Optimal Robot Recharging Strategies For Time Discounted Labour

Jens Wawerla and Richard T. Vaughan
Autonomy Lab, Simon Fraser University
Burnaby, BC V5A 1S6, Canada
{jwawerla, vaughan}@sfu.ca

Abstract

Energy is defined as the potential to perform work: every system that does some work must possess the required energy in advance. An interesting class of systems, including animals and recharging robots, has to actively choose when to obtain energy and when to dissipate energy in work. If working and collecting energy are mutually exclusive, as is common in many animal and robot scenarios, the system faces an essential two-phase action selection problem: (i) how much energy should be accumulated before starting work; (ii) at what remaining energy level should the agent switch back to feeding/recharging? This paper presents an abstract general model of a energy-managing agent that does time-discounted work. Analyzing the model, we find solutions to both questions that optimise's the value of the work done. This result is validated empirically by simulated robot experiments that agree closely with the model.

Introduction

Since the early years of robotics, e.g. Walter (1963), roboticists have been challenged by the need to supply robots with energy. For mobile robots that carry their own limited energy store, the key question is "when should the robot refuel/recharge?". A standard approach in the literature and in commercial robots is to set a fixed threshold and refuel whenever the robot's energy supply drops below this threshold (Silverman et al. (2002), Silverman et al. (2003)). The simplicity of this approach is appealing, but it may not be the optimal strategy. For example Wawerla and Vaughan (2007) showed that, in a realistic surveying task, an adaptive threshold produces a higher overall work rate than a static threshold, and a rate maximizing approach outperforms both by a large margin. The biologically-inspired rate maximizing method performs well, but it has some important limitations discussed below. This paper provides a more sophisticated model for rational recharging robots.

Robots, by their name[1] and nature, are supposed to perform some form of labour, e.g. space exploration, entertainment, rescue missions, clean-up, assembly, etc. Most tasks require execution in a timely manner, though some more then others. For example, we may be willing to wait for a day or two for the latest geological observations from Mars, while waiting the same time for the rescue of trapped miners or ordnance disposal might not be acceptable. The standard method to model the decreasing value of work over time, and thereby encourage timely execution, is to discount by some factor β in discrete timesteps the reward given in exchange for investment (Varian (1992)), where investment here is energy dissipated in labour. The inverse of discount ($1/\beta$) is the familiar interest rate, of savings accounts and credit cards.

The laws of physics dictate that energy cannot be transferred instantaneously, or in other words, refuelling takes time and this time cannot be spent working. If a robot spends an hour refuelling, it starts to work one hour later and since the reward is discounted over time, it receives a smaller payment then if it would have started working immediately. But the initial charging period is strictly required, as no work can be done without previously obtaining energy. This conflict between the mutually exclusive tasks of refuelling and working raises two interesting questions:

Q1 How much energy should be accumulated before starting work?

Q2 At what remaining energy level should the agent switch back to obtaining energy?

Most real-world robot systems avoid these questions by maintaining a permanent connection to an energy source, e.g. industrial robotic manipulators wired into the mains power grid, or solar powered robots which are capable of gathering energy while performing some task at the same time. This paper addresses the more interesting class of machine, including animals and mobile robots, that must obtain and store energy prior to working. The Q1,Q2 action selection problem must be solved by every animal and long-lived robot in some way or another. Further we are only considering rational agents. Any introductory textbook on decision making (e.g. Stuart and Peter (2003)) defines an agent to be rational if it always selects that action, i.e. an answer to Q1 and Q2, that returns the highest expected utility. Here we

[1] Webster: Czech, from *robota*: compulsory labour

assume that utility is proportional to the reward obtained by working, which is discounted over time.

After considering related work, we analyze the problem in terms of a simplified abstract model which, when parametrized to approximate a particular robot system, predicts the optimal answers to Q1 and Q2. We validate the model by comparing its predictions with data empirically obtained from a simulated robot.

To the best of our knowledge, this is the first proposed solution to this general robotics problem.

Related Work

The literature on robotic energy management has many aspects ranging from docking mechanisms, energy-efficient path-planning, to fuel types. The most relevant aspect to this work is on action-selection. Perhaps the most standard and simple way to determine when it is time for the robot to recharge is to set a fixed threshold. This can either be a threshold directly on the energy supply as in Silverman et al. (2002) or on time elapsed since last charging as in Austin et al. (2001). The latter is usually easier to implement but less accurate, because one has to have some model of the energy supply. However, Wawerla and Vaughan (2007) showed a fixed threshold policy can be improved upon. While it is true that maximizing the energy intake rate maximizes potential work rate, it does not optimize with respect to when the work is done. It also *assumes* that recharging is always valuable. This is often true, but not always, as we show below. Also notable is that all of the above papers refuel the robot to maximum capacity at each opportunity, and do not consider that this may not be the best policy.

Litus et al. (2007) consider the problem of energy efficient rendezvous as an action selection problem, and so investigate the where, but not the when and how long, to refuel. Birk (1996) had robots living in a closed ecosystem learn to ‘survive’. Here robots learned to choose between recharging and fighting off competitors. Birk’s agents’ value function of ‘survivability’ is different to that considered here. The rational robot and its owner are interested in gaining maximum reward by working at the robot’s task, and are indifferent to the lifespan of the robot. This is a key difference between the purpose of robots and animals.

Although intended as a wake-up call for psychology research, Toda’s Fungus Eater thought experiment (Toda (1982)) has been influential in the robotics literature. The survival quest of a mining robot on a distant planet contributed significantly to ideas of embodiment and whole agents (Pfeifer (1996)), but the action selection problem presented has yet to be solved in more then the trivial way of a fixed threshold policy.

Spier and McFarland (1997) and McFarland and Spier (1997) investigate work - refuel cycles, or as they call it ’basic cycles’, and show a simple rule, based on cue and deficit, can solve a two resource problem. The cue-deficit policy is inherently reactive and thus fails to cope with the cost of switching between behaviours. Lacking any form of look ahead or planning, so it is difficult to see how it would handle discounted labour situations.

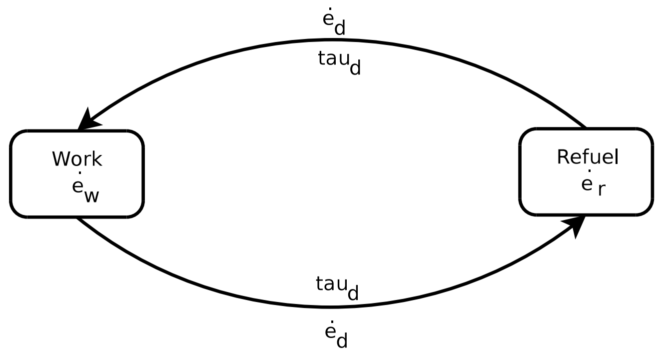

Figure 1: Refuel and time discounted labour model, see text for details

From McFarland’s work, it is a small step into the vast literature on behavioural ecology, from which Houston and McNamara (1999), Stephens and Krebs (1986) and Stephens et al. (2007) are strongly recommended starting points. Due to the biological background of these publications, the analogy to labour and reward in a robotic case is not obvious. The majority of this work uses dynamic programming (DP) as a means of evaluating models of animal behaviour. An exception that does not rely on DP is Hedenström (2003), who investigates the bio-mechanics of land based animals to derive models of optimal fuel load during migration. The model optimizes for migration time, does not map directly into discounted labour or the cyclic work-charge lifestyle of the long-lived robot.

It is known that animals prefer a small, immediate reward to a large delayed reward. So animals seem to do some form of time discounting. According to Kacelnik and Bateson (1996) the reason seems to be that animals in general are risk averse. From a robotics point of view, one major issue with the descriptive models of behavioural ecology is that they lack the ‘how does the animal actually do it’ prescriptive description and hence do not translate readily into robot controllers. Work that tries to bridge the gap between ecology and robotics is Seth (2007). Here Seth uses ALife methods to evolve controllers that obey Herrnstein’s matching law (roughly: relative rate of response matches relative rate of reward), which again is in the domain of rate maximization.

The Model

In this section we describe the behavioural model used in this work. In order to keep the analysis tractable we choose an abstract, slightly simplified model. The world is modelled as two distinct spatially separated sites: a work site and a refuelling site. Moving between sites has a non-zero cost (Figure 1). The robot has a energy storage of $E(t)$, where $0 \leq E(t) \leq E_{max}$. If the energy supply drops to zero anywhere but at the refuelling site, the robot looses is ability to move or work and can gain no more reward. The robot can be in one of four states:

- **refuelling** with a refuelling rate of $\dot{e_r}$, to do so the robot has to be at the refuelling site.
- **transitioning from the refuelling site to the work site**, the duration of this transition is τ_d and the robot has to spend energy at a rate of $\dot{e_d}$, so the transitions cost in term of energy is $\tau_d \dot{e_d}$
- **working**, which gets the robot a reward of $R = \int_{t_0}^{t_0+\tau_w} \beta^t dt$, where t_0 is it time when the robot starts to work and τ_w is the duration the robot works for. Therefore the reward the robot earns by working is discounted with a discount factor $0 < \beta < 1$. While working the robot spends energy with a rate of $\dot{e_w}$. In other words, the robot turns energy into work and therefore reward. In case where the robot performs several work sessions, the reward is accumulated and only the overall reward is at interest to the owner of the robot. As with refuelling, work can only be performed at the work site.
- **transitioning from the work site to the refuelling site**, the duration of this transition is τ_d and the robot has to spend energy at a rate of $\dot{e_d}$

The robot's goal is to achieve as much reward as possible. To do so, it has to make two decisions, (1) when to stop refuelling and resume work and (2) when to stop working and refuel. We mostly refer to the action of accumulating energy as *refuelling* and not as *recharging* because we want to emphasis the general nature of our model.

It is worth pointing out that in a real world scenario all important variables, namely the energy rates, could be known in advance or are easily measured by the robot. Here we assume these variables to be constant, though in an actual implementation we would use averages as approximations. It would also be feasible to do some form of piece-wise linear approximation of the energy rates. The discount factor can also be assumed to be known, since this factor is task dependent and, hence, is set by the owner or designer of the robot or by some external market. As we show below, even if all else is fixed, the robot owner can use the discount factor as a control variable that can be tweaked to fine tune the robot's behaviour. Everything else is predefined by the tasks, the robot's construction or the environment.

In order to improve readability, we need to introduce some additional notation. $k_1 = \frac{\dot{e_r}}{\dot{e_w}}$ is the ratio of the energy rate while refuelling to the rate while working. Similarly, $k_2 = \frac{\dot{e_d}}{\dot{e_w}}$ is the energy rate while transitioning to the energy rate while working. $k_3 = \frac{\dot{e_d}}{\dot{e_r}} = \frac{k_2}{k_1}$ is the ratio of the energy rate while in transition and the energy rate refuelling. τ_r is the time spent refuelling during one refuel-work cycle. The amount of work the robot can perform is limited by the energy supply the robot has, so we express the potential work duration as a function of refuelling and transitioning time where $\tau_w = \tau_r k_1 - 2\tau_d k_2$, which is basically the amount of time the robot can work for, given the amount of energy the robot got from refuelling minus the energy the robot has to spend to travel to the work site and back to the charging station. We also introduce the period of time $T = \tau_r + 2\tau_d + \tau_w = \tau_r(1+k_1) + 2\tau_d(1-k_2)$ as the length of one refuel-work cycle.

When to stop working

Let $e(t)$ be the energy in the robot's storage at time t. At what energy level $e(t) < \epsilon_{w\to r}$ should the robot stop working and transition to the refuelling site? Since the value of work is time discounted, work that the robot performs now is always more valuable then the same amount of work performed later. This creates an inherent opportunity cost in transitioning from the work site to the refuelling site because it takes time and costs energy $\tau_d \dot{e_w}$ that cannot be spent working. This implies that the robot needs to work as long as possible now and not later. Hence the only two economically rational transitioning thresholds are:

- $\epsilon_{w\to r} = \tau_d \dot{e_d}$
  The robot stops working when it has just enough energy left to make it to the refuelling station. The robot will spend the maximum amount of energy, and therefore time, working, ensuring the highest reward before refuelling. Comparatively, should a higher transitional threshold be used, the robot would stop working earlier and refuel earlier, but discounting results in a smaller reward. Should the transitioning threshold be smaller, the robot would have insufficient energy to reach the refuelling station. In this case, the robot cannot gain any further reward because it runs out of fuel between the work and refuel sites.
- $\epsilon_{w\to r} = 0$
  The robot spends all of its energy working and terminates its functionality while doing so. At first glance this option seems counter intuitive, but one can imagine highly discounted labour situations, such as rescue missions, where the energy that would otherwise be spent on approaching a refuelling site is better spent on the task at hand. This might also be a rational option if the transition cost is very high, e.g. NASA's Viking Mars lander took a lot of energy to Mars in the form of a small nuclear reactor, rather than returning to Earth periodically for fresh batteries (the recent Mars rovers employ solar cells to recharge their batteries originally charged on Earth).

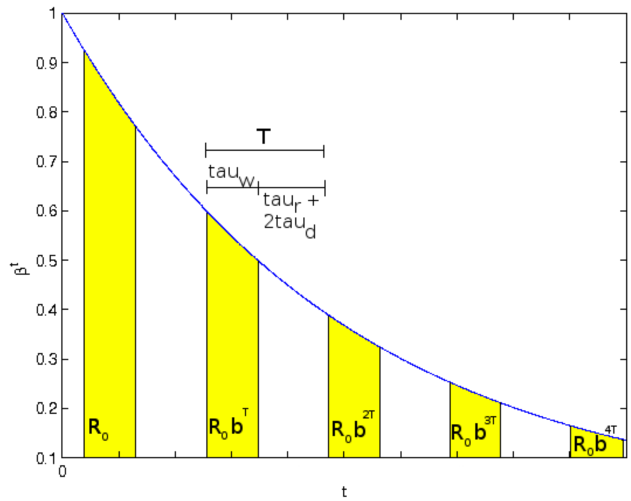

Figure 2: General discounting in refuel - work cycles. The shaded areas are periods in which the robot works and thus earns a reward. The white areas correspond with time in which the robot does not obtain any rewards because it either travels or refuels

Suicide or live forever?

Using our simple model, we can determine whether a robot in a given scenario should terminate while working or continue indefinitely with the work-refuel cycle. Let

$$R_0 = \int_{\tau_r+\tau_d}^{\tau_r+\tau_d+\tau_w} \beta^t dt = \beta^{\tau_r+\tau_d} \frac{b^{\tau_w}-1}{\ln(\beta)} \tag{1}$$

be the reward obtained from spending $E_{max} - \epsilon_{w \to r}$ energy or τ_w time during the first working period (see figure 2). In this figure the shaded areas correspond to time in which the robot performs work and thus obtains a reward proportional to the size of the shaded area. Later work periods are discounted more strongly and hence provide a smaller reward. White areas correspond to times in which no reward is earned because the robot either travels between the work and refuelling site or it refuels. The size of this area is proportional to the opportunity cost, that is, reward that, in principal, could have been obtained if the time had been spent working.

Let T be the duration of one full work-refuel cycle, that is working - transition - refuel - transition, or $T = \tau_w + \tau_d + \tau_r + \tau_d$. Therefore, the reward gained in the next cycle is the initial reward R_0 discounted by T and becomes $\beta^{\tau} R_0$. Subsequent rewards are again discounted by T and so the reward for the third cycle is $\beta^{2T} R_0$. The sum of all rewards if working infinitely, that is choosing $\epsilon_{w \to r} = \tau_d k_2$, is

$$R_\infty = R_0 \sum_{i=0}^{\infty} \beta^{iT} = R_0 \frac{1}{1-\beta^T} \tag{2}$$

In practice no system will last forever, so this analysis is slightly biased towards infinite life histories.

If the robot chooses $\epsilon_{w \to r} = 0$ it gains the initial reward R_0 plus a one time bonus of

$$R_+ = \int_{\tau_r+\tau_d+\tau_w}^{\tau_r+\tau_d+\tau_w+\tau_d k_2} \beta^t dt = \beta^{\tau_r+\tau_d+\tau_w} \frac{\beta^{\tau_d k_2}-1}{\ln(\beta)} \tag{3}$$

by spending the energy required for transitioning on working. The reward gained over the live time of the robot (which is fairly short) is $R_{rip} = R_0 + R_+$.

So the answer to Q2 is that the rational robot selects that threshold $\epsilon_{w \to r}$ that achieves the higher overall reward, so it picks

$$\epsilon_{w \to r} = \begin{cases} 0 & : \quad R_{rip} \geq R_\infty \\ \tau_d \dot{e_d} & : \quad R_{rip} < R_\infty \end{cases} \tag{4}$$

Since the discount function $\int \beta^t dt$ belongs to the class of memory-less functions, we only have to calculate eq. 4 once, in other words if it is the best option to refuel after the first work cycle it is always the best option to do so and vice versa.

How much energy to store

We have shown how to determine a threshold for transitioning from work to refuelling. In this section we will analyze when to stop refuelling and resume work, or phrased differently, how much energy to accumulate before starting to work. Energy and time are interchangeable elements, provided that we know the rate at which energy is spent and gained. Since discounting is done in the time domain, our analysis equates energy with time for simplicity. Based on this, we can ask the time equivalent of Q1: 'how long should the robot refuel for?' We call this refuelling duration τ_r. To be rational, the robot must refuel long enough to gain enough energy to make the trip to the work site and back, that is $2\dot{e_d}\tau_d$, otherwise it would have to turn around before reaching the work site and thus will not gain any reward. Refuelling after the storage tank is full is time wasted that would better be spent obtaining a reward. Therefore the refuelling time is limited to $2\tau_d k_3 \leq \tau_r \leq E_{max}\dot{e_r}^{-1}$. In the following we assume, without loss of generality, the robot starts at the refuelling site with an empty fuel tank. Assuming differently will just result in shift of the analysis by a constant factor, but will not change the overall conclusions.

Acyclic tasks

First we examine situations in which the robot has to refuel for a task that has to be done only once, that is the robot refuels, performs the task, and returns to the refuelling site. Depending on the time spent refuelling the robot obtains the following reward during the upcoming work period.

$$R(\tau_r) = \int_{\tau_r+\tau_d}^{\tau_r+\tau_d+\tau_w} \beta^t dt = \beta^{\tau_r+\tau_d} \int_0^{\tau_r k_1 - 2\tau_d k_2} \beta^t dt \tag{5}$$

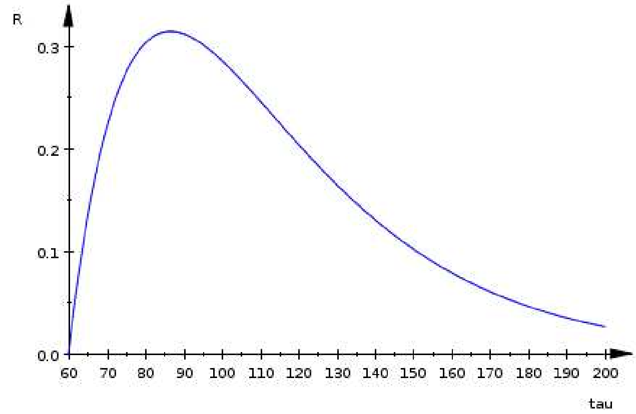

Figure 3: Reward depending on refuelling time with an example configuration $k_1 = 0.5, k_2 = 0.5, \beta = 0.97, \tau_d = 85s$

Next we need to find the $\hat{\tau_r}$ that maximizes $R(\tau_r)$, which is

$$\hat{\tau_r} = argmax(R(\tau_r)) = \frac{\log_\beta\left(\frac{1}{k_1+1}\right) + 2\tau_d k_2}{k_1} \qquad (6)$$

Figure 3 shows an example reward function (eq. 5) depending on the refuelling duration τ_r. Using eq. 6 we calculate that for this particular configuration the reward is maximized when the robot refuels for $\hat{\tau_r} = 86.6234....$ If the fuel tank is filled before that time, the best the robot can do is return to work. This will give it the highest reward achievable, but the designer should keep in mind that there might exist a class of robots with a larger fuel tank that will achieve a higher reward. Note that if the robot stops refuelling at $\hat{\tau_r}$ even if its energy store is not full to capacity, and transitions to working, it earns the highest reward possible. To our knowledge this has not been stated explicitly in the robotics literature before. It is generally assumed that robots should completely recharge at each opportunity, but this is not always the optimal strategy.

Cyclic tasks

In cyclic tasks a robot is required to always return to work after resupplying with energy. Here the analysis is slightly different then in the acyclic case because the refuelling time of the current cycle not only influences the duration and length of the work period of this cycle but of all cycles to come. Hence, we should select a refuelling threshold that maximizes the overall reward. The overall reward is calculated by (see fig. 2)

$$R_\infty(\tau_r) = R_0 \sum_{i=0}^{\infty} \beta^{iT} = \frac{(\beta^{\tau_w} - 1)\beta^{\tau_r+\tau_d}}{(1-\beta^T)\ln(\beta)} \qquad (7)$$

Unfortunately, it seems impossible to find a closed form solution to $\hat{\tau_r} = argmax(R_\infty(\tau_r))$. However, eq. 7 can easily be evaluated for a given τ_r and so calculating the reward for each of a finite set of values for τ_r and selecting the one that maximizes the reward is quite practical. In any real application the number of τ_r to be tested is limited and possibly rather small, in the order of a few thousand. This is because any real robot will have a finite energy storage and any practical scenario will require only limited sampling due to the resolution of the fuel gauge, the uncertainty in the environment, etc. In the case of our Chatterbox Robot (see below), the battery capacity is 2.8 Ah and the fuel gauge has a resolution of 1mA, resulting in less than 3000 calculations for an exhaustive search.

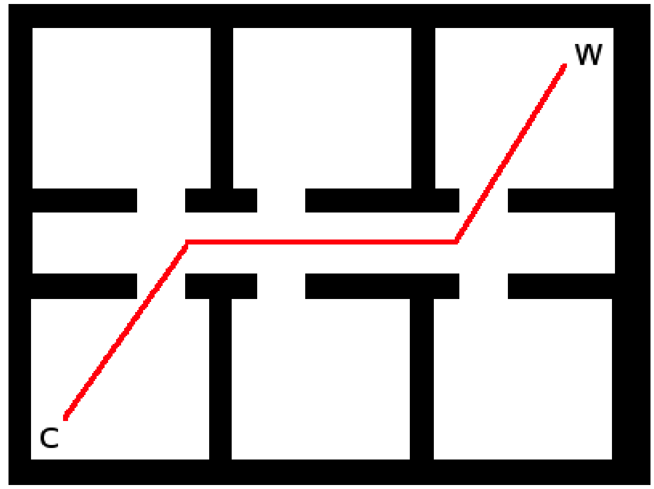

Figure 4: Office like environment with a charging station 'C' and a work site 'W'. The red line is the stylized path the robot travels on.

Experiments

In this section we present experiments to validate the theoretical results described in detail above. All experiments were performed using the robot simulator Stage[2]. The simulated robot uses simulated electrical energy, where we assume charging and discharging to be linear, with constant current for charging I_c, working I_w and driving I_d. We further ignore any effects caused by the docking mechanism, change in battery chemistry or ambient temperature.

In all experiments we roughly model a Chatterbox robot, a robot designed and built at SFU based on an iRobot Create platform. This robot has a battery capacity of approximately 2.8Ah and draws about 2A while driving. We defined an abstract work task which consumes 4A of current. Once at a charging station, the robot docks reliably and recharges with 2A. The world the robot operates in is office-like with one charging station and one work site shown in fig. 4. The obstacle avoidance and navigation controller drives the robot from the charging station to the work site and vice versa in

[2]http://playerstage.sourceforge.net

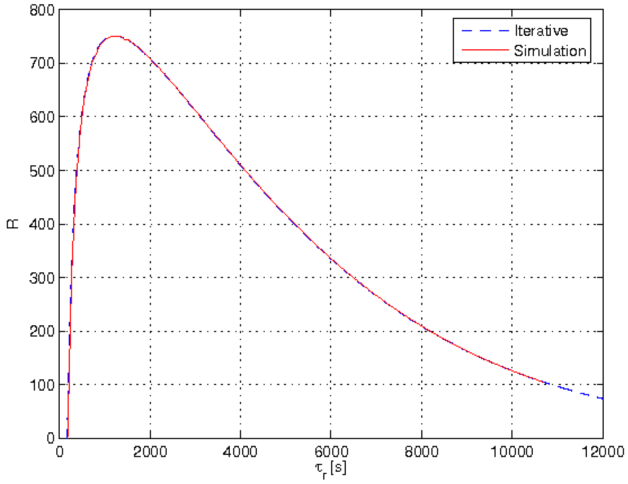

Figure 5: Comparing analytical and simulation results for accumulated reward from a **cyclic** task depending on refuelling time with an example configuration $I_c = 2.0, I_d = 2.0, I_w = 4.0, \beta = 0.9997, \tau_d \approx 85$

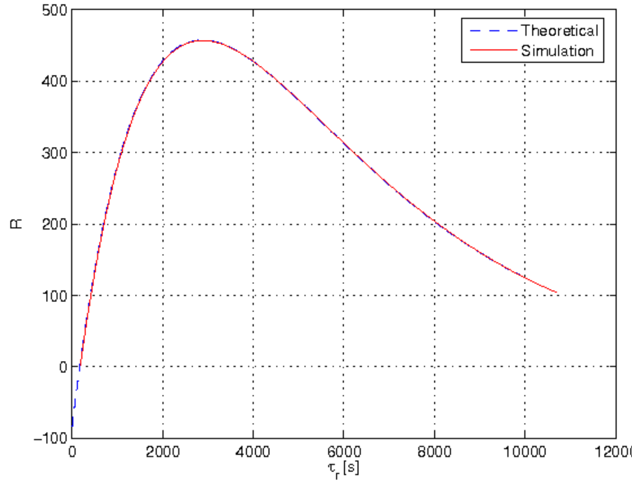

Figure 6: Comparing analytical and simulation results for accumulated reward from an **acyclic** task depending on refuelling time with an example configuration $I_c = 2.0, I_d = 2.0, I_w = 4.0, \beta = 0.9997, \tau_d \approx 85$

approximately $\tau_d = 85s$. Due to naturally occuring noise in the experimental setup the travel time may vary by up to 6 seconds. While working, the robot receives one unit of reward per second, discounted by β. Discounting occurs on a one second basis.

Cyclic Task

The goal of this experiment is to evaluate how closely our analysis from section matches a robot in a simulated environment. In this experiment the robot's task is to recharge for some time τ_r, proceed to the work site, work until the battery energy drops to $\epsilon_{w \to r} = \tau_d I_d$, return to the charging station, and repeat the process. The reward for work is discounted by $\beta = 0.9997$. To find out which τ_r maximizes the reward we varied the threshold for leaving the charging station $\epsilon_{r \to w} = \tau_r I_c$ in each try. A trial lasted for 50000 seconds ($\approx$ 13.8 hours). Figure 5 compares the accumulated reward gained over τ_r from the simulation and from the best solution obtained from the model by iterating over eq. 7. The recharging time that maximizes the reward is predicted by the model to be $\hat{\tau_r} = 1219$ and in the simulation $\hat{\tau_r} = 1170$. The difference comes from the variation in time, and therefore energy, the robot requires to travel between the charging station and the work site. Not only does this change the starting time of work which influences the reward, it also makes it necessary to give the robot a small amount of spare energy to ensure it would not run out of battery. This, in turn, delays charging and thereby influences the reward gained. However, the empirical results agree qualitatively with the values predicted by the model, and the optimal recharging time predicted by the model was within 4% of that observed in the simulation.

Acyclic Task

As before, we perform this experiment in order to compare the theoretical results with a simulation. The setup is the same as in the cyclic task experiment with the difference that the robot only has to perform one charge-work cycle. Figure 6 compares the simulation results to the analytical results. Where the general shape of the curve is similar to that in the cyclic task, it is worth to point out that the maximum reward is gained with a larger charging threshold. This is intuitively correct as the robot has only once chance to obtain a reward. It can be (depending on the discount factor) beneficial to begin work later, but to work for a longer period. For our configuration, the most profitable theoretical charging time is $\hat{\tau_r} = 2872.7$ and the best simulation results were obtained with $\hat{\tau_r} = 2880$. Again the difference between the theoretical and experimental results, barely visible in the plot, are due to imprecision in the robot simulation.

Once or forever

In a further experiment we investigate the circumstances under which it is more profitable, and hence rational, for the robot to fully deplete its energy supply while working and when it is better to choose a perpetual refuelling policy. As in the previous experiments we use a simulated Chatterbox robot with the previously described parameters in the office-like environment. For this scenario, we vary the discount rate between 0.9850 and 0.9999 in 0.0005 increments and run two sets of simulations. In the first, the robot depletes it's energy supply while working, that is, we choose the leave work threshold $\epsilon_{w \to r} = 0$. For the second set, we choose $\epsilon_{r \to w} = \tau_d \dot{e_d}$, a leave work threshold that causes the robot to keep performing work-refuel cycles forever. Since

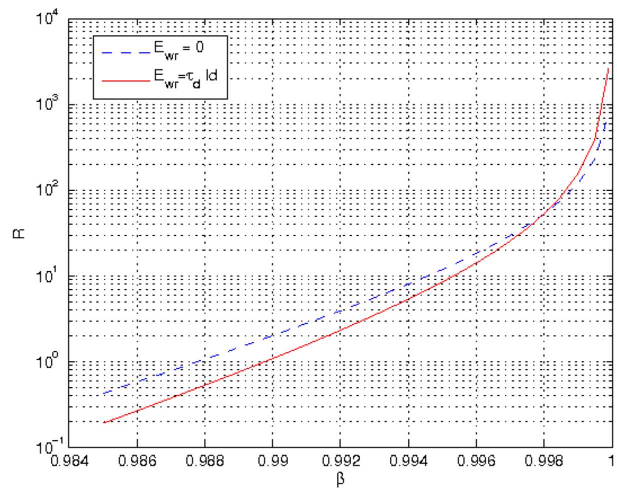

Figure 7: Reward obtained for different discount factors with two leave work thresholds. Configuration $I_c = 2.0, I_d = 2.0, I_w = 4.0, \tau_d \approx 85$

we change the discount rate we have to adapt the leave refuelling site threshold in order for the robot to earn the highest possible reward. For this determine the optimal threshold in the same way as for the previous experiments. Figure 7 depicts the rewards obtained for different discount factors with each policy. As the graph further shows, for higher discounting (smaller discount rate), it is beneficial for the robot to choose a one time work policy. Conversely, for smaller discounting (higher β), it pays to keep working. The theoretical discount rate for switching the policy from one work period to an infinite work refuel cycle is $\beta = 0.9979$, which, as the graph shows, closely resembles the experimental result.

Discussion and Conclusion

We outlined a theoretical analysis of when to refuel and for how long to refuel a robot in situations where the reward for the robot's objective is discounted over time. This discounting is, more often then not, ignored in robotics literature, although it is at the very base of rational behaviour (Stuart and Peter (2003)). We took theoretical results and demonstrated that they apply to a simulated robot. In these simulations we assumed the location of and the distance between work and refuelling station to be known. This is reasonable in the state of the art in mapping and localization, in a wide range of scenarios. We further assumed the average energy spending rates to be constant and known, something achievable in most cases. One assumption made that simplifies a real-world robot scenario is the refuelling rate. Gasoline-powered vehicles which refuel from a standard gas station have a constant refuelling rate, or close to it. However, the charging rate of a battery may depend on many factors including the charging method used, temperature, battery chemistry, and the current capacity of the battery. One useful extension of our model would be to include a realistic chemical battery recharge transfer function.

This paper has presented and analyzed a core action selection problem for autonomous agents such as animals and mobile robots: how much to fuel before working, and when to abandon working and return to fuelling, such that the value of discounted work is maximized. A simple model readily provides answers to these questions and closely predicts the observed behaviour of a robot simulation. While the model is simple, it is very general, and these results suggest that it could be of practical as well as theoretical interest. We propose it as a baseline to build upon.

Acknowledgements

The authors would like to thank Yaroslav Litus, Sarah Brown and Adam Lein for the inspiring discussions that informed this work.

References

Austin, D., Fletcher, L., and Zelinsky, A. (2001). Mobile robotics in the long term - exploring the fourth dimension. In *Proceedings IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 613–618, Wailwa, Hawaii.

Birk, A. (1996). Learning to survive. In *Procs. of the 5th European Workshop on Learning Robots*, pages 1–8, Bari, Italy.

Hedenström, A. (2003). Optimal migration strategies in animals that run: a range equation and its consequences. *Animal Behavior*, 66:631–636.

Houston, A. I. and McNamara, J. M. (1999). *Models of Adaptive Behaviour*. Cambridge University Press.

Kacelnik, A. and Bateson, M. (1996). Risky theories: The effects of variance on foraging decisions. *American Zoologist*, 36:402–434.

Litus, Y., Vaughan, R. T., and Zebrowski, P. (2007). The frugal feeding problem: Energy-efficient, multi-robot, multi-place rendezvous. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pages 27–32, Roma, Italy.

McFarland, D. and Spier, E. (1997). Basic cycles, utility and opportunism in self-sufficient robots. *Robotics and Autonomous Systems*, 20:179–190.

Pfeifer, R. (1996). Building fungus eaters: Design principles of autonomous agents. In *From Animals to Animats, Cambridge, MA: MIT Press*, volume 4.

Seth, A. K. (2007). The ecology of action selection: Insights from artificial life. *Philos. Trans. R. Soc. B*.

Silverman, M., Jung, B., Nies, D. M., and Sukhatme, G. S. (2003). Staying alive longer: Autonomous robot recharging put to the test. Cres-03-015, Center for Robotics and Embedded Systems, University of Southern California.

Silverman, M., Nies, D. M., Jung, B., and Sukhatme, G. S. (2002). Staying alive: A docking station for autonomous robot recharging. In *IEEE International Conference on Robotics and Automation*, pages 1050–1055, Washington D.C.

Spier, E. and McFarland, D. (1997). Possibly optimal decision-making under self-sufficiency and autonomy. *Journal of Theoretical Biology*, 189(3):317–331.

Stephens, D. W., Brown, J. S., and Ydenberg, R. C. (2007). *Foraging*. University of Chicago Press.

Stephens, D. W. and Krebs, J. R. (1986). *Foraging Theory*. Princton University Press.

Stuart, R. J. and Peter, N. (2003). *Artificial Intelligence: A Modern Approach*. Prentice Hall.

Toda, M. (1982). *Man, Robot and Society*. Martinus Nijhoff Publishing.

Varian, H. R. (1992). *Microeconomic Analysis*. W. W. Norton.

Walter, W. G. (1963). *The Living Brain*. W. W. Norton, New York.

Wawerla, J. and Vaughan, R. T. (2007). Near-optimal mobile robot recharging with the rate-maximizing forager. In *European Conference On Artifical Life (ECAL)*, pages 776–785, Lisbon, Portugal.