**Skill Value Determination by Mapping Conceptual Hierarchy**

Skills, by definition, are concepts meant to describe a particular ability developed through training and experience.

This ability is composed of several components, or sub-skills. For example, the ability of a surgeon to perform **surgery** requires, among other skills, the skills of **stitching**, **anatomy** **proficiency** and **dexterity**. Without one of those, the doctor cannot solely perform **surgery**. If one of those is not automated, the skill is not automated.

The amount and composition of those sub-components is important to determine how likely a skill is to be automated. Repetitive jobs requiring non-complex skills disappear rapidly, while complex ones enjoy higher demand than ever.

To detect long-term valuable skills, a method to determine skill complexity and skill long-term value in the job market is needed. Such a method is presented here, named **mapping conceptual hierarchy**.

**Mapping conceptual hierarchy** is the use of the links between concepts to understand their relation.

It does it by determining how many concepts are needed to be known in advance in order to define the current connect, and how many concepts rely on it, or in other words, how abstract and how fundamental a concept is.

As skills are concepts by definition, the method can be applied on them.

The method is relatively simple and effective, can project skill value for long periods of time, and can map projected directions of technological advancement, and with it the resulting skill demand changes.

**Abstraction** is the ability to use concepts to define more complex concepts, enabling humans to deliver huge amounts of information in one sentence, or even one word.

For example, we use the concept **dexterity** to describe **coordination** of **small** **muscles**, and particularly **fingers**, with **human** **eyes**. Each of these **concepts** is also defined by **other concepts,** until raw sensual data (precept) is reached.

The amount of concept layers needed to define a concept is called “abstraction level”. The higher the abstraction level, the more **complex** the concept is and the harder it is to automate it.

The more concepts rely on another concept, the more **fundamental** it is.

Technological advancement is the main force behind skill demand shifts- process automation due to technological advancement, obsolescence of skill due to technological advancement, and demand for new skills to exploit technological advancement are the main driving factors behind such shifts.

So, for a skill to have long-term value, is must be either protected from technological advancement or needed when it occurs.

This happens the skill is or has one of the followings:

1. Regulation protecting its holder (like entry license), i.e. **protected skill**.

2. High complexity that prevents automation, i.e. **complex skill**.

3. It is a foundation of other needed skills, i.e. **foundational skill.**

**Protected skills** get their value from restriction rather than demand. Therefore, they are hard to predict and not covered by this method.

**Complex skills** are hard to automate. Therefore, ensuring proper supply of skill holders is expected to be economically valuable, especially in the long-run.

**Foundational skills** are key to learning other needed skills. Therefore, instruction of such skills is expected to be economically valuable as technology progresses, as more and more layers of skills are expected to be built on them.

**Mapping conceptual hierarchy** maps how much technological advancement is expected to affect a field, and in which fashion, enabling us to project what skill have long-term value.

This particular implementation analyzes how **complex** and **fundamental** is a skill using the **NLTK natural language toolkit**, and particularly the **WordNet lexical database** components **NLTK** provides, as its **data sources**.

**Complex skills** are found by analyzing abstraction level, i.e. how many concepts are needed to define the examined skill.

In this solution, it is done by extracting concepts from lexical definitions, in the following fashion and implementation:

1. Search **skill (S)** lexical definition for **concepts**.

2. Search **each found concept (SC)** definition for **concepts**.

3. Repeat the process until you get only **precepts** (building blocks of concepts) or **circular concepts**.

4. The amount of **concept layers** needed to build **your** **concept** is the **abstraction level**.

**Foundational skills** are found by analyzing reliance level, i.e. how many other skills need to use the examined skill in order to be defined.

In this solution, it is done by mapping concept prevalence in lexical dataset in the following fashion and implementation:

1. Search **all** **skill** lexical definitions for definitions containing the **concept (S)**.

2. For each one found (SD), search **all skill** lexical definitions and check which contain **it (SD)**.

3. Repeat the process until you reach the **highest** **abstraction level** or **circular definition.**

4. The amount of **concept layers** relying on **your concept** for definition is the **reliance level**.

The dataset can be optimized to ease skill searches, by finding the **abstraction level** and **reliance level** for **each concept** in advance and adding it to the dataset. This will reduce needed computing resources dramatically, as only new, unmapped concepts will be analyzed, and analysis will be limited in size to the unmapped concepts chain size.

To save resources, a recursion depth limit can be set in both **abstraction level** and **reliance level** analyses. Once this limit is reached, the appropriate **compound abstraction level (CAL)** or **compound reliance level (CRL)** is increased by one.

If such limit is set, the **total abstraction/reliance level** will be a multiplication of the **depth limit value** by the appropriate **compound parameter level.**

Adjusting for previously automated concepts is rather straightforward- if a concept has been automated, its **abstraction level** will be zero.

Generating a dataset of previously automated concepts can be done by finding concepts that everybody spoke of automating, but no does anymore, that are also part of the definition of a concept everybody speak of automating now.

The reason is that such a concept is probably automated, as automation of concept of higher abstraction level demands automation of the previous abstraction level.

One potential source for such data can be a dataset of software and hardware product selling points over time. The reason is that the highlights the manufacturer tries to present to prospective clientele are expected to be advantages other products dont have, and the fact the seller stopped mentioning then indicates they are already standard.

This is targeted to be explored in detail in future work, and is out-of-scope in this one.

A skill containing only **zero-level abstraction concepts** can be automated.

If it is not, either it is not economical to automate it or it is protected by regulation.

The way to differentiate the two is to check the skill's **reliance level**, i.e. how many skills depend on it- if it is high, is it safe to assume regulation is protecting it, and vice-versa.

This is targeted to be explored in detail in future work, and is out-of-scope in this one.

This method has another significant advantage worthy of discussion- it can be used to determine what concepts one needs to learn to know each skill given one's current skillset, as all skills are composed of sub-concepts.

By subtracting sub-concepts of previously-learned skills that overlap those of the desired skill, a map of the exact additional concepts one needs to learn to acquire the new skill is made.

This creates roadmap from any skill to another, similar in function to a skill tree in computer games, but for real people, with real data, in the real world.

If applied correctly, this can potentially revolutionize learning program planning.

**The work of Mrs. Ayn Rand regarding human concept formation is one of the main pillars of the logical-epistemological basis of this work.**