Knowledge sharing as the key to a sustainable energy balance in embodied AI

Fernando Díaz Ledezma and Sami Haddadin[1]

I. Mathematical framework

In this section we model the scaling of energy and time demand that results from the use of a set of several robots performing numerous skills.

A. Energy and time demand for learning skills

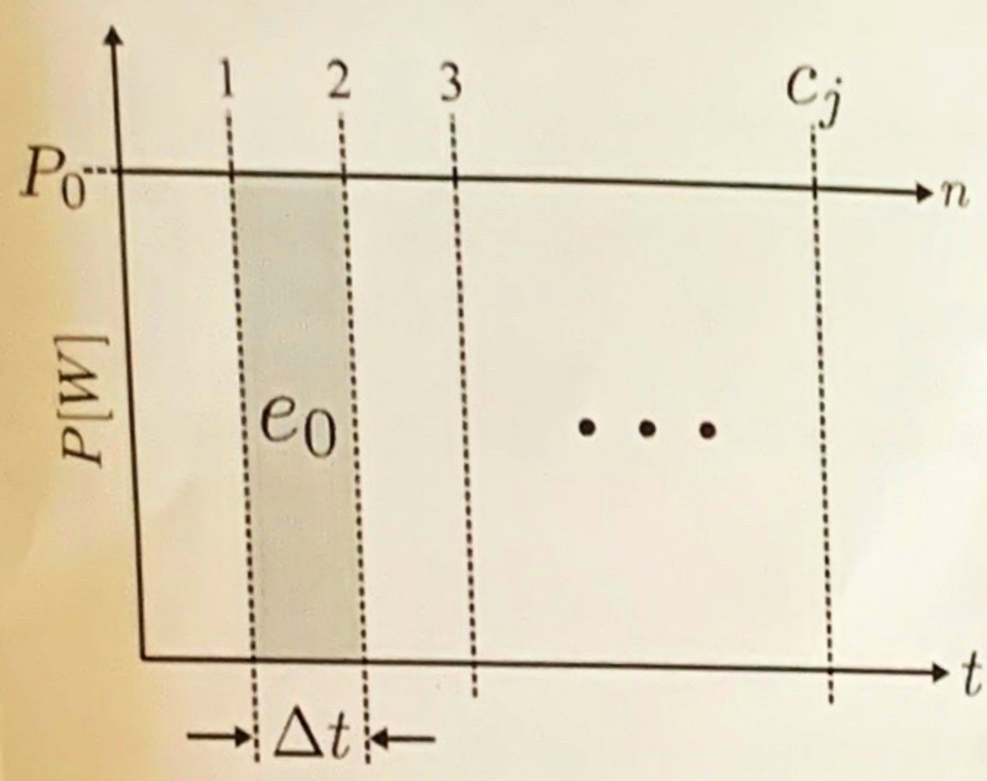


Fig. 1: Power consumption per episode.

> *Definition 1:* The complexity c_j of a skill s_j is ~~understood as~~ the number of trial episodes n needed to successfully learn the skill; i.e., all actions and states visited until a stopping criterion is reached.

Consider P_0 to be the constant power demand of a given robot during the learning of a skill. Furthermore,

> *Assumption 1:* the same amount of time Δt is allocated for the execution of each trial episode n (see Fig. 1).

1) Energy requirement: Under Asm. 1, the energy consumption of the n-th episode $e_j(n)$ is simply

$$e_j(n) = \underbrace{P_0 \cdot \Delta t}_{\text{constant}} = e_0. \tag{1}$$

Consequently, the energy consumed by a robot learning the skill s_j is directly proportional to the complexity; i.e.

$$E_j = \sum_{n=1}^{c_j} e_j(n) = e_0 \cdot c_j. \tag{2}$$

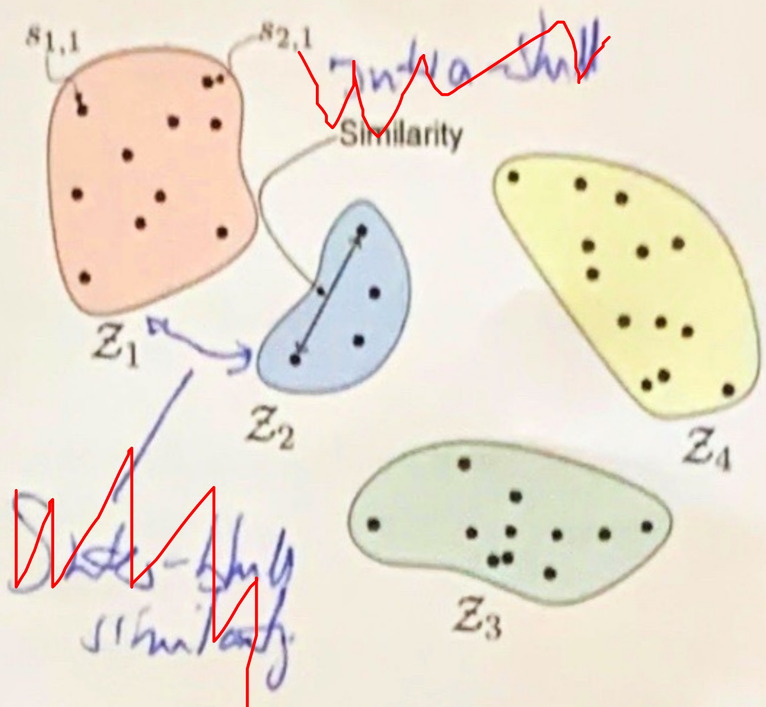


Fig. 2: Similar skills in $\mathcal{S}$ can be grouped into clusters $\mathcal{Z}_k$.

Having $\mathcal{S}$ represent the set of all *to-be-learned* skills, with $|\mathcal{S}| = N_{\mathcal{S}}$; then, the energy spent on learning $\mathcal{S}$ is

$$E_{\mathcal{S}} = \sum_{j=1}^{N_{\mathcal{S}}} E_j = e_0 \sum_{j=1}^{N_{\mathcal{S}}} c_j \tag{3}$$

2) Time requirement: Similarly, the total time $T_{\mathcal{S}}$ is simply

$$T_{\mathcal{S}} = \Delta t \sum_{j=1}^{N_{\mathcal{S}}} c_j. \tag{4}$$

B. Similarity and knowledge

> *Assumption 2:* ~~When~~ the degree of similarity among a set of skills is comparable, they can be clustered together.

Following Asm. 2, let $\mathcal{Z}_k \subset \mathcal{S}$ be a subset of $N_{\mathcal{Z}_k}$ skills that share high similarity; i.e. a *cluster* of skills, see Fig. 2. Furthermore, consider another set $\zeta_k \subset \mathcal{Z}_k$ that denotes already learned skills from $\mathcal{Z}_k$. Asm. 2 implies that the j-th skill in the k-th cluster $s_{j,k} \in \mathcal{Z}_k$ can always benefit from the knowledge contained in ζ_k. This implies that the more skills enter ζ_k, the less knowledge about $s_{j,k}$ will remain to be learned.

Now, we introduce a function $\bar{\sigma}_{j,k}(n) \in [0,1]$ that expresses the knowledge about a skill $s_{j,k} \in \mathcal{Z}_k$ that **is not** contained in ζ_k; i.e. $s_{j,k} \notin \zeta_k$. The function $\bar{\sigma}_{j,k}(\cdot)$ satisfies:

- $\bar{\sigma}_{j,k}(n) = 1$, if $\zeta_k = \emptyset$ or if it does not contain knowledge about the skill $s_{j,k}$

- $\bar{\sigma}_{j,k}(n) = 0$, if *all* the knowledge about skill $s_{j,k}$ is contained in $\mathcal{Z}_k$

Conceptually, $\bar{\sigma}_{j,k}(\cdot)$ is the fraction of knowledge from $\mathcal{Z}_k$ that remains to be learned.

*1) **Leveraging the acquired knowledge**:* To simplify the analysis, we introduce a fundamental complexity.

> *Assumption 3:* The fundamental complexity c_0 describes the maximum number of episodes required to learn *any* skill.

If, in learning a skill $s_{j,k}$, a robot uses the knowledge contained in ζ_k; then, two effects take place:

1) There is less remaining knowledge and this is reflected in the initial value; i.e. $\bar{\sigma}_{j,k}(0) < 1$
2) Knowledge is acquisition is expedited

This implies that the remaining knowledge scales down as a function of the number of learned skills $N_{\zeta_k} = |\zeta_k|$.

> *Assumption 4:* The knowledge function $\bar{\sigma}_{j,k}(\cdot)$ has a monotonically decreasing behavior.

An idealization of the behavior described in Asm. 4 can be modeled via the linear homogeneous differential equation

$$\dot{\bar{\sigma}}_{j,k}(n) = -\alpha f(N_{\zeta_k}, r)\bar{\sigma}_{j,k}(n) \tag{5a}$$

$$\bar{\sigma}_{j,k}(0) = g(N_{\zeta_k}, r) \tag{5b}$$

which is a function of the trial episodes n and is parameterized by the number of already learned skills N_{ζ_k}, the factor $\alpha > 0$ that models the rate at which a robot in isolation learns any given skill and the number r of robots exchanging knowledge about concurrently learned skills. The solution to (5) is given by

$$\boxed{\bar{\sigma}_{j,k}(n) = g(N_{\zeta_k}, r)\left(e^{-\alpha n}\right)^{f(N_{\zeta_k}, r)n} \in (0,1]} \tag{6}$$

and meets the desired behavior, see Fig. 3. The function $f(N_{\zeta_k})$ expresses the form in which the learned skills affect the original rate α. A possible ~~realization~~ choice is to make this effect proportional; i.e.

$$f(N_{\zeta_k}, r) = \frac{1}{(1-\beta_k)}(\eta r N_{\zeta_k} + 1), \tag{7}$$

where η represents how efficient is the exchange of knowledge from ζ_k to $s_{j,k}$. Similarly, $g(N_{\zeta_k})$ models the change in the initial remaining knowledge that results from ζ_k. Although many potential models might be used, we envision the effect to be of exponential nature; i.e.

$$g(N_{\zeta_k}, r) = (1-\beta_k)e^{-\delta r N_{\zeta_k}}, \tag{8}$$

again with a term δ controlling the rate at which the exponential converges as a function of the number of seen skills N_{ζ_k}. In both (7) and (8), β_k is the head start granted by knowledge transfer from other clusters to the skills in $\mathcal{Z}_k$.

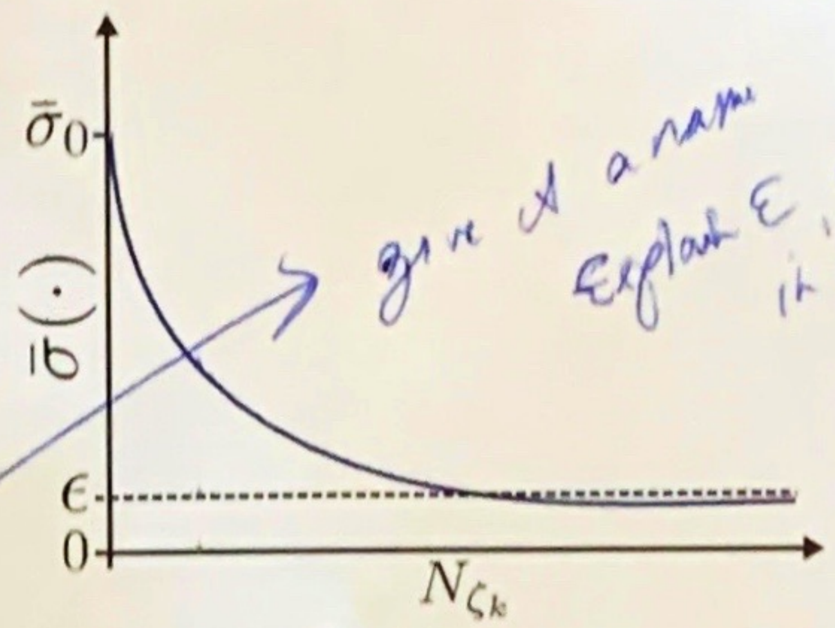


Fig. 3: The ~~idealized~~ remaining knowledge to learn a new skill $s_{j,k}$ as a function of the number of learned skills N_{ζ_k}.

C. Knowledge sharing under different learning paradigms

The following assumptions are analogous to average behavior and imply that a suitable scaling strategy is available and executed.

> *Assumption 5:* Every agent has the same capabilities and energetic cost.

> *Assumption 6:* Every cluster $\mathcal{Z}_k$ contains the same number $N_{\mathcal{Z}}$ of skills.

> *Assumption 7:* The knowledge transferability between skills is assumed to be equal; ~~therefore, also~~ the transferability between clusters ~~is assumed to be equal~~.

*1) **Isolated learning (Iso)**:* A robot learns each skill in $\mathcal{Z}_k$ one after another from the ground up, disregarding the accumulating knowledge from already learned skills; in other words, this implies that $N_{\zeta_k} = 0$ in (6). Therefore, the energy required by one single robot to learn the skills in the cluster $\mathcal{Z}_k$ is given by

$$E^{(Iso)}_{\mathcal{Z}_k} = \sum_{j=1}^{N_{\mathcal{Z}}} E^{(Iso)}_j = N_{\mathcal{Z}_k} e_0 c_{j,k} = N_{\mathcal{Z}} e_0 c_0 \tag{9}$$

If $\mathcal{S}$ is divided into $\{\mathcal{Z}_k\}_{k=1}^{N_{\mathcal{K}}}$ clusters, the total energy to learn the universe of skills is simply

$$E^{(Iso)}_{\mathcal{S}} = N_{\mathcal{K}} N_{\mathcal{Z}} e_0 c_0. \tag{10}$$

Notice that using a batch of m robots to learn the same number of skills in parallel does not bring about any energy reductions since

$$\| E^{(Iso)}_{\mathcal{S}} = \underbrace{m}_{\text{robots}} \cdot \overbrace{\frac{1}{m} E^{(Iso)}_{\mathcal{S}}}^{\text{energy per robot}} = E^{(Iso)}_{\mathcal{S}}. \tag{11}$$

- $\bar{\sigma}_{j,k}(n) = 0$, if *all* the knowledge about skill $s_{j,k}$ is contained in $\mathcal{Z}_k$

Conceptually, $\bar{\sigma}_{j,k}(\cdot)$ is the fraction of knowledge from $\mathcal{Z}_k$ that remains to be learned.

1) ***Leveraging the acquired knowledge:*** To simplify the analysis, we introduce a fundamental complexity.

> *Assumption 3:* The fundamental complexity c_0 describes the maximum number of episodes required to learn *any* skill.

If, in learning a skill $s_{j,k}$, a robot uses the knowledge contained in ζ_k; then, two effects take place:

1) There is less remaining knowledge and this is reflected in the initial value; i.e. $\bar{\sigma}_{j,k}(0) < 1$
2) Knowledge is acquisition is expedited

This implies that the remaining knowledge scales down as a function of the number of learned skills $N_{\zeta_k} = |\zeta_k|$.

> *Assumption 4:* The knowledge function $\bar{\sigma}_{j,k}(\cdot)$ has a monotonically decreasing behavior.

An idealization of the behavior described in Asm. 4 can be modeled via the linear homogeneous differential equation

$$\dot{\bar{\sigma}}_{j,k}(n) = -\alpha f\left(N_{\zeta_k}, r\right)\bar{\sigma}_{j,k}(n) \tag{5a}$$

$$\bar{\sigma}_{j,k}(0) = g\left(N_{\zeta_k}, r\right) \tag{5b}$$

which is a function of the trial episodes n and is parameterized by the number of already learned skills N_{ζ_k}, the factor $\alpha > 0$ that models the rate at which a robot in isolation learns any given skill and the number r of **robots exchanging knowledge** about concurrently learned skills. The solution to (5) is given by

$$\boxed{\bar{\sigma}_{j,k}(n) = g\left(N_{\zeta_k}, r\right)\left(e^{-\alpha n}\right)^{f\left(N_{\zeta_k}, r\right)n} \in (0,1]} \tag{6}$$

and meets the desired behavior, see Fig. 3. The function $f(N_{\zeta_k})$ expresses the form in which the learned skills affect the original rate α. A possible ~~realization~~ choice is to make this effect proportional; i.e.

$$f\left(N_{\zeta_k}, r\right) = \frac{1}{(1-\beta_k)}\left(\eta r N_{\zeta_k} + 1\right), \tag{7}$$

where η represents how efficient is the exchange of knowledge from ζ_k to $s_{j,k}$. Similarly, $g(N_{\zeta_k})$ models the change in the initial remaining knowledge that results from ζ_k. Although many potential models might be used, we envision the effect to be of exponential nature; i.e.

$$g\left(N_{\zeta_k}, r\right) = (1-\beta_k)e^{-\delta r N_{\zeta_k}}, \tag{8}$$

again with a term δ controlling the rate at which the exponential converges as a function of the number of seen skills N_{ζ_k}. In both (7) and (8), β_k is the head start granted by knowledge transfer from other clusters to the skills in $\mathcal{Z}_k$.

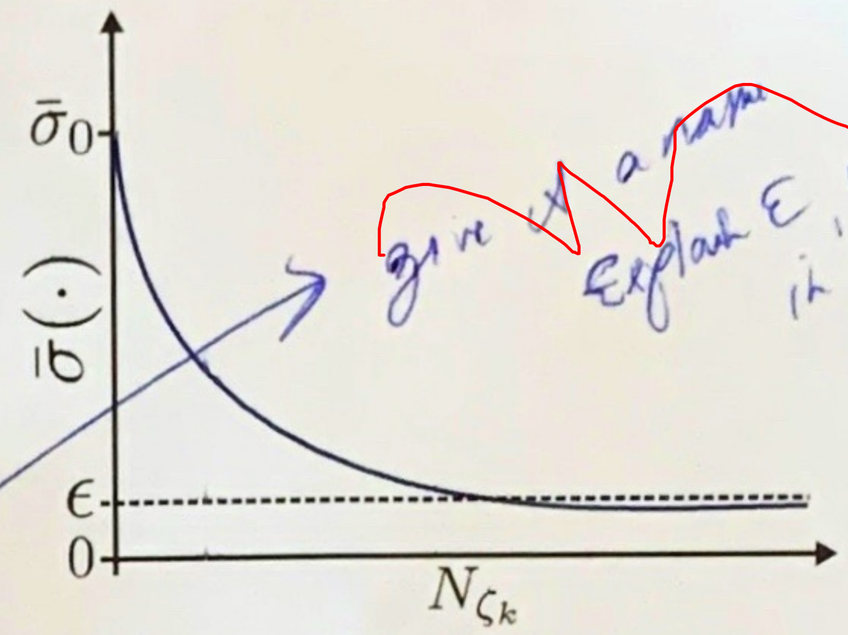


Fig. 3: The ~~idealized~~ remaining knowledge to learn a new skill $s_{j,k}$ as a function of the number of learned skills N_{ζ_k}.

C. Knowledge sharing under different learning paradigms

The following assumptions are analogous to average behavior and imply that a suitable scaling strategy is available and executed.

> *Assumption 5:* Every agent has the same capabilities and energetic cost.

> *Assumption 6:* Every cluster $\mathcal{Z}_k$ contains the same number $N_{\mathcal{Z}}$ of skills.

> *Assumption 7:* The knowledge transferability between skills is assumed to be equal; ~~therefore, also~~ the transferability between clusters ~~is assumed to be equal~~.

1) ***Isolated learning (Iso):*** A robot learns each skill in $\mathcal{Z}_k$ one after another from the ground up, disregarding the accumulating knowledge from already learned skills; in other words, this implies that $N_{\zeta_k} = 0$ in (6). Therefore, the energy required by one single robot to learn the skills in the cluster $\mathcal{Z}_k$ is given by

$$E^{(Iso)}_{\mathcal{Z}_k} = \sum_{j=1}^{N_{\mathcal{Z}}} E^{(Iso)}_j = N_{\mathcal{Z}_k} e_0 \cancel{c_{j,k}}\, c_0 = N_{\mathcal{Z}} e_0 c_0 \tag{9}$$

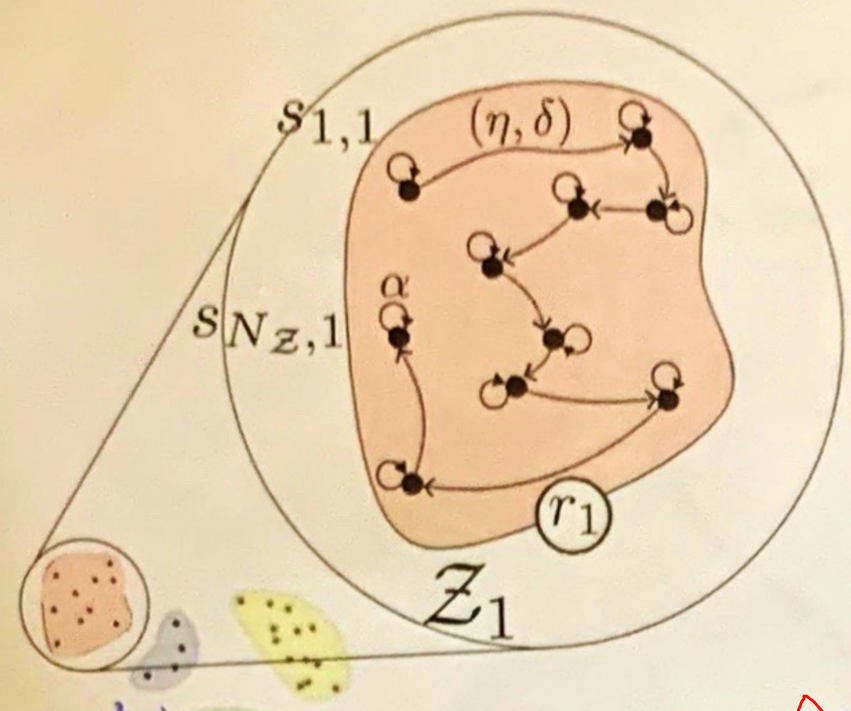
If $\mathcal{S}$ is divided into $\{\mathcal{Z}_k\}_{k=1}^{N_{\mathcal{K}}}$ clusters, the total energy to learn the universe of skills is simply

$$E^{(Iso)}_{\mathcal{S}} = N_{\mathcal{K}} N_{\mathcal{Z}} e_0 c_0. \tag{10}$$

Notice that using a batch of m robots to learn the same number of skills in parallel does not bring about any energy reductions since

$$\| E^{(Iso)}_{\mathcal{S}} = \underbrace{m}_{\text{robots}} \cdot \overbrace{\frac{1}{m} E^{(Iso)}_{\mathcal{S}}}^{\text{energy per robot}} = E^{(Iso)}_{\mathcal{S}}. \tag{11}$$

Assumptions: ML algorithm are able to do control



Assumption: ML + control are able to do it.

Fig. 4: Incremental learning benefits from the high similarity of skills belonging to the same cluster.

significant

significant

2) Incremental learning (I): **INTRA cluster** Referring back to Asm. 2, the knowledge from skills belonging to a cluster $\mathcal{Z}_k$ can be be leveraged by an agent in virtue of their high similarity. As depicted in Fig. 4, a robot (r_1 in this case) learns every skill in $\mathcal{Z}_1$ with a rate *alpha* —the self loops— but also passes the acquired knowledge to the subsequent skills, via the exchange factors (η,δ). Therefore, in incremental learning the knowledge about every new skill gets gradually increased by leveraging previous knowledge, resulting in the expression

$$\bar{\sigma}^{(I)}_{i,j}(n) = e^{-\alpha\left(\eta N_{\zeta_k}+1\right)n} e^{-\delta N_{\zeta_k}},$$

which is equivalent to (6) with $r=1$ since no inter-agent exchange of knowledge occurs and $\beta_k = 0$,, as knowledge from other cluster cannot be used. As the complexity $c_{j,k}$ of a skill can also be interpreted as the number of trial episodes such that the remaining knowledge goes below a threshold ϵ; i.e.

$$\bar{\sigma}^{(I)}_{i,j}(n)|_{n\geq c_{j,k}} \leq \epsilon,$$

then under this scheme the complexity $c_{j,k}$ to learn a new is skill in the cluster results in

$$c^{(I)}_{j,k} = -\frac{\log(\epsilon)-\log(\bar{\sigma}^{(I)}_{j,k}(0))}{\alpha(\eta N_{\zeta_k}+1)} = -\frac{\log(\epsilon)+\delta N_{\zeta_k}}{\alpha(\eta N_{\zeta_k}+1)}. \quad (12)$$

dual behavior are of behavior adversed

The total number of trial episodes C_k that an agent following an incremental learning strategy needs to learn the $N_{\mathcal{Z}_k}$ skills in a cluster $\mathcal{Z}_k$ is given by

Paragraph. Multi-agent case

$$C^{(I)}_k = \sum_{j=1}^{N_Z} c^{(I)}_{j,k}. \quad (13)$$

If m robots are used in parallel to divide the load of learning the tasks then

$$^{\parallel}C^{(I)}_k = m\sum_{j=1}^{\frac{N_Z}{m}} c^{(I)}_{j,k}. \quad (14)$$

learning

In essence, using m robots without exchanging knowledge only subdivides the learning in every cluster into m smaller problems without adding any additional benefit to the rate at which knowledge is acquired.

up to here all fine

we did not yet solve the schedule

3) Transfer learning (TL): **INTER cluster** Considering $\mathcal{K} = \{\mathcal{Z}_k\}_{k=1}^{N_\mathcal{K}}$ to be the set of all available skill clusters (see Fig. 5a), TL represents the exchange of knowledge from the skills learned in different *origin* clusters $\mathcal{O} = \{\mathcal{Z}_1, \mathcal{Z}_2, \ldots, \mathcal{Z}_{k-1}\}$ to the skills that will be learned in a *destination* cluster $\mathcal{Z}_k$. Concretely, the effect that TL has on the skills of the destination cluster is the reduction of the initial remaining knowledge and the initial collection rate for all the skills in the cluster via the parameter β_c in (6), where $0<\beta_c<1$ is the transfer coefficient from the different origin clusters $\mathcal{Z}_c$ to the target cluster $\mathcal{Z}_k$[1]. Asm. 7 implies that

$$\sum_{k=1}^{N_\mathcal{K}} \beta_k \leq 1 \quad (15)$$

really? why?

cannot be implicitly and choice all the same time.

and

$$\beta_k = \frac{k-1}{N_\mathcal{K}}. \quad (16)$$

Notice that this choice of β_k requires that all the skills from the origin clusters were learned. The remaining knowledge when transfer and incremental learning are used in conjunction is

$$\bar{\sigma}^{(I)}_{k,k}(n) = (1-\beta_k)\, e^{-\alpha\left[\frac{\left(\eta N_{\zeta_k}+1\right)}{1-\beta_k}\right]n} e^{-\delta N_{\zeta_k}},$$

Similar to incremental learning the complexity to learn a skill in transfer learning is

$$c^{(IT)}_{j,k} = -\frac{1-\beta_k}{\alpha(\eta N_{\zeta_k}+1)}\left[\log(\epsilon)+\delta N_{\zeta_k}-\log(1-\beta_k)\right]. \quad (17)$$

Similar to incremental learning the total number of episodes C_k that an agent requires to learn the $N_{\mathcal{Z}_k}$ skills is merely their sum

paragraph. Multi-agent case

$$C^{(IT)}_k = \sum_{j=1}^{N_Z} c^{(T)}_{j,k}. \quad (18)$$

If m robots are used in parallel to divide the load of learning the tasks then

across agents

$$^{\parallel}C^{(IT)}_k = m\sum_{j=1}^{\frac{N_Z}{m}} c^{(IT)}_{j,k}. \quad (19)$$

mem-copy → impossible in humans

The effect of this having m robots learning in parallel is depicted on Fig. 5c, where two robots r_1 and r_2 learn skills in four different clusters. The shaded areas are the subclusters of skills learned by each robot. Since they do not share knowledge between them, each robot has access only to the knowledge it has collected and cannot benefit from one another. → humans

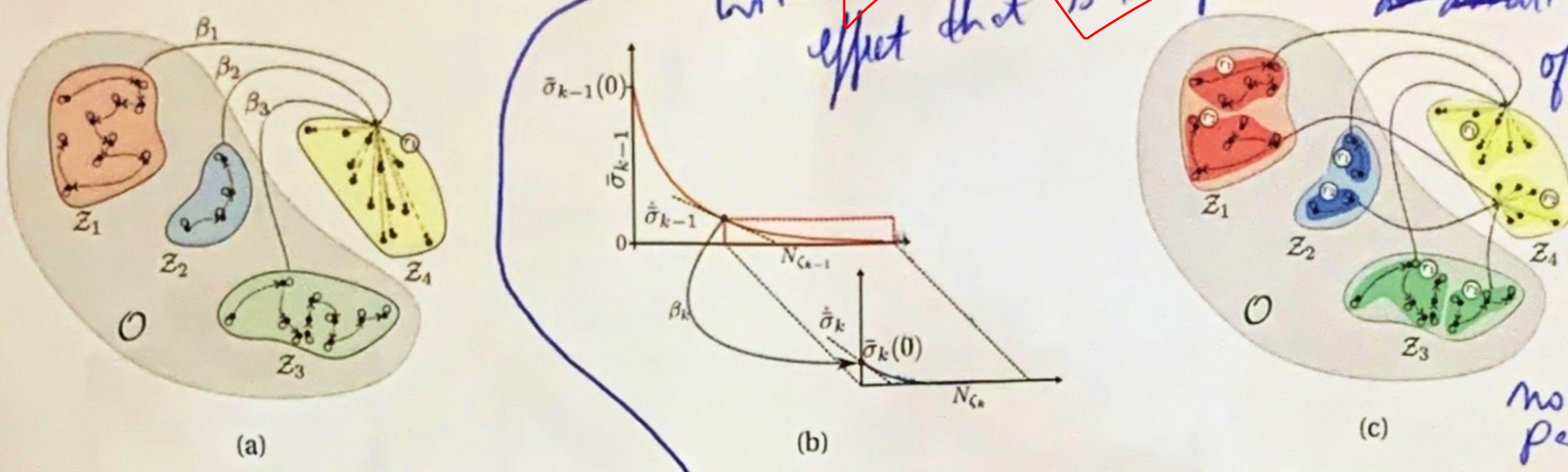
4) Collective learning (C): The key aspect in collective learning is that the exchange of knowledge between a batch of robots $\{r_i\}_1^m$ is now enabled. Fig. 6 illustrates the concept, the self loop represents the dynamics of a single robot learning, expressed by (5). The exchange of knowledge is represented via the cross-couplings weighted by a parameter γ that models how efficient is the bidirectional pairwise

[1] Once again, since there is no inter-agent exchange of knowledge, $r=1$ in this case.

(very complex, occurs both)

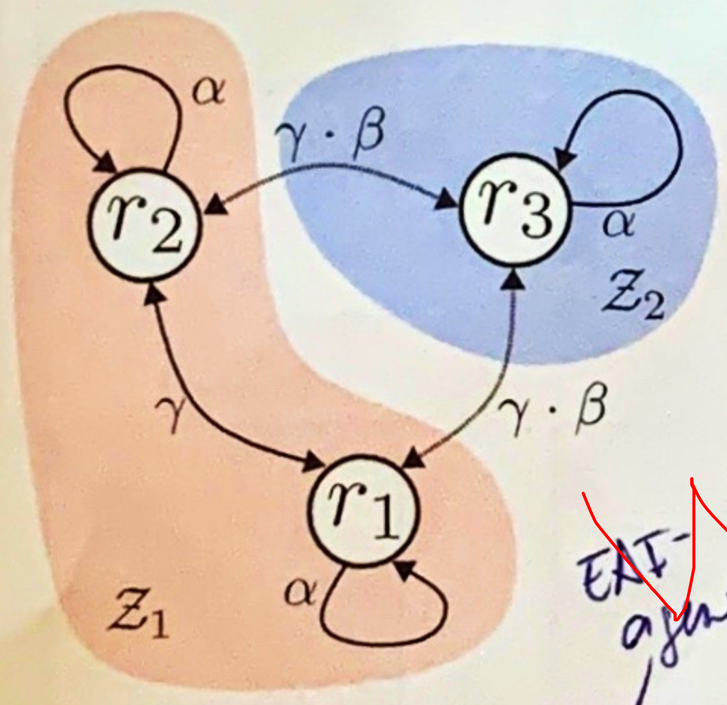
can 2 assumptions: the learning algorithm
- transfer (communication)
- incremental (speaking)

Without that? ⇒ it shows that collective is an effect that is independent of the embodiment but ~~an robot~~ an energy effic of network of transferability (the agents are simply the means to do so, like a novel compute paradigm)



(a) (b) (c)

Fig. 5: Transfer learning. (a) Transfer of knowledge from different origin clusters to the target cluster, (b) the effect of transfer learning, (c) using several robots only subdivides the problem.



EXT-agents ≠ robots

Fig. 6: Exchange of knowledge between ~~robots~~ enables collective learning.

knowledge exchange. Similar to transfer learning, if two robots exchange knowledge about skills with low similarity (i.e. skills in different clusters), then γ is scaled by the inter-cluster transferability parameter β. To account for the couplings (5) is extended to

specify meaning.

$$\dot{\bar{\sigma}}^{(C)}_{j,k}(n) = \left[-\alpha f\left(N_{\zeta_k}, r\right) I + \gamma A \odot B\right] \bar{\sigma}_{j,k}(n) \tag{20a}$$

$$\bar{\sigma}_{j,k}(0) = g\left(N_{\zeta_k}, r\right) I, \tag{20b}$$

with $\bar{\sigma}_{j,k} \in \mathbb{R}^r$ and $A \in \mathbb{R}^{r \times r}$ being the zero-diagonal symmetric adjacency matrix whose entry $(A)_{i,j} = 1$ if robot i exchanges knowledge with robot j an $(A)_{i,j} = 0$ if it does not. The term $\gamma \in \mathbb{R}_+$ weighs the knowledge exchange strength among robots. Furthermore, since robots may learn skills in different clusters at the same time, the matrix B, whose entries are $(B)_{i,j} \in \{1, 1/N_{\mathcal{K}}\}$, scale down the knowledge contributions between robots from different clusters.

Given the dynamics matrix of the collective system

$$\bar{A}\left(N_{\zeta_k}\right) = -\alpha f\left(N_{\zeta_k}, r\right) I + \gamma A \odot B \tag{21}$$

it can be proven that there is a coupling strength γ that ensures that the remaining knowledge for all skills converges asymptotically to zero. Notice that (20) is a skill-varying system by virtue of the dependence of $\bar{A}$ on the number of seen skills N_{ζ_k}.

? What does this mean?

discuss and rewriting

TODO is there an approximation of the convergence rate of all the skills?

D. The effects of the different paradigms

To show ~~what is~~ the effect of the different learning paradigms ~~consider~~ a scenario in which the skill fundamental complexity is $c_0 = 100$ episodes with $N_{\mathcal{S}} = 512$ skills are divided into $N_{\mathcal{K}} = 4$ clusters of $N_{\mathcal{Z}} = 128$ skills each. Additionally, we consider that $m = 32$ robots are available to learn the skills and that a skill is considered learn once the remaining knowledge goes below $\epsilon = 0.01$. In this scenario the robots are used to learn in parallel the skills of each cluster in succession. The remaining parameters are selected as follows;

- $\alpha = 0.0461$
- $\delta = 0.0360$
- $\eta = 0.1$

Why? ⇒ rationalize with 3 scenarios!

The skills learned per robot are shown in Fig. 7, notices that as expected, since incremental learning does not benefit from knowledge from other clusters the robot needs to start accumulating knowledge from the beginning every time it moves to a different cluster. This is not the case in transfer learning, as the more cluster the robot has visited the faster a new skill is learned. The speed of knowledge collection is exponentiated with collective learning thanks to the exchange of knowledge among robots. To assess how the number of robots affects the total number of trial episodes required to learn all the $N_{\mathcal{S}}$ skills we use the same parameters as before but vary the number of robots in $m \in \{2, 4, 8, 16, 32, 64\}$. Additionally we considered an additional collective learning scenario in which, unlike the previous case, the total number of available robots is distributed equally among the clusters to benefit from transfer learning at an earlier time during learning. The results are shown in Fig. 8.

discuss representation

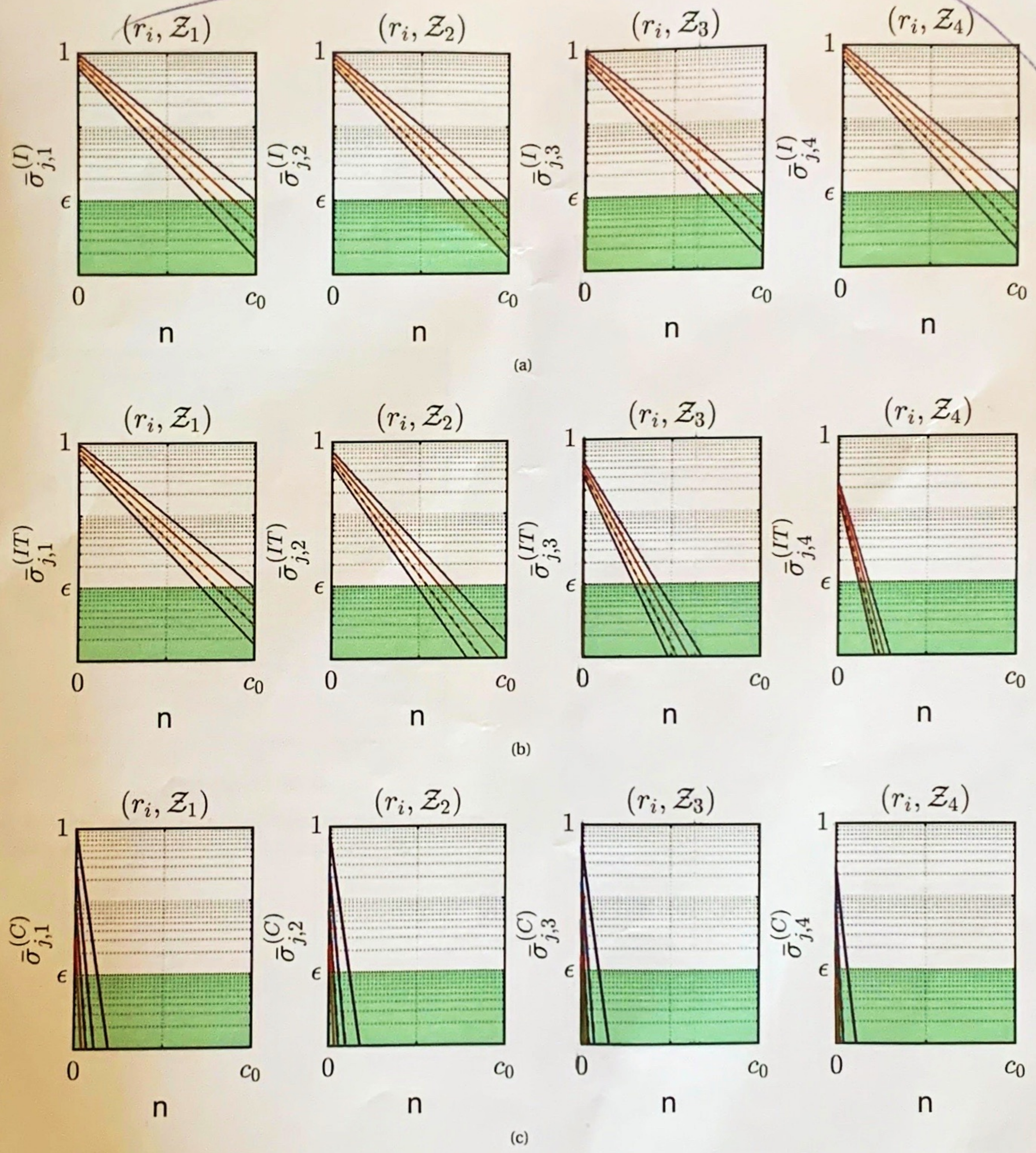


Fig. 7: Scenario 1: the skills of each cluster are learned by the m robots in succession. (a) incremental learning, (b) incremental + transfer learning, (c) collective learning.

Isolated learning is missing.