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

Fernando Díaz Ledezma[*,1], Carlos Magno C. O. Valle[*,1] , Sami Haddadin[1]

I. Mathematical framework

In this section, the goal is to model the demand scaling of energy and time that results from the use of a large number of robots performing a large number of skills.

A. Energy and time demand for learning skills

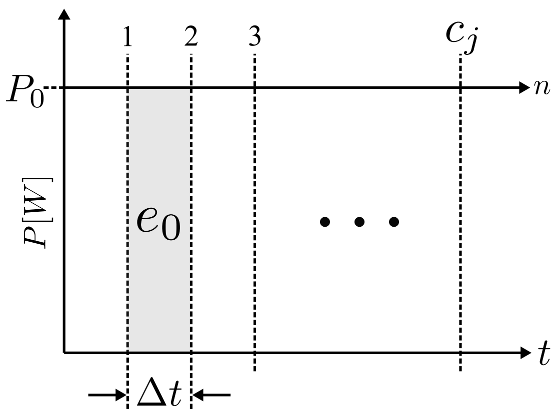


Fig. 1: Power consumption per episode.

> *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.

Now, let P_0 be the total power[1] required by the robot to sustain the learning. Furthermore,

> *Assumption 1:* Every trial episode n takes the same amount of time Δt to be executed (Fig. 1).

*1) **Energy requirement**:* Under Assumption 1, the energy consumption of the n-th episode $e_j^{(n)}$ is simply

$$e_j^{(n)} = \cancelto{\text{const}}{P_0 \cdot \Delta t} = e_0 \tag{1}$$

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

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

[1] P_0 is assumed to be constant.

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. Skill knowledge

Consider a knowledge function $\bar{\sigma}_j(\mathcal{Z}) \in [0, 1]$ that expresses the knowledge from a skill $s_j \in \mathcal{S}$ that **is not** contained in a set of already learned skills $\mathcal{Z} \subset \mathcal{S}$; i.e. $s_j \notin \mathcal{Z}$. The function $\bar{\sigma}_j(\cdot)$ satisfies:

- $\bar{\sigma}_j(\mathcal{Z}) = 1$, if $\mathcal{Z} = \emptyset$ or if it does not contain knowledge about the skill s_j
- $\bar{\sigma}_j(\mathcal{Z}) = 0$, if all the knowledge about skill s_j is contained in $\mathcal{Z}$

Conceptually, $\bar{\sigma}_j(\mathcal{Z})$ is the fraction of knowledge that remains to be learned.

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

> *Assumption 2:* The fundamental complexity c_0 describes the maximum number of episodes required to learn *any* skill.

If, in learning a skill s_j, a robot uses the knowledge contained in $\mathcal{Z}$ about s_j; then, its associated complexity c_j is a scaled down verison of the complexity c_0 and is given by:

$$c_j = c_0 \cdot \bar{\sigma}_j(\mathcal{Z}) \in [0, c_0]. \tag{5}$$

> *Assumption 3:* The skills in $\mathcal{S}$ share a high degree of similarity (have the same complexity).

Assumption 3 implies that a new skill $s_j \in \mathcal{S}$ can always benefit from the knowledge contained in $\mathcal{Z} \subset \mathcal{S}$. This implies that the more skills enter $\mathcal{Z}$ (with $|\mathcal{Z}| = N_j$), the less knowledge about s_j will remain to be learned. Thus, according to (5) the complexity scales down as a function of the number of learned skills, as exemplified in Fig. 2. Alternatively,

$$\lim_{N_j \to N_{\mathcal{S}}} \bar{\sigma}_j = 0 \implies \lim_{N_j \to N_{\mathcal{S}}} c_j = 0. \tag{6}$$

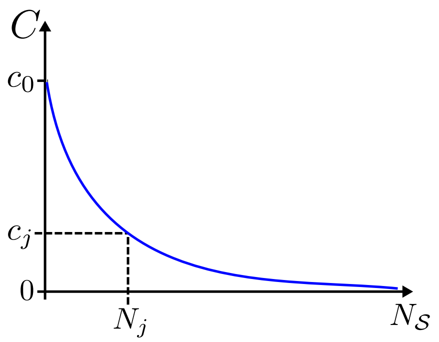


Fig. 2: The complexity of a skill c_j (number of trial episodes) decreases exponentially with the number of learned skills $|\mathcal{Z}| = N_j$.

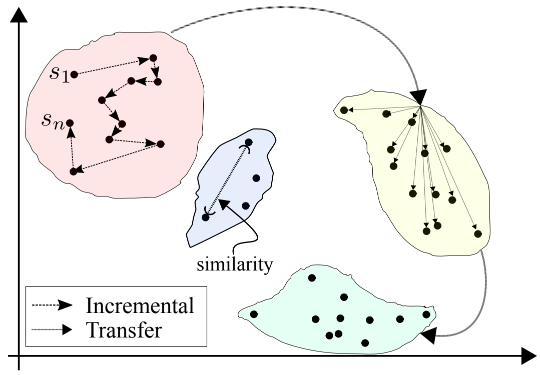


Fig. 3: Incremental and transfer learning and its relation to similarity.

> *Assumption 4:* The knowledge function $\bar{\sigma}(\cdot)$ has an exponentially decreasing behavior.

Considering Assumption 4, an idealization of the behavior described by (6) can be modeled as a decreasing exponential, which is a function of the number of already learned skills N_j:

$$\bar{\sigma}_j = e^{-\alpha \cdot N_j} \in (0, 1], \tag{7}$$

where $0 < \alpha << 1$ models how effectively the knowledge contained in $\mathcal{Z}$ is shared with s_j.

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 abilities and energetic cost.

> *Assumption 6:* When the degree of similarity among a set of skills is comparable, they can be clustered together.

> *Assumption 7:* Every cluster contains the same number of skills.

> *Assumption 8:* 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 all the skills in $\mathcal{S}$ one after another from scratch, disregarding the knowledge from all other learned skills in $\mathcal{Z}$ when learning a new skill. In other words, this implies that $N_j = 0$ in (7). The energy required by the robot to learn all skills in $\mathcal{S}$ is simply

$$\begin{aligned} E_{\mathcal{S}}^{(Iso)} &= N_{\mathcal{S}} \cdot E_j^{(Iso)} \\ &= N_{\mathcal{S}} \cdot e_0 \cdot \cancel{c_j}^{\,c_0} \end{aligned} \tag{8}$$

Furthermore, using batches of m robots to learning m skills in parallel, needs $m \cdot E_j^{(Iso)}$ J. Therefore, there are no energy reductions under this scheme.

2) ***Incremental learning (I):*** Following Assumption 6, skills with comparable similarity are clustered in $\mathcal{K} = \{k_i\}_{i=1}^{N_{\mathcal{K}}}$ clusters, with $N_{\mathcal{K}}$ being the total number of clusters, see Fig. 3. Any two skills belonging to different clusters cannot profit from incremental learning algorithms in virtue of their relatively low similarity. Thus, with (7), the scaling effect that incremental learning has on the skill complexity c_j for the skills contained in the i-th cluster is

$$c_{j,k}^{(I)} = c_0 \cdot \bar{\sigma}_{j,k} = c_0 \left(e^{-\alpha \cdot N_{j,k}}\right)^r = c_0 \left(e^{-\alpha \cdot N_{j,k}}\right), \tag{9}$$

where $r = 1$ and $N_{j,k}$ indicates the number of already learned skills in the k-th cluster. The effect on complexity is shown in Fig. 2. Similarly, $\bar{\sigma}_{j,k}$ indicates the knowledge that is yet to be acquired about a given skill in cluster k. In virtue of the high similarity of the skills in the cluster, the term $\left(1 - \bar{\sigma}_{j,k}\right)$ indirectly reflects the knowledge collected from the skills in the k-th cluster.

The total number of trial episodes C_k that an agent following an incremental learning strategy needs to learn the skills in a cluster k is given by

$$\begin{aligned} C_k^{(I)} &= \sum_{j=1}^{N_{\mathcal{S}}/N_{\mathcal{K}}} c_{j,k}^{(I)} \\ &= \sum_{j=1}^{N_{\mathcal{S}}/N_{\mathcal{K}}} c_0 \cdot \bar{\sigma}_{j,k} = \sum_{j=1}^{N_{\mathcal{S}}/N_{\mathcal{K}}} c_0 e^{-\alpha \cancel{N_{j,k}}^{\,j-1}} \\ &= c_0 \sum_{j=1}^{N_{\mathcal{S}}/N_{\mathcal{K}}} e^{-\alpha(j-1)}. \end{aligned} \tag{10}$$

Likewise, the total number of episodes to learn the skills in all the $N_{\mathcal{K}}$ clusters is

$$\begin{aligned} C_{\mathcal{S}}^{(I)} &= N_{\mathcal{K}} \cdot c_0 \frac{1 - e^{-\alpha \frac{N_{\mathcal{S}}}{N_{\mathcal{K}}}}}{1 - e^{-\alpha}} \\ &= N_{\mathcal{K}} \cdot c_0 \frac{1 - \bar{\sigma}_k^{(I)}}{1 - e^{-\alpha}}. \end{aligned} \tag{11}$$

with $\bar{\sigma}_k^{(I)} = e^{-\alpha \frac{N_{\mathcal{S}}}{N_{\mathcal{K}}}}$ representing the knowledge remaining after learning all the skills in a cluster. If

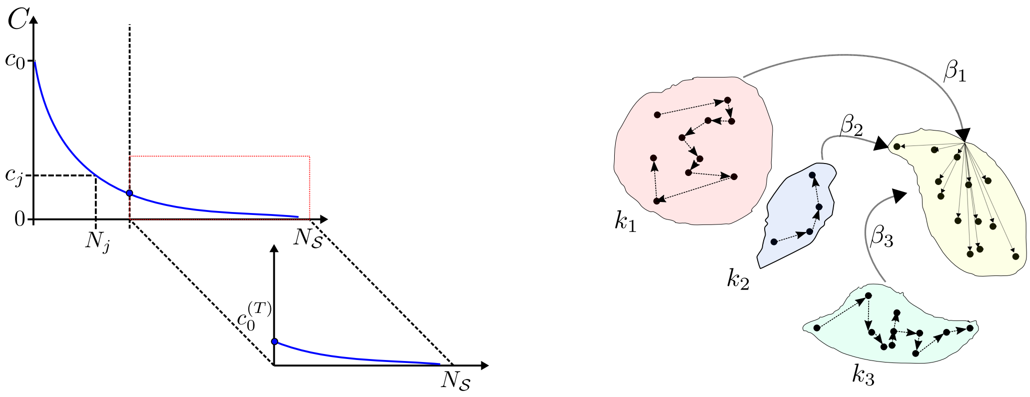


Fig. 4: Transfer learning. (a) The effect of transfer learning, (b) transfer of knowledge from different origin clusters to the target cluster.

m robots are available, each agent can take care of learning $\frac{N_{\mathcal{S}}}{m\cdot N_{\mathcal{K}}}$ skills in parallel with the rest. Under such conditions, the total number of episodes is simply

$$
\begin{aligned}
{}^{\|}C_{\mathcal{S}}^{(I)} &= m\cdot N_{\mathcal{K}}\cdot c_0 \frac{1-e^{-\alpha\frac{N_{\mathcal{S}}}{m\cdot N_{\mathcal{K}}}}}{1-e^{-\alpha}} \\
&= m\cdot N_{\mathcal{K}}\cdot c_0 \frac{1-\bar{\sigma}_m^{(I)}}{1-e^{-\alpha}}
\end{aligned}
\tag{12}
$$

*3) **Transfer learning (TL):*** TL represents the exchange of knowledge from the skills learned in different *origin* clusters $\mathcal{O}$ to the skills that will be learned in a *target* cluster $\mathcal{T}$, see Fig. 3. In general, the effect that TL has on the skills of another cluster is the reduction of the total remaining knowledge to be learned. Referring to (7), it means that its value when $N_{j,k}=0$ will be reduced (Fig. 4a). Considering $\mathcal{K}=\{k_i\}_{i=1}^{N_{\mathcal{K}}}$ to be the set of all available clusters of knowledge (see Fig. 4b), the TL effect can be modeled as follows

$$
\begin{aligned}
\bar{\sigma}_{j,k_{\mathcal{T}}}^{(T)} &= \left\{ e^{-\alpha\left(N_{j,k_{\mathcal{T}}}-\frac{1}{\alpha\cdot r}\log\left[1-\sum_{k\in\mathcal{K}\setminus k_{\mathcal{T}}}\beta_k(1-\bar{\sigma}_{j,k})\right]\right)}\right\}^r \\
&= \left(e^{-\alpha N_{j,k_{\mathcal{T}}}}\right)^r e^{\log\left(1-\sum_{k\in\mathcal{K}\setminus k_{\mathcal{T}}}\beta_k(1-\bar{\sigma}_{j,k})\right)} \\
&= \underbrace{\left[1-\sum_{k\in\mathcal{K}\setminus k_{\mathcal{T}}}\beta_k\left(1-\bar{\sigma}_{j,k}\right)\right]}_{\text{Transfer}}\left(e^{-\alpha N_{j,k_{\mathcal{T}}}}\right)^r,
\end{aligned}
\tag{13}
$$

where $0<\beta_k<<1$ is the transfer coefficient from the different $\mathcal{O}$ clusters to the $\mathcal{T}$ cluster. Additionally,

$$
\sum_{k\in\mathcal{K}\setminus k_{\mathcal{T}}}\beta_k \leq 1,
\tag{14}
$$

assuming that the origin clusters have unique knowledge contributions to the target cluster. Again, $r=1$ in this case. Furthermore, considering Assumptions 7 and 8 $\beta_k = 1/|\mathcal{K}\setminus k_{\mathcal{T}}|$ and $\bar{\sigma}_{j,k}=\bar{\sigma}_m$. Where $\bar{\sigma}_m$ represents the knowledge collected from the other clusters; i.e.

$$
\bar{\sigma}_m = e^{-\alpha\frac{N_{\mathcal{S}}}{mN_{\mathcal{K}}}}.
\tag{15}
$$

Similar to incremental learning, the total number of trial episodes required to learn all the skills in the $N_{\mathcal{K}}$ clusters is

$$
C_{\mathcal{S}}^{(T)} = \left[1-\frac{(1+N_{\mathcal{K}})}{2}\beta\left(1-\bar{\sigma}_k^{(I)}\right)\right]N_{\mathcal{K}}\cdot c_0\frac{1-e^{-\alpha\frac{N_{\mathcal{S}}}{N_{\mathcal{K}}}}}{1-e^{-\alpha}}
\tag{16}
$$

$$
= \left[1-\frac{(1+N_{\mathcal{K}})}{2}\beta\left(1-\bar{\sigma}_k^{(I)}\right)\right]N_{\mathcal{K}}\cdot c_0\frac{1-\bar{\sigma}_k^{(I)}}{1-e^{-\alpha}}
\tag{17}
$$

Furthermore, if m robots can be used in parallel to learn th skills in the clusters; then

$$
{}^{\|}C_{\mathcal{S}}^{(T)} = \left[1-\frac{(1+N_{\mathcal{K}})}{2}\beta\left(1-\bar{\sigma}_m\right)\right]m\cdot N_{\mathcal{K}}\cdot c_0\frac{1-e^{-\alpha\frac{N_{\mathcal{S}}}{m\cdot N_{\mathcal{K}}}}}{1-e^{-\alpha}},
\tag{18}
$$

*4) **Collective learning (TL):*** Finally, in collective learning the notion of cluster is not necessarily applicable anymore, thus $\mathcal{K}=k_{\mathcal{T}}$ which corresponds to the existence of only one big cluster.

$$
\begin{aligned}
\bar{\sigma}^{(C)} &= \cancelto{1}{\left[1-\sum_{k\in\mathcal{K}\setminus k_{\mathcal{T}}}\beta_k\left(1-\bar{\sigma}_{j,k}\right)\right]}\left(e^{-\alpha^{\mathcal{T}}N_j}\right)^r \\
&= \left(e^{-\alpha N_{j,k_{\mathcal{T}}}}\right)^r = \left(e^{-\alpha N_j}\right)^r
\end{aligned}
\tag{19}
$$

Furthermore, now m robots are learning (potentially) m different skills in parallel while exchanging knowledge.

$$
\begin{aligned}
\bar{\sigma}^{(C)} &= \left(e^{-\alpha\cdot N_j}\right)^r \\
&= \left(e^{-\alpha\cdot N_j}\right)^m.
\end{aligned}
\tag{20}
$$

The total number episodes when using collective learning is given by

$$
C_{\mathcal{S}}^{(C)} = m\cdot C_0\frac{1-e^{-\alpha N_{\mathcal{S}}}}{1-e^{-\alpha m}}
\tag{21}
$$

D. Comparison

The general complexity scaling based on knowledge sharing expression is

TABLE I: The total complexity of the different learning schemes.

$c_0\left[1-\sum_{k\in\mathcal{K}\setminus k_{\mathcal{T}}}\beta_k\left(1-\bar{\sigma}_{j,k}\right)\right]\left(e^{-\alpha N_{j,k_{\mathcal{T}}}}\right)^r$			
	Incremental $r=1,\quad \beta_k=0$	**Transfer** $r=1$	**Collective** $r=m,\quad \mathcal{K}\setminus k_{\mathcal{T}}=\emptyset,\quad N_{j,k_{\mathcal{K}}}=N_j$
Single	$C_{\mathcal{S}}^{(I)}=N_{\mathcal{K}}\cdot c_0\frac{1-e^{-\alpha\frac{N_{\mathcal{S}}}{N_{\mathcal{K}}}}}{1-e^{-\alpha}}$	$C_{\mathcal{S}}^{(T)}=\left[1-\frac{(1+N_{\mathcal{K}})}{2}\beta(1-\bar{\sigma})\right]C_{\mathcal{S}}^{(I)}$	${}^{\|\|}C_{\mathcal{S}}^{(C)}=m\cdot c_0\frac{1-e^{-\alpha N_{\mathcal{S}}}}{1-e^{-\alpha m}}$
Parallel	${}^{\|\|}C_{\mathcal{S}}^{(I)}=m\cdot N_{\mathcal{K}}\cdot c_0\frac{1-e^{-\alpha\frac{N_{\mathcal{S}}}{m\cdot N_{\mathcal{K}}}}}{1-e^{-\alpha}}$	${}^{\|\|}C_{\mathcal{S}}^{(T)}=\left[1-\frac{(1+N_{\mathcal{K}})}{2}\beta(1-\bar{\sigma}_m)\right]{}^{\|\|}C_{\mathcal{S}}^{(I)}$	

$$c_j=c_0\left[1-\sum_{k\in\mathcal{K}\setminus k_{\mathcal{T}}}\beta_k\left(1-\bar{\sigma}_{j,k}\right)\right]\left(e^{-\alpha N_{j,k_{\mathcal{T}}}}\right)^r \tag{22}$$

where according to (23), the effects of the different learning schemes are reflected.

$$c_j=\begin{cases}\text{Incremental} & \alpha\neq 0,\beta=0,r=1\\ \text{Transfer} & \alpha\neq 0,\beta\neq 0,r=1\\ \text{Collective} & \alpha\neq 0,\beta=0,r=m\end{cases} \tag{23}$$

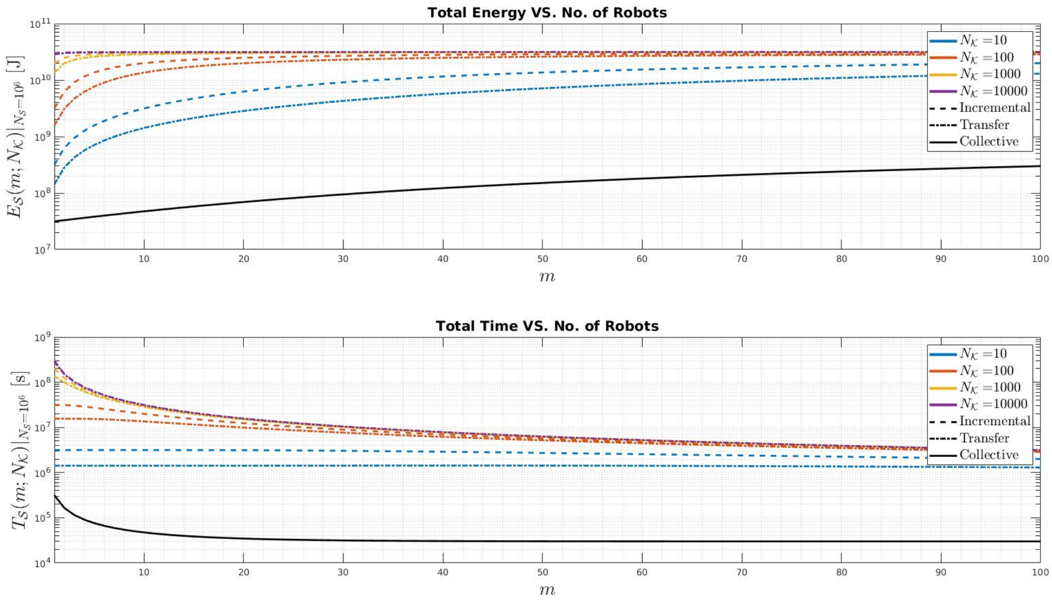
From (12) and (18) it is clear that they differ only by the factor $\left[1-\frac{(1+N_{\mathcal{K}})}{2}\beta(1-\bar{\sigma}_m)\right]\leq 1$, showing that TL is a scaled down version of IL. Table I shows the comparison between the total number of trial episodes required by the three learning schemes.

The total energy and time consumption in the different learning paradigms is

$$\begin{aligned}E_{\mathcal{S}}^{(\star)}&=e_0\cdot{}^{||}C_{\mathcal{S}}^{(\star)}\\ T_{\mathcal{S}}^{(\star)}&=\frac{\Delta t}{m}\cdot{}^{||}C_{\mathcal{S}}^{(\star)},\end{aligned} \tag{24}$$

with $\star$ representing the total complexity corresponding the considered learning scheme. Fig. 5 shows the effect in the energy and time demand from the three different learning paradigms as a function of the number of robots learning, considering $c_0=100$, $P_0=100\,\mathrm{W}$, $\Delta t=300\,\mathrm{s}$ and $N_{\mathcal{S}}=10^6$. Some insights can be extracted...

⚠ **IMPORTANT:** Discuss the figure.



(a)

Fig. 5: The total energy and time demands for the different learning paradigms, considering $N_{\mathcal{S}} = 10^6$.