., Berlin/Heidelberg: Springer, 2009, pp 48–70.
- [8] O. Baumann, “Distributed Problem Solving in Modular Systems: the Benefit of Temporary Coordination Neglect,” *Systems Research and Behavioral Science*, vol. 32, pp. 124–136, Jan./Feb. 2015.
- [9] F. Wall, “Organizational dynamics in adaptive distributed search processes: Effects on performance and the role of complexity,” in *Frontiers of Information Technology & Electrical Engineering*, vol. 17, pp. 283–295, Apr. 2016.
- [10] F. Wall, “Effects of organizational dynamics in adaptive distributed search

- processes,” in *Distributed Computing and Artificial Intelligence, 12th Int. Conf.*, S. Omatu, Q.M. Malluhi, S. R. Gonzalez et al., Ed., *Advances in Intelligent Systems and Computing* vol. 373, Berlin: Springer, 2015, pp. 121–128.
- [11] F. Wall, “Beneficial Effects of Randomized Organizational Change on Performance,” in *Advances in Complex Systems*, vol. 18, 05n06:1550019, Nov. 2015.
- [12] F. Wall, “Self-adaptive organizations for distributed search: The case of reinforcement learning,” in *Distributed Computing and Artificial Intelligence, 13th International Conference*, S. Omatu, A. Semalat, G. Bocewicz et al., Ed., *Advances in Intelligent Systems and Computing*, vol. 474, Cham: Springer International Publishing, 2016, pp 23–32.
- [13] S. A. Kauffman and S. Levin, “Towards a general theory of adaptive walks on rugged landscapes,” *Journal of Theoretical Biology*, vol. 128, pp. 1145, Jan. 1987.
- [14] S. A. Kauffman, *The Origins of Order: Self-Organization and Selection in Evolution*. Oxford: Oxford University Press, 1993.
- [15] R. Li, M. M. Emmerich, J. Eggermont, E. P. Bovenkamp, T. Bäck, J. Dijkstra and J. C. Reiber, “Mixed-integer NK landscapes,” in *Parallel Problem Solving from Nature IX*, T. Runarsson, H.-G. Beyer, E. Burke, J. Merelo-Guervós, L. D. Whitley and X. Yao, Ed., *Lecture Notes in Computer Science*, vol. 4193, Berlin: Springer, 2006, pp. 42–51.
- [16] B. Levitan and S.A. Kauffman, “Adaptive walks with noisy fitness measurements.” *Molecular Diversity*, vol. 1, pp. 53–68, Jan. 1995.
- [17] F. Wall, “The (beneficial) role of informational imperfections in enhancing organisational performance,” in *Progress in Artificial Economics, 6th Artificial Economics*, M. Li Calzi, L. Milone and P. Pellizzari, Ed., *Lecture Notes in Economics and Mathematical Systems*, vol. 645, Berlin: Springer, 2010, pp. 115–126.
- [18] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, 2nd ed., Cambridge (Mass.): MIT Press, 2012.
- [19] L. P. Kaelbling, M. L. Littman and A. W. Moore, “Reinforcement learning: a survey,” *Journal of Artificial Intelligence Research*, vol. 4, pp. 237–285, May 1996.
- [20] R. R. Bush and F. Mosteller, *Stochastic Models for Learning*, Oxford (Engl.): Wiley, 1955.
- [21] T. Brenner, “Agent learning representation: Advice on modelling economic learning,” in *Handbook of Computational Economics*, vol. 2, L. Tesfatsion and K. L. Judd, Ed., Amsterdam: Elsevier, 2006, pp. 895–947.
- [22] R. W. Rivkin and N. Siggelkow, “Patterned interactions in complex systems: Implications for exploration,” in *Management Science*, vol. 53, pp. 1068–1085, July 2007.
- [23] A. Lembacher, “The Application of Learning Algorithms in Job Scheduling Problems” Master Thesis (supervisor F. Wall), Universitaet Klagenfurt 2015 (unpublished).



Friederike Wall earned her Diploma for Business Economics in 1988 and her Doctoral degree in 1991, both at the Georg-August-Universität Göttingen, Germany. In 1996 she received the “*venia legendi*” (Habilitation for Business Economics) from the Universität Hamburg, Germany. After being a scientific Project Referent for Accounting in the context of the implementation of SAP R/3 at Max-Planck Society, Munich, Germany she became Full Professor of Business Administration, esp. Controlling and Information Management at the Universität Witten/Herdecke, Germany. Since 2009 she is Full Professor and Head of the Department of Controlling and Strategic Management at the Alpen-Adria-Universität Klagenfurt, Austria. Prof. Wall’s scientific work is focused on distributed decision-making, management control systems and the quality of the information provided by these systems. Her main research approach is defined by agent-based simulation methods and agent-based technologies.