Imperial College London

# TRANSFER LEARNING: AN APPLICATION OF FOLIATIONS

JANITH PETANGODA, NICK A. M. MONK & MARC DEISENROTH

The University Of Sheffield.

## GOAL

Our goal is to introduce foliations as a useful framework to formalise transfer learning (TL), and foster more precise, foundational and deliberate discourse and enquiry of the TL problem.

## TRANSFER LEARNING

Transfer learning is an aspect of machine learning (ML) that has garnered relatively recent interest. It deals with the problem of transferring learned knowledge between *related tasks*. For example, consider the RL problem in Figure 1.

Many experimentally successful works exist in this space [1, 2]. However, these do not present transfer (or meta) learning through a formal framework. We cannot understand why nor when they would work.

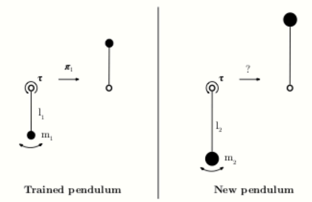


**Figure 1:** Example of transfer with a pendulum

## REPRESENTATION, RELATEDNESS AND SIMILARITY

**Representation** A representation is a scheme by which we can describe a set of abstract objects. For example, the abstract number 42 can be written as 'forty-two' in English, 42 in decimal and 101010 in binary. In ML, we want to find an *approximate description* of a map in terms of a *chosen representation* scheme (the model), from the given data. The representation scheme reflects our assumptions about the problem.

**Relatedness** This definition of relatedness borrows from [3]. We define 2 tasks $f, g$ in a set of tasks $T$ to be related if, given a group of transformations $\Pi$ that act on $T$, there exists $\pi \in \Pi$ such that $\pi(f) = g$, and $\pi^{-1}(g) = f$. Relatedness is a transformative notion.

**Equivariance** We can further assume that if two tasks are *related*, then their solution in the space of models $M$ are also related. That is, the learning algorithm, is an *equivariant map*. See Figure 2.

**Similarity** We can make similarity distinct from relatedness. 2 tasks $f, g \in T$ are $\epsilon$-similar, if, under some metric $\rho$ on $T$, $\rho(f, g) \leq \epsilon$. Similarity is a geometric notion. See Figure 3.

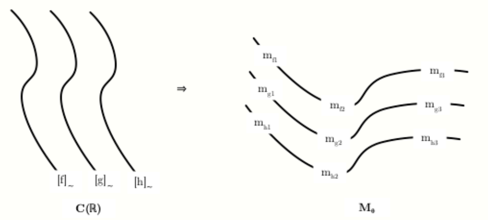


**Figure 2:** Assumption of equivariance

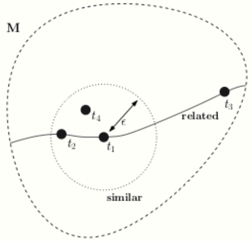


**Figure 3:** Relatedness vs Similarity

## RELATED TASKS, INVARIANT QUANTITIES AND PARALLEL SPACES

How can we build a representation scheme that reflects a notion of relatedness, as defined?

**Set of related tasks** Under $\Pi$, we can find an equivalence class of tasks. That is, $[f]_\sim \subseteq T$, is an equivalence class containing $f \in T$ where equivalence relation $\sim$ is given by $f \sim g \iff f = \pi(g)$ for some $\pi \in \Pi$. These sets are the orbits of the group action of $\Pi$.

**Invariant quantities** A quantity on $T$ is a map $q : T \to \mathbb{R}$. This quantity is invariant w.r.t a group of transformations $\Pi$ if the value of the quantity does not change under any $\pi \in \Pi$.

If the action of the transformation group is *regular*, then the number of independent invariant quantities is completely determined by the dimension of the orbit. The collection of invariant quantities then give us a notion of *what is the invariant structure within a set of related tasks.*

**Parallel Spaces** If we take the union all the sets of related tasks, we obtain a partition of the space of tasks $T$. In particular, since these are the orbits of group actions, each partition looks like a *parallel space* in $T$, as shown in Figure 4.

If we considered similarity instead of relatedness, we can also obtain a partition of $T$. However, this partition would look like a tessellation.

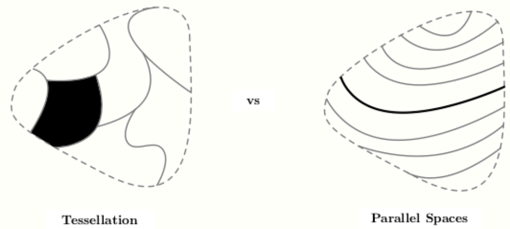


**Figure 4:** Partitioning of $T$ under similarity vs relatedness.

## FOLIATIONS

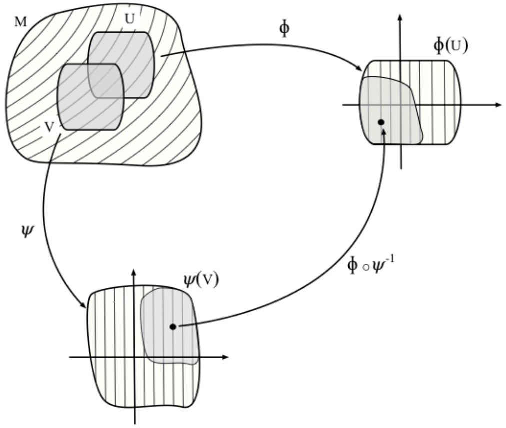


**Figure 5:** Foliations

A foliation is a formal way of mathematically describing a notion of parallel spaces. It is an additional structure that can be placed on a (smooth) manifold. A foliation is defined by a restriction on the atlas of the manifold.

In particular, a foliation defines a way to consistantly define a locally rectified coordinate system on the manifold. That is, we choose an atlas, where the chart transitions keep some coordinates fixed. These consistency requirements mean that when the local open sets are glued together, we get back parallel spaces. In its theory, these parallel spaces are called *leaves*.

Foliations naturally define locally invariant quantities. This is because, locally on a leaf, a subset of the dimensions of the coordinates must, by definition, remain constant.

## REFERENCES

[1] Sinno Jialin Pan and Qiang Yang. A survey on transfer learning. *IEEE Transactions on knowledge and data engineering*, 22(10):1345–1359, 2009.

[2] Timothy Hospedales, Antreas Antoniou, Paul Micaelli, and Amos Storkey. Meta-learning in neural networks: A survey. *arXiv preprint arXiv:2004.05439*, 2020.

[3] Shai Ben-David and Reba Schuller. Exploiting task relatedness for multiple task learning. In *Learning Theory and Kernel Machines*, pages 567–580. Springer, 2003.

## CONTACT INFORMATION

**Email:** jpetangoda@gmail.com
**Twitter:** @janithpet