OPEN

Correspondence and requests for materials should be addressed to S.-J.K. (KIM.Songju@nims.go.jp)

Decision Maker based on Nanoscale Photo-excitation Transfer

Song-Ju Kim[1], Makoto Naruse[2], Masashi Aono[3], Motoichi Ohtsu[4] & Masahiko Hara[5]

[1]WPI Center for Materials Nanoarchitectonics (MANA), National Institute for Materials Science (NIMS), 1-1 Namiki, Tsukuba, Ibaraki 305-0044, Japan, [2]Photonic Network Research Institute, National Institute of Information and Communications Technology, 4-2-1 Nukui-kita, Koganei, Tokyo 184-8795, Japan, [3]Earth-Life Science Institute, Tokyo Institute of Technology, 2-12-1 Ookayama, Meguro-ku, Tokyo 152-8550, Japan, [4]Department of Electrical Engineering and Information Systems / Nanophotonics Research Center, Graduate School of Engineering, The University of Tokyo, 2-11-16 Yayoi, Bunkyo-ku, Tokyo 113-8656, Japan, [5]Department of Electronic Chemistry, Interdisciplinary Graduate School of Science and Engineering, Tokyo Institute of Technology, 4259 Nagatsuta, Midori-ku, Yokohama 226-8503, Japan.

Decision-making is one of the most important intellectual abilities of the human brain. Here we propose an efficient decision-making system which uses optical energy transfer between quantum dots (QDs) mediated by optical near-field interactions occurring at scales far below the wavelength of light. The simulation results indicate that our system outperforms the softmax rule, which is known as the best-fitting algorithm for human decision-making behaviour. This suggests that we can produce a nano-system which makes decisions efficiently and adaptively by exploiting the intrinsic spatiotemporal dynamics involving QDs mediated by optical near-field interactions.

Which clothes should I wear? Which restaurant should we choose for lunch? Which article should I read first? We make many decisions in our daily lives. Can we look to nature to find a method for 'efficient decision-making'? For formal discussion, let us focus on the multi-armed bandit problem (BP), stated as follows. Consider a number of slot machines. Each machine when pulled, rewards the player with a coin at a certain probability P_k ($k \in \{1, 2, \ldots, N\}$). For simplicity, we assume that the reward from one machine is the same as that from another machine. To maximise the total amount of reward, it is necessary to make a quick and accurate judgment of which machine has the highest probability of giving a reward. To accomplish this, the player should gather information about many machines in an effort to determine which machine is the best; however, in this process, the player should not fail to exploit the reward from the known best machine. These requirements are not easily met simultaneously, because there is a trade-off between 'exploration' and 'exploitation'. The BP is used to determine the optimal strategy for maximising the total reward with incompatible demands, either by exploiting rewards obtained owing to already collected information or by exploring new information to acquire higher pay-offs in risk taking. Living organisms commonly encounter this 'exploration-exploitation dilemma' in their struggle to survive in the unknown world.

This dilemma has no known generally optimal solution. What strategies do humans and animals exploit to resolve this dilemma? Daw et al. found that the softmax rule is the best-fitting algorithm for human decision-making behaviour in the BP task[1]. The softmax rule uses the randomness of the selection specified by a parameter analogous to the temperature in the Boltzmann distribution (see Methods). The findings of Daw et al. raised many exciting questions for future brain research[2]. How humans and animals respond to the dilemma and the underlying neural mechanisms still remain important and open questions.

The BP was originally described by Robbins[3], although the same problem in essence was also studied by Thompson[4]. However, the optimal strategy is known only for a limited class of problems in which the reward distributions are assumed to be 'known' to the players, and an index called 'the Gittins index' is computed[5,6]. Furthermore, computing the Gittins index in practice is not tractable for many problems. Auer et al. proposed another index expressed as a simple function of the total reward obtained from a machine[7]. This upper confidence bound (UCB) algorithm is used worldwide for many applications, such as Monte-Carlo tree searches[8,9], web content optimization[10] and information and communications technology (ICT)[11–14].

Kim et al. proposed a decision-making algorithm called the 'tug-of-war model' (TOW); it was inspired by the true slime mold *Physarum*[15,16], which maintains a constant intracellular resource volume while collecting environmental information by concurrently expanding and shrinking its branches. The conservation law entails a 'nonlocal correlation' among the branches, that is, the volume increment in one branch is immediately

compensated for by volume decrement(s) in the other branch(es). This nonlocal correlation was shown to be useful for decision-making. Thus, the TOW is a dynamical system which describes spatiotemporal dynamics of a physical object (i.e. an amoeboid organism). The TOW connected 'natural phenomena' to 'decision-making' for the first time. This approach enables us to realise an 'efficient decision maker'–an object which can make decisions efficiently.

Here we demonstrate the physical implementation of the TOW with quantum dots (QDs) and optical near-field interactions by using numerical simulations. Semiconductor QDs have been used for innovative nanophotonic devices[17,18] and optical near-field interactions have been successfully applied to solar cells[19], LEDs[20], diode lasers[21] etc. We have already proposed QD systems for computing applications, such as the constraint satisfaction problem (CSP) and the satisfiability problem (SAT)[22,23]. We introduce a new application for decision-making by making use of optical energy transfer between QDs mediated by optical near-field interactions.

We use three types of cubic QDs with side lengths of a, $\sqrt{2}a$ and $2a$, which are respectively represented by $\mathbf{QD}_S$, $\mathbf{QD}_M$ and $\mathbf{QD}_L$ [Fig. 1(a)]. We assume that five QDs are one-dimensionally arranged in 'L-M-S-M-L' or 'M-L-S-L-M' as shown in Fig. 1(b), where $\mathbf{S}$, $\mathbf{M}$ and $\mathbf{L}$ are simplified representations of $\mathbf{QD}_S$, $\mathbf{QD}_M$ and $\mathbf{QD}_L$, respectively. Owing to the steep electric fields in the vicinity of these QDs, an optical excitation can be transferred between QDs through resonant energy levels mediated by optical near-field interactions[24–28]. Here we should note that an optical excitation is usually transferred from smaller QDs to larger ones owing to energy dissipation processes occurring at larger QDs (details are described in Supplementary Information). In addition, an optical near-field interaction follows Yukawa-type potential, meaning that it could be engineered by inter-dot distances,

$$U(r) = A\exp(-r/a)/r, \qquad (1)$$

where r is the inter-dot distance, and A and a are constants[29].

When an optical excitation is generated at $\mathbf{QD}_S$, it is transferred to the lowest energy levels in $\mathbf{QD}_L$s; we observe negligible radiation from $\mathbf{QD}_M$s. However, when the lowest energy levels of $\mathbf{QD}_L$s are occupied by control lights, which induce state-filling effects, the optical excitation at $\mathbf{QD}_S$ is more likely to be radiated from $\mathbf{QD}_M$s[30].

Here we consider the photon radiation from either left $\mathbf{QD}_{ML}$ or right $\mathbf{QD}_{MR}$ as the decision of selecting slot machine A and B, respectively. The intensity of the control light to induce state-filling at the left and right $\mathbf{QD}_L$s is respectively modulated on the basis of the resultant rewards obtained from the chosen slot machine. We call such a decision-making system the 'QD-based decision maker (QDM)'. The QDM can be easily extended to N-armed ($N > 2$) cases, although we demonstrate only the two-armed case in this study.

It should be noted that the optical excitation transfer between QDs mediated by optical near-field interactions is fundamentally probabilistic; this is described below in detail on the basis of density matrix formalism. Until energy dissipation is induced, an optical excitation simultaneously interacts with potentially transferable destination QDs in the resonant energy level. We exploit such probabilistic behaviour for the function of exploration for decision-making.

It also should be emphasised that conventionally, propagating light is assumed to interact with nanostructured matter in a spatially uniform manner (by a well-known principle referred to as long-wavelength approximation) from which state transition rules for optical transitions are derived, including dipole-forbidden transitions. However, such an approximation is not valid for optical near-field interactions in the subwavelength vicinity of an optical source; the inhomogeneity of optical near-fields of a rapidly decaying nature makes even conventionally dipole-forbidden transitions allowable[17,22,23].

We introduce quantum mechanical modelling of the total system based on a density matrix formalism. For simplicity, we assume one

excitation system. There are in total 11 energy levels in the system: S_1 in $\mathbf{QD}_S$; ML_1 and ML_2 in $\mathbf{QD}_{ML}$; LL_1, LL_2 and LL_3 in $\mathbf{QD}_{LL}$; MR_1 and MR_2 in $\mathbf{QD}_{MR}$; LR_1, LR_2 and LR_3 in $\mathbf{QD}_{LR}$. Therefore the number of different states occupying these energy levels is 12 including the vacancy state, as schematically shown in Fig. 1(b). Because a fast inter-sublevel transition in $\mathbf{QD}_L$s and $\mathbf{QD}_M$s is assumed, it is useful to establish theoretical treatments on the basis of the exciton population in the system composed of $\mathbf{QD}_S$s, $\mathbf{QD}_M$s and $\mathbf{QD}_L$s, where 11 basis states are assumed, as schematically illustrated in Fig. 1(c).

Results

Here we propose the QDM that is based on the five QD system, as shown in Fig. 1(b). Through optical near-field interactions, an optical excitation generated at $\mathbf{QD}_S$ is transferred to the lowest energy levels in the largest-size QD, namely the energy level LL_1 or LR_1. However, when LL_1 and LR_1 are occupied by other excitations, the input excitation generated at S_1 should be relaxed from the middle-sized QD, that is, ML_1 and MR_1. The idea of the QDM is to induce state-filling at LL_1 and/or LR_1 while observing radiations from ML_1 and MR_1. If radiation occurs in ML_1 (MR_1), then we consider that the system selects machine A (B). We can modulate these radiation probabilities by changing the intensity of the incident light.

We adopt the intensity adjuster (IA) to modulate the intensity of incident light to large QDs, as shown at the bottom of Fig. 2. The initial position of the IA is zero. In this case, the same intensity of light is applied to both LL_1 and LR_1. If we move the IA to the right, the intensity of the right increases and that of the left decreases. In contrast, if we move the IA to the left, the intensity of the left increases and that of the right decreases. This situation can be described by the following relaxation rate parameters as functions of the IA position j: $\Gamma_{LR_2} = \frac{1}{100} - \frac{1}{10000}j + \frac{1}{100000}$, and $\Gamma_{LL_2} = \frac{1}{100} + \frac{1}{10000}j + \frac{1}{100000}$.

In advance, we calculated two (right and left) radiation probabilities from ML_1 and MR_1, namely $S_A(j)$ and $S_B(j)$, in each of the 201 states ($-100 \leq j \leq 100$) by using the quantum master equation of the total system (see Supplementary Information). There are 101 independent values. Results are shown in Fig. 2 as the solid red line. If $j > 0$, the intensity of incident light to the LR_1 increases (the Γ_{LR_2} decreases by the amount of $j/10,000$), while that to the LL_1 decreases (the Γ_{LL_2} increases by the amount of $j/10,000$). Correspondingly, the radiation probability $S_B(j)$ is larger than $S_A(j)$ for $j > 0$, while the radiation probability $S_A(j)$ is larger than $S_B(j)$ for $j < 0$. For $j < -100$, we used probabilities $S_A(-100)$ and $S_B(-100)$. Similarly, we used $S_A(100)$ and $S_B(100)$ for $j > 100$.

The dynamics of the IA is set up as follows. The IA position is moved according to the reward from a slot machine. Here the parameter D is a unit of the move.
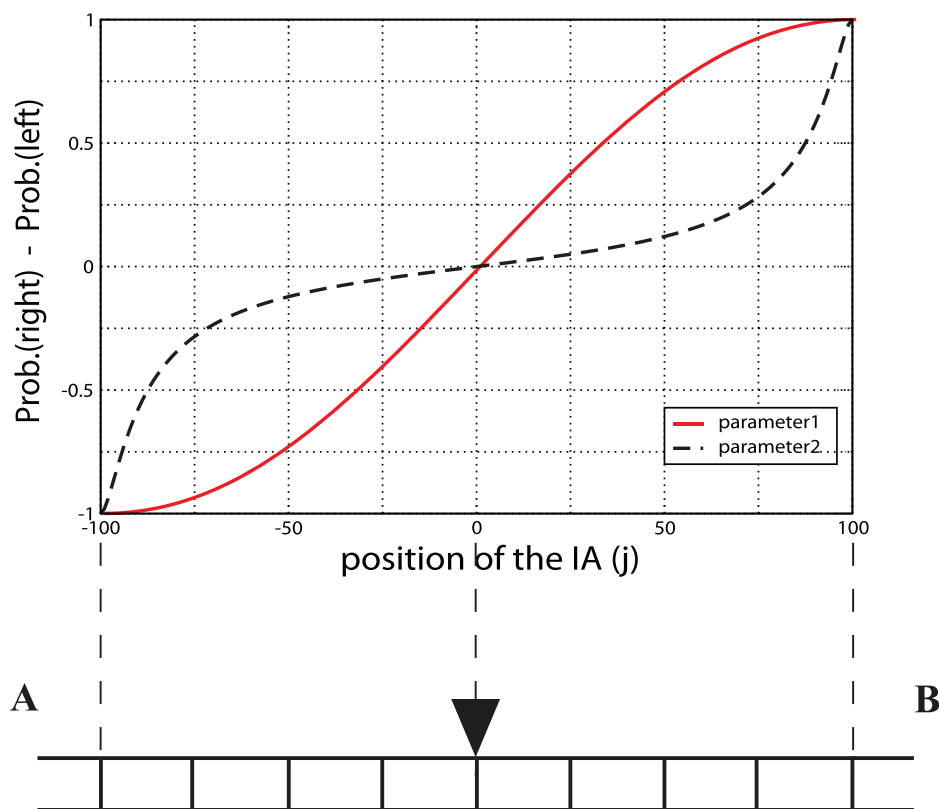
1. Set the IA position j to 0.
2. Select machine A or B by using $S_A(j)$ and $S_B(j)$.
3. Play on the selected machine.
4. If a coin is dispensed, then move the IA to the selected machine's direction, that is, $j = j - D$ for A, and $j = j + D$ for B.
5. If no coin is dispensed, then move the IA to the inverse direction of the selected machine, that is, $j = j + D$ for A, and $j = j - D$ for B.
6. Repeat step (2).

In this way, the system selects A or B, and the IA moves to the right or left according to the reward.

Figures 3(a) and (b) demonstrates the efficiency (cumulative rate of correct selections) for the QDM (solid red line) and the softmax rule with optimised parameter τ (dashed line) in the case where the reward probabilities of the slot machines are (a) $P_A = 0.2$ and $P_B =$
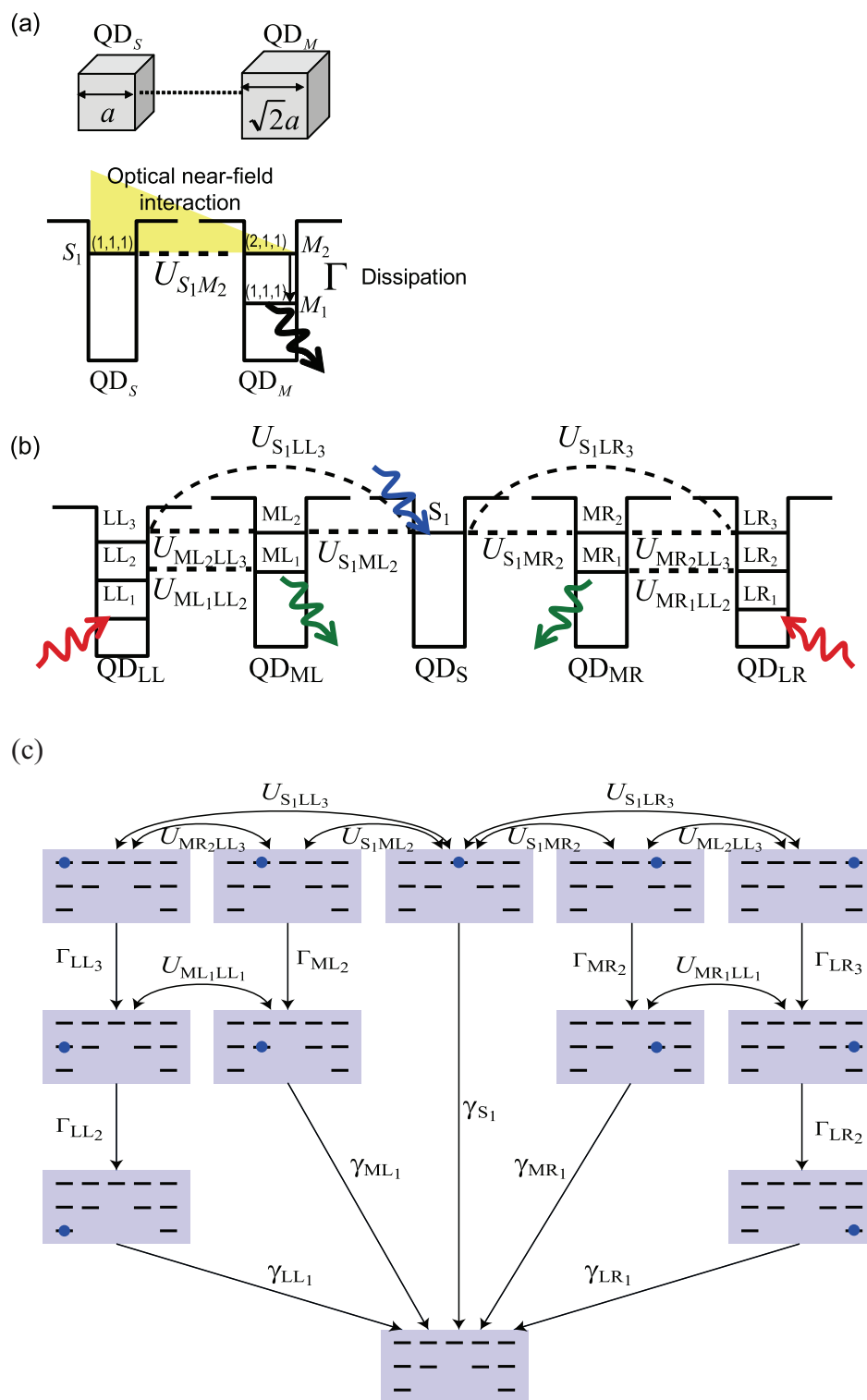
Figure 1 | (a) Energy transfer between quantum dots (QDs). Two cubic quantum dots **QD**$_S$ and **QD**$_M$, whose side lengths are a and $\sqrt{2}a$, respectively, are located close to each other. Optical excitations in **QD**$_S$ can be transferred to neighbouring structures **QD**$_M$ via optical near-field interactions, denoted by $U_{S_1M_2}$[29], because there exists a resonance between the level of quantum number $(1, 1, 1)$ for **QD**$_S$ (denoted by S_1) and that of quantum number $(2, 1, 1)$ for **QD**$_M$ (M_2). (b) QD-based decision maker. The system consists of five QDs denoted **QD**$_{LL}$, **QD**$_{ML}$, **QD**$_S$, **QD**$_{MR}$ and **QD**$_{LR}$. The energy levels in the system are summarised as follows. The $(2, 1, 1)$-level of **QD**$_{ML}$, **QD**$_{MR}$, **QD**$_{LL}$ and **QD**$_{LR}$ is respectively denoted by ML_2, MR_2, LL_2 and LR_2. The $(1, 1, 1)$-level of **QD**$_{ML}$, **QD**$_{MR}$, **QD**$_{LL}$ and **QD**$_{LR}$ is respectively denoted by ML_1, MR_1, LL_1 and LR_1. The $(2, 2, 2)$-level of **QD**$_{LL}$ and **QD**$_{LR}$ is respectively denoted by LL_3 and LR_3. The optical near-field interactions are $U_{S_1Mi_2}$, $U_{S_1Li_3}$, $U_{Mi_2Li_3}$ and $U_{Mi_1Li_2}(i=L, R)$. (c) Schematic summary of the state transitions. Shown are the relaxation rates Γ_{LL_3}, Γ_{LL_2}, Γ_{ML_2}, Γ_{MR_2}, Γ_{LR_3} and Γ_{LR_2}, and the radiative decay rates γ_{LL_1}, γ_{ML_1}, γ_{S_1}, γ_{MR_1} and γ_{LR_1}.

Figure 2 | Intensity adjuster (IA) and difference between radiation probabilities from ML_1 and MR_1. The difference between radiation probabilities $S_B(j) - S_A(j)$ as a function of the IA position j, which are calculated from the quantum master equation of the total system, is denoted by the solid red line. Here we used the parameters $\Gamma_{LR_2} = \frac{1}{100} - \frac{1}{10000}j + \frac{1}{100000}$, and $\Gamma_{LL_2} = \frac{1}{100} + \frac{1}{10000}j + \frac{1}{100000}$. As supporting information, the dashed line denotes the case where $\Gamma_{LR_2} = \frac{1}{10+500(100+j)}$, and $\Gamma_{LL_2} = \frac{1}{10-500(100-j)}$.

0.8 and (b) $P_A = 0.4$ and $P_B = 0.6$. In these cases, the correct selection is 'B' because P_B is larger than P_A. These cumulative rates of correct selections are average values for each 1, 000 samples. Hence, each value corresponds to the average number of coins acquired from the slot machines. Even with a non-optimised parameter D, the performance of the QDM is higher than that of the softmax rule with optimised parameter τ, in a wide parameter range of $D = 10$–100 although we show only the $D = 50$ case in Figs. 3(a) and (b).

Figure 3(c) shows the adaptability (percentage of correct selections) for the QDM (red line) and the softmax rule (black line). The parameter τ of the softmax rule was optimised to obtain the fastest adaptation. The switching occurs at time steps $t = 3, 000$ and 6, 000. Up to $t = 3, 000$, the systems adapt to the initial environment ($P_A = 0.4$ and $P_B = 0.6$). Between $t = 3, 000$ and 6, 000, the systems adapt to the new environment ($P_A = 0.6$ and $P_B = 0.4$). Beyond $t = 6, 000$, the systems adapt to the first environment again ($P_A = 0.4$ and $P_B = 0.6$). It is noted that we set $\beta = 0$ at the switching points because the softmax rule cannot detect the environmental change. Nevertheless, the adaptation (slope) of the QDM was found to be faster (steeper) than that of the softmax rule to the initial environment ($P_A = 0.4$ and $P_B = 0.6$) as well as to the new environment ($P_A = 0.6$ and $P_B = 0.4$). Thus, we can conclude that the QDM has better adaptability than the softmax rule.

Discussion

In summary, we have demonstrated a QDM on the basis of the optical energy transfer occurring far below the wavelength of light by using computer simulations. This paves a new way for utilising quantum nanostructures and inherent spatiotemporal dynamics of optical near-field interactions for totally new application, that is, 'efficient decision-making'. Surprisingly, the performance of the QDM is better than that of the softmax rule in our simulation studies. Moreover, the QDM exhibits superior flexibility in adapting to environmental changes, which is an essential property for living organisms to survive in uncertain environments.

Finally, there are a few additional remarks. First, it has been demonstrated that the optical energy transfer is about 10^4 times more energy efficient as compared with the bit-flip energy of conventional electrical devices[31]. Furthermore, nanophotonic devices based on optical energy transfer with such energy efficiency have been experimentally demonstrated, including input and output interfaces with the optical far-field[32]. These studies indicate that the QDM is highly energy efficient. The second remark is about the experimental implementation of size- and position-controlled quantum nanostructures. Kawazoe et al. successfully demonstrated a room-temperature-operated two-layer QD device by utilising highly sophisticated molecular beam epitaxy (MBE)[17]. In addition, Akahane et al. succeeded in realising more than 100 layers of size-controlled QDs[33]. Besides, DNA-based self-assembly technology can also be a solution in realising controlled nanostructures[34]. Other nanomaterial systems, such as nanodiamonds[35], nanowires[36], nanostructures formed by droplet epitaxy[37], have also been showing rapid progress. These technologies provide feasible solutions. Finally, in the demonstration of decision-making in this study, for simplicity, we dealt with restricted problems, namely, $P_A + P_B = 1$. However, general problems can also be solved by the extended QDM although the IA dynamics becomes a bit more complicated. The IA and its dynamics can also be implemented by using QDs.

(a)

(b)

(c)
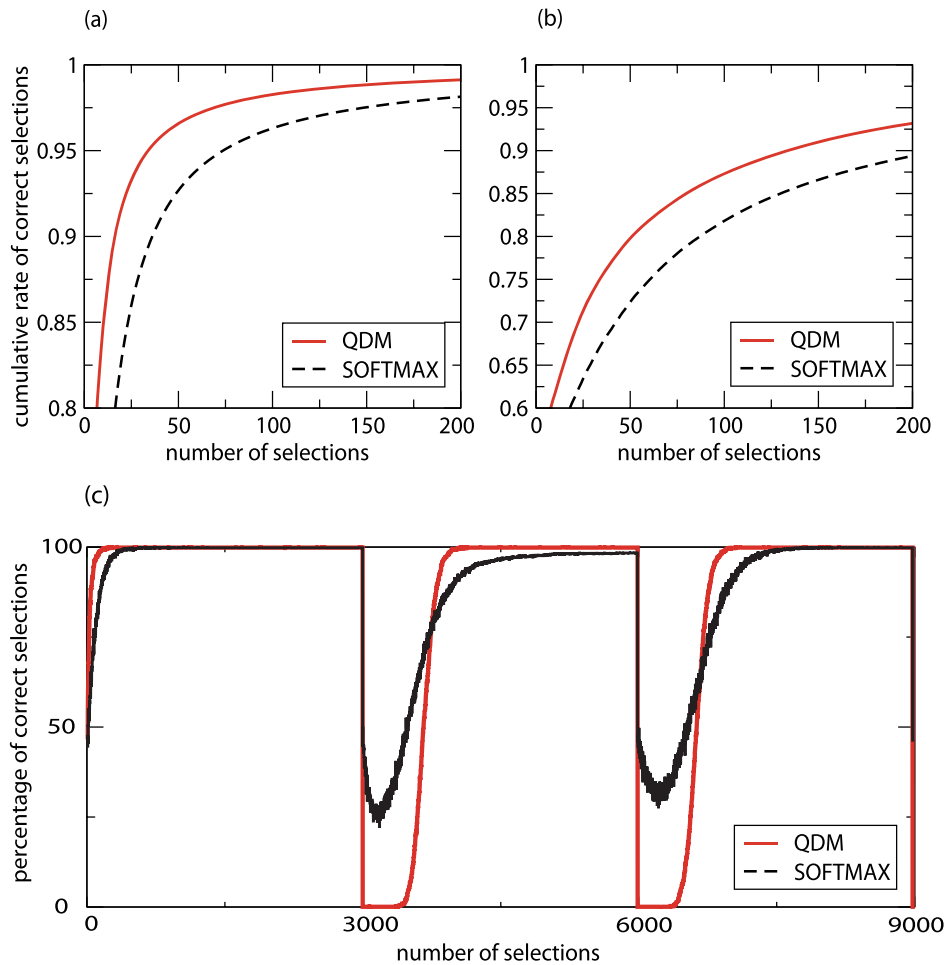
Figure 3 | (a) Efficiency comparison 1. The efficiency comparison between the QDM and the softmax rule is for slot machine reward probabilities of $P_A = 0.2$ and $P_B = 0.8$. The cumulative rate of correct selections for the QDM with fixed parameter $D = 50$ (solid red line) and the softmax rule with optimised parameter $\tau = 0.40$ (dashed line) are shown. (b) Efficiency comparison 2. The efficiency comparison between the QDM and the softmax rule for $P_A = 0.4$ and $P_B = 0.6$. The cumulative rate of correct selections for the QDM with fixed parameter $D = 50$ (solid red line) and the softmax rule with optimised parameter $\tau = 0.25$ (dashed line) are shown. (c) Adaptability comparison. The adaptability comparison between the QDM and the softmax rule for $P_A = 0.4$ and $P_B = 0.6$. In every 3,000 steps, two reward probabilities switch. The percentage of correct selections for the QDM with fixed parameter $D = 50$ (red line), and the softmax rule with the optimised parameter $\tau = 0.08$ (black line) are shown. In this simulation, we used the forgetting parameter $\alpha = 0.999$ (see Methods).

Methods

Calculation of radiation probabilities $S_A(j)$ and $S_B(j)$. Using the quantum Liouville equation (see Supplementary Information), we can calculate the probabilities of radiation from ML_1 and MR_1. We used the following relaxation rate and radiative decay parameters as shown in Fig. 1(c), $\gamma_{S_1} = \frac{1}{(2)^3 1,000}$, $\gamma_{MR_1} = \frac{1}{(\sqrt{2})^3 1,000}$,

$\gamma_{ML_1} = \frac{1}{(\sqrt{2})^3 1,000}$, $\gamma_{LR_1} = \frac{1}{1,000}$, $\gamma_{LL_1} = \frac{1}{1,000}$, $\Gamma_{MR_2} = \frac{1}{10}$, $\Gamma_{LR_3} = \frac{1}{10}$,

$\Gamma_{LR_2} = \frac{1}{100} - \frac{1}{10,000}j + \frac{1}{100,000}$, $\Gamma_{ML_2} = \frac{1}{10}$, $\Gamma_{LL_3} = \frac{1}{10}$ and $\Gamma_{LL_2} = \frac{1}{100} + \frac{1}{10,000}j + \frac{1}{100,000}$, where j represents the position of the intensity adjuster (IA).

Note that we used the following optical near-field interaction parameters so that the QDs are one-dimensionally arranged in "M-L-S-L-M" order for this calculation: $U_{S_1 MR_2} = \frac{1}{10,000,000}$, $U_{S_1 LR_3} = \frac{1}{100}$, $U_{MR_2 LR_3} = \frac{1}{10,000}$, $U_{MR_1 LR_2} = \frac{1}{10,000}$, $U_{S_1 ML_2} = \frac{1}{10,000,000}$, $U_{S_1 LL_3} = \frac{1}{100}$, $U_{ML_2 LL_3} = \frac{1}{10,000}$ and $U_{ML_1 LL_2} = \frac{1}{10,000}$. Here we used $A = 0.1$ and $a = 10^{-9}$ for the parameters in eq.(1). Then, the distance between QD_S and QD_L (r_1) is 20 nm, and the distance between QD_L and QD_M (r_2) is 24.5 nm, such that $U(r_1) = \frac{1}{100}$ and $U(r_2) = \frac{1}{10,000}$. Finally, we obtained two radiation probabilities whose maximum ratio is 8520.82. The difference between radiation probabilities, $S_B(j) - S_A(j)$, is shown by the solid red line in Fig. 2.

Algorithms. *Softmax rule.* In the softmax rule, the probability of selecting A or B, $P'_A(t)$ or $P'_B(t)$, is given by the following Boltzmann distributions:

$$P'_A(t) = \frac{\exp[\beta Q'_A(t)]}{\exp[\beta Q'_A(t)] + \exp[\beta Q'_B(t)]}, \qquad (2)$$

$$P'_B(t) = \frac{\exp[\beta Q'_B(t)]}{\exp[\beta Q'_A(t)] + \exp[\beta Q'_B(t)]}, \qquad (3)$$

where $Q'_k (k \in \{A, B\})$ is the estimated reward probability of slot machines k, denoted by P_k, and 'temperature' β is a parameter. These estimates are given as,

$$Q'_k(t) = \frac{W_k(t)}{M_k(t)}, \qquad (4)$$

$$W_k(t) = w_k(t) + \alpha W_k(t-1), \qquad (5)$$

$$M_k(t) = m_k(t) + \alpha M_k(t-1). \qquad (6)$$

Here $w_k(t)$ ($k \in \{A, B\}$) is 1 if a reward is dispensed from machine k at time t, otherwise 0, and $m_k(t)$ ($k \in \{A, B\}$) is 1 if machine k is selected at time t, otherwise 0. α is a parameter that can control the forgetting of past information. We used $\alpha = 1$ for Figs. 3(a) and (b), and $\alpha = 0.999$ for (c).

In our study, the 'temperature' β was modified to a time-dependent form as follows:

$$\beta(t) = \tau \cdot t. \qquad (7)$$

Here, τ is a parameter that determines the growth rate. $\beta = 0$ corresponds to a random selection, and $\beta \to \infty$ corresponds to a greedy action. Greedy action means that the player selects A if $Q'_A > Q'_B$, or selects B if $Q'_A < Q'_B$.

Dynamics of the intensity adjuster. We adopt the IA to modulate the intensity of incident light to large QDs as shown at the bottom of Fig. 2. The dynamics of the IA also uses estimates Q_k $(k \in \{A, B\})$ which are different from those of the softmax rule. The IA position j is determined by

$$j = floor(Q_B - Q_A), \tag{8}$$

$$Q_k(t) = r_k(t) + \alpha Q_k(t-1), \tag{9}$$

where the function $floor(x)$ truncates the decimal point. $r_k(t)$ $(k \in \{A, B\})$ is D if a reward is dispensed from the machine k at time t, $-D$ if a reward is not dispensed from the machine k and 0 if the system does not select the machine k. Here D is a parameter and α is also a parameter that can control forgetting of past information. We used $\alpha = 1$ for Figs. 3(a) and (b), and $\alpha = 0.999$ for (c).

If we set $\alpha = 1$, the dynamics can be stated by the following rules.

1. If a coin is dispensed, then move the IA to the selected machine's direction, that is, $j = j - D$ for A and $j = j + D$ for B.
2. If a coin is not dispensed, then move the IA to the inverse direction of the selected machine, that is, $j = j + D$ for A and $j = j - D$ for B.

1. Daw, N., O'Doherty, J., Dayan, P., Seymour, B. & Dolan, R. Cortical substrates for exploratory decisions in humans. *Nature* **441**, 876–879 (2006).
2. Cohen, J., McClure, S. & Yu, A. Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Phil. Trans. R. Soc. B* **362**(1481), 933–942 (2007).
3. Robbins, H. Some aspects of the sequential design of experiments. *Bull. Amer. Math. Soc.* **58**, 527–536 (1952).
4. Thompson, W. On the likelihood that one unknown probability exceeds another in view of the evidence of two samples. *Biometrika* **25**, 285–294 (1933).
5. Gittins, J. & Jones, D. A dynamic allocation index for the sequential design of experiments. In: Gans, J. (Eds.), *Progress in Statistics* North Holland, 241–266 (1974).
6. Gittins, J. Bandit processes and dynamic allocation indices. *J. R. Stat. Soc. B* **41**, 148–177 (1979).
7. Auer, P., Cesa-Bianchi, N. & Fischer, P. Finite-time analysis of the multiarmed bandit problem. *Machine Learning* **47**, 235–256 (2002).
8. Kocsis, L. & Szepesvári, C. Bandit based monte-carlo planning. *ECML2006, LNAI 4212,* Springer, 282–293 (2006).
9. Gelly, S., Wang, Y., Munos, R. & Teytaud, O. Modification of UCT with patterns in Monte-Carlo Go. *RR-6062-INRIA,* 1–19 (2006).
10. Agarwal, D., Chen, B.-C. & Elango, P. Explore/exploit schemes for web content optimization. *Proc. of ICDM2009,* http://dx.doi.org/10.1109/ICDM.2009.52 (2009).
11. Gai, Y., Krishnamachari, B. & Jain, R. Learning multiuser channel allocations in cognitive radio networks: A combinatorial multi-armed bandit formulation. *Proc. of DySPAN2010,* http://dx.doi.org/10.1109/DYSPAN.2010.5457857 (2010).
12. Lai, L., Gamal, H., Jiang, H. & Poor, V. Cognitive medium access: Exploration, exploitation, and competition. *IEEE Trans. on Mobile Computing* **10**, 239–253 (2011).
13. Kim, S.-J., Aono, M., Nameda, E. & Hara, M. Tug-of-war model for competitive multi-armed bandit problem: Amoeba-inspired algorithm for cognitive medium access. *Proc. of NOLTA2012,* 590–593 (2012).
14. Lazaar, N., Hamadi, Y., Jabbour, S. & Sebag, M. Cooperation control in parallel SAT solving: A multi-armed bandit approach. *RR-8070-INRIA,* 1–15 (2012).
15. Kim, S.-J., Aono, M. & Hara, M. Tug-of-war model for multi-armed bandit problem. *UC2010, LNCS 6079,* Springer, 69–80 (2010).
16. Kim, S.-J., Aono, M. & Hara, M. Tug-of-war model for the two-bandit problem: Nonlocally-correlated parallel exploration via resource conservation. *BioSystems* **101**, 29–36 (2010).
17. Kawazoe, T. *et al.* Two-dimensional array of room-temperature nanophotonic logic gates using InAs quantum dots in mesa structures. *Appl. Phys. B* **103**, 537–546 (2011).
18. Crooker, S. A., Hollingworth, J. A., Tretiak, S. & Klimov, V. I. Spectrally resolved dynamics of energy transfer in quantum-dot assemblies: Towards engineered energy flows in artificial materials. *Phys. Rev. Lett.* **89**, 186802 (2002).
19. Yukutake, S., Kawazoe, T., Yatsui, T., Nomura, W., Kitamura, K. & Ohtsu, M. Selective photocurrent generation in the transparent wavelength range of a semiconductor photovoltaic device using a phonon-assisted optical near-field process. *Appl. Phys. B* **99**, 415–422 (2010).
20. Kawazoe, T., Mueed, M. A. & Ohtsu, M. Highly efficient and broadband Si homojunction structured near-infrared light emitting diodes based on the phonon-assisted optical near-field process. *Appl. Phys. B* **104**, 747–754 (2011).
21. Kawazoe, T., Ohtsu, M., Akahane, K. & Yamamoto, N. Si homojunction structured near-infrared laser based on a phonon-assisted process. *Appl. Phys. B* **107**, 659–663 (2012).
22. Naruse, M. *et al.* Spatiotemporal dynamics in optical energy transfer on the nanoscale and its application to constraint satisfaction problems. *Physical Review B* **86**, 125407 (2012).
23. Aono, M. *et al.* Amoeba-inspired nanoarchitectonic computing: Solving intractable computational problems using nanoscale photoexcitation transfer dynamics. *Langmuir* **29**, 7557–7564 (2013).
24. Ohtsu, M., Kawazoe, T., Yatsui, T. & Naruse, M. Nanophotonics: Application of dressed photons to novel photonic devices and systems. *IEEE JSTQE* **14**, 1404–1417 (2008).
25. Nomura, W., Yatsui, T., Kawazoe, T., Naruse, M. & Ohtsu, M. Structural dependency of optical excitation transfer via optical near-field interactions between semiconductor quantum dots. *Appl. Phys. B* **100**, 181–187 (2010).
26. Unold, I. T., Mueller, K., Lienau, C., Elsaesser, T. & Wiek, A. D. Optical control of excitons in a pair of quantum dots coupled by the dipole-dipole interaction. *Phys. Rev. Lett.* **94**, 137404 (2005).
27. Vasa, I. P. *et al.* Coherent excitonsurface-plasmon-polariton interaction in hybrid metal-semiconductor nanostructures. *Phys. Rev. Lett.* **101**, 116801 (2008).
28. Franzl, T., Klar, T. A., Schietinger, S., Rogach, A. L. & Feldmann, J. Exciton recycling in graded gap nanocrystal structures. *Nano Lett.* **4**, 1599–1603 (2004).
29. Ohtsu, M., Kobayashi, K., Kawazoe, T., Yatsui, T. & Naruse, M. *Principles of Nanophotonics.* Taylor and Francis (2008).
30. Naruse, M. *et al.* Skew Dependence of Nanophotonic Devices based on Optical Near-Field Interactions. *ACM JETC* **8**, 4:1–4:12 (2012).
31. Naruse, M., Hori, H., Kobayashi, K., Holmstrom, P., Thylen, L. & Ohtsu, M. Lower bound of energy dissipation in optical excitation transfer via optical near-field interactions. *Optics Express* **18**, A544–A553 (2010).
32. Naruse, M. *et al.* Energy dissipation in energy transfer mediated by optical near-field interactions and their interfaces with optical far-fields. *Appl. Phys. Lett.* **100**, 241102 (2012).
33. Akahane, K., Yamamoto, N. & Tsuchiya, M. Highly stacked quantum-dot laser fabricated using a strain compensation technique. *Appl. Phys. Lett.* **93**, 041121 (2008).
34. Pistol, C., Dwyer, C. & Lebeck, A. R. Nanoscale Optical Computing Using Resonance Energy Transfer Logic. *IEEE Micro.* **28**, 7–18 (2008).
35. Cuche, A. *et al.* Near-field optical microscopy with a nanodiamond-based single-photon tip. *Opt. Express* **17**, 19969–19980 (2009).
36. Benson, O. Assembly of hybrid photonic architectures from nanophotonic constituents. *Nature* **480**, 193–199 (2011).
37. Mano, T. & Koguchi, N. Nanometer-scale GaAs ring structure grown by droplet epitaxy. *J. Cryst. Growth* **278**, 108–112 (2005).

Acknowledgements

Author contributions

S.-J.K., M.N. and M.A. designed research, S.-J.K. performed the computer simulations and analysed the data, M.N. and M.A. helped with the modelling. M.O. and M.H. helped with the analysis. S.-J.K. and M.N. wrote the manuscript. All atuthors reviewed the manuscript.

Additional information