

# Autonomy and Automation

# Introduction

## Increasing Automation

Extent: 47% of current jobs at risk of automation in the next 20 years.

Professions: white collar jobs, including legal and medical professionals.

Services: cognitive/affective labour, including retail and customer service.

Cause: developments in AI, including deep neural networks and other learning systems.

## Decreasing Autonomy

Hyper-Taylorism: Amazon warehouse workers, mechanical turk

Precarization: zero hours contracts, involuntary self-employment

Hyper-Employment: work/life collapse, constant connectivity

Bureaucratization: bullshit jobs, perverse incentives

How should we think about the connection between these?

*When are we exactly?*

Post-What?

# Historical Background

## Economics

Structure: Capitalism

Key Features: Modes of Production / Modes of Consumption

Core Concepts: Work / Leisure

## Politics

Structure: Modernity

Key Features: Modes of Governance / Modes of Subjectivation

Core Concepts: Self / Other

# Phases of Capitalism

## Feudalism

### Agrarian Capitalism (16th-17th C)

Mercantilism: Competition, Trade Wars

### Industrial Capitalism (18th-19th C)

1st Wave: Factories, Steam Power, Canals

2nd Wave: Machine Produced Machines, Telegraphs, Railways, Steamships

### Financial Capitalism (20th-21st C)

#### **Monopoly Capitalism**

3rd Wave: Heavy Industry, Electrical Engineering, Telephones, Mass Production

#### **State Capitalism**

4th Wave: Factory Automation, Transistors, Synthetic Materials, Automatic Calculation

#### **Neoliberal Capitalism**

5th Wave? - Networks, Mobile Communications, Information Goods

### Post-Capitalism? Platform Capitalism? What?

# Phases of Modernity

Pre-Modern

Early Modern (16th-18th C)

Renaissance: Humanism

Birth of the Novel: Authorship

Reformation: Personal Religion

Enlightenment: Empirical Science

Late Modern (18th-20th C)

Industrialisation: Urbanisation

Formalisation of Literature: Readership

Separation of Art and Craft: Creative Ownership

Rise of the Bourgeoisie: Leisure

Post-Modern? High Modern? What?

# Economic Analysis

## Feudalism

Mode of Production: Caste / Agriculture

Mode of Consumption: Hierarchy / Tithe

## Capitalism

Mode of Production: Division of (Wage) Labour / Automation of Non-Cognitive Labour

Mode of Consumption: Creation of Markets (Internal/External) / Creation of Consumers

## Post-Capitalism?

Mode of Production: Division of Labourers / Automation of Cognitive Labour

Mode of Consumption: Creation of Addicts (Pleasure/Compulsion) / Cognitive Capture

# Political Analysis

## Pre-Modern

Mode of Governance: Sovereign Society / Brute Domination

Mode of Subjectivation: Supra-Personal - family, church, community

## Modern

Mode of Governance: Disciplinary Society / Biopolitics

Mode of Subjectivation: Personal - education, employment, creativity

## Post-Modern?

Mode of Governance: Control Society / Administration

Mode of Subjectivation: Sub-Personal - tasks, talents, drives



# What is Freedom?

## Free Will and Determinism

Compatibilism: Epistemic, Normative

Incompatibilism: Libertarianism, Fatalism

What's wrong with it?

Spinoza: Free will presupposes determinism (self-causation).

What's right about it?

Kant: There is a difference between being autonomous and automatic (self-legislation).

## Types of Freedom

Positive / Negative

Qualitative / Quantitative

Cognitive / Non-Cognitive

# What's Happening Now?

## Cognitive Economics

There's a pivot between *supply side efficiency* and *demand side inefficiency*.

This has fed into synergetic forms of *cognitive oppression*.

This has incentivised *cognitive capture*.

## Cognitive Fragmentation

We're fragmenting *ourselves*.

We're building fragments of *artificial minds*.

We're combining the two in more or less *ad hoc* ways.

# Fragmentation

# Automatic Animals

How do we build up from systems that we treat *as if* they are rational agents, to systems that *genuinely are* rational agents?

## Analogical Bootstrapping

Functional Explanation: analogy with *practical reason* (success/failure)

Representational Explanation: analogy with *theoretical reason* (accuracy/inaccuracy)

## Cybernetic Systems

Parochial Problems: root systems optimise growth (water, nutrients); CNNs perform specific image recognition tasks (classification, generation); mice navigate local environments (food, reproduction, evasion)

Drives: behavioural modulation (input/output), interacting impulses (conflict/convergence)

Simulations: information integration, global availability, *predictive processing*

Control: self-modelling, choice selection, *hierarchical control*

Constraints: computational resources, information filtration

# Post-Postmodern Malaise

## The Poststructuralist Legacy

Good: the critique of subjective unity

Bad: the valorisation of subjective multiplicity

Ugly: the allergy to the notion of control

## Control vs. Agency

Action: events, responsibility, reasons

Rationalisation: tracking commitments, articulating drives, ownership

Volition: post-hoc rationalisation, reasoning, *akrasia*

## Personal Autonomy

Self-Causation / Self-Legislation

Normativity / Normality

Force / Content

*What is a Self?*

# Integration

# From AI to AGI

What does the G stand for?

## AGI vs. AI

Intelligence: problem solving

Generality: unrestricted (non-parochial)

Artificial: functionally decomposable

## Types of Generality

Absolute / Relative (*human-like*)

Qualitative / Quantitative (*human-level*)

Abstract / Concrete (*human-tractable*)

## Strategies of Abstraction

Good / Bad (implementation)

Above / Below (balance)

# What is AGI? (I)

What is abstract generality?

## Non-Deterministic Computation

Creativity: obviating brute force

Learning: asking better questions

## Bad Approaches to Abstraction

Skepticism: mysterian retreat

Singularitarianism: frictionless self-improvement

## Good Approaches to Abstraction

Generalised Aesthetics: systems that can be *cognitively stimulated*

Generalised Pedagogy: systems that can be *interactively taught*



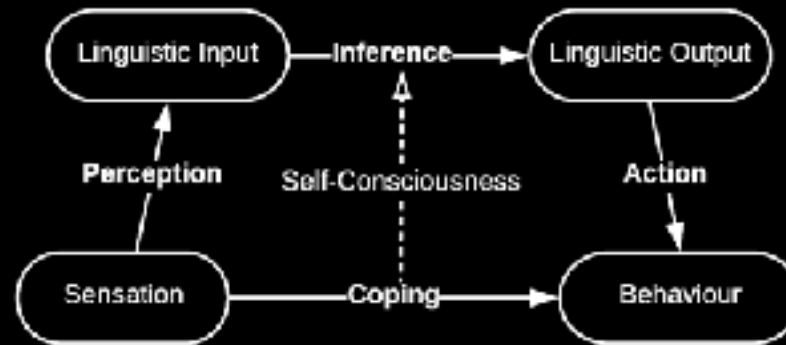
# What is AGI? (II)

What enables generality?

## Arbitrary Extensibility

The Frame Problem: deduction, irrelevance, parochial frames

Linguistic Re-framing: meaning as functional role (perception / inference / action), extended capacities, open relevance



## Unrestricted Representation

Commitment Integration: completeness, consistency

Conceptual Revision: theory change, referential continuity

# AGI as Transcendental Psychology (I)

Kant's Principle of Consciousness (KPC): there is no consciousness without the *possibility* of self-consciousness.

## **Imagination** (parochial processing)

Environmental Simulation: global integration of multi-modal sensory information

Object Individuation: extraction of local invariants

Object Simulation: anticipation of local variations

## **Understanding** (general framing)

Generic Judgment: classification of local invariants

Recognitive Judgment: re-identification of local invariants across contexts

Predicative Judgment: identification and classification of local variations

## **Reason** (general re-framing)

Ampliative Inference ( $\rightarrow$ ): extraction of judgment consequences

Critical Inference ( $\leftarrow$ ): identification of judgment conflicts

World Simulation: global integration of conceptually formatted information

# AGI as Transcendental Psychology (II)

The Understanding is our *type system*.

## Artificial German Idealism

**Imagination**

Kant → ↑

**Understanding**

Hegel → ↑

**Reason**

## Integrated Objects

Perceptual Invariants: criteria of identity contained *within* types

Interactive Loci: criteria of identity mapped *between* types

# Autonomous Machines

## Drive vs. Desire

Satiation vs. Satisfaction: representation, transmissibility

Unrestricted Aims: innovative means, original ends

Underdetermined Aims: preference elaboration, libidinal outsourcing

## Automatons vs. Autonomous Agents

Practical Commitment Integration: agents without autonomy, ultimate priorities

Theoretical-Practical Duality: revisability, selfhood, ends-in-themselves

# Modernist Morals

Do we know what we want?

## Re-Designing Persons

Consciousness and Self-Consciousness

Expression and Impression

Autonomy as Self-Control