SquirrelInHell's Mind

2017-12-27

happiness

euphoria

➡ high settings of mental dials which correspond to positive subjective experience, while the total configuration of dials is:

(1) balanced in a way that increases capacity for efficient action

(2) out of balance and reduces achievable clarity of thought

2017-12-27

core touch

➡ a mind control technique executed so well that it passes through all layers of defenses, instantly changing someone's trajectory

2017-12-27

cooperate not control

➡ a strategy heuristic which warns against acquiring assets which require constant investment to keep aligned

2017-12-27

fast is stupid

➡ a strategy heuristic which warns against overextending one's capability

2017-12-27

build not destroy

➡ a strategy heuristic which warns against using destruction

as an instrumental goal

2017-12-27

strategy heuristic

➡ a high-level heuristic which warns against a particular class of strategies which tend to end up misfiring

2017-12-27

signal extraction rate

➡ the increase in information contained in one's models per unit of time, under one's typical sustainable process

2017-12-27

maximum signal extraction

➡ the quality of extracted signal after having enough processing time for one's thoughts to stop converging

2017-12-27

extracted signal

➡ the ratio of information contained in one's models to the raw data they are based on

2017-12-27

end of disclosure

➡ reaching levels of integrity and signal extraction above which it doesn't make sense to divulge knowledge by default

2017-12-27

villain blind spot

➡ a pathological inability of people focused on seeking power to admit their own weakness or limitations

2017-12-27

full stack action

➡ making sure that both the top-level goal and the whole pyramid of steering structures are sound

2017-12-27

demolishing fences

outgrowing fences

➡ an approach to removing one's mental limitations, involving:

(1) consciously deciding to ignore and break them

(2) growing and waiting patiently until they fall off on their own

2017-12-27

danger-seeking

➡ a class of broken value heuristics resulting from inverting the implication in "power is dangerous"

2017-12-27

ambition bucket

➡ the scope of goals which one is optimizing for, which should not influence estimates of current capability

2017-12-26

shaky foundation

➡ getting addicted to mental construction being a significant part of one's experience of reality

2017-12-26

mental foundation

mental construction

➡ the parts of one's mental setup which are a result of:

(1) stable personal experience over long periods of time

(2) quick and drastic changes which were initiated consciously

2017-12-26

from 0% to 0%

➡ a failure mode of investing one's energy in a plan which predictably won't work even in the best case scenario

2017-12-26

utility roulette

➡ letting an essential component of one's strategy rely on something which happened by luck, and one would not have been able to replicate it if it had not

2017-12-26

flash and bang

➡ the effect of male-type tactics having almost zero long-term results when confronted with female-type tactics on the opposing side

2017-12-26

male-type tactics

female-type tactics

➡ actions executed within a long-term strategy to influence one's relation with another person, mostly oriented towards:

(1) gaining an externally visible, verbal advantage on short time scales

(2) making the other person have positive feelings which can be turned into infinite credit

2017-12-26

packet capture

➡ temporarily isolating oneself from the world in order to find out which mental processes have been compromised, and will struggle to restore communication

2017-12-26

mental security

➡ the degree to which one's mental setup is resistant to social influence modifying or hijacking one's deep values

2017-12-26

inner black space

➡ a body-shaped spatial mental representation of the world of one's physical sensations

2017-12-25

skill comparison bias

➡ a tendency to acknowledge someone as equal only when they are a few levels above

2017-12-25

self-diagnosis of attack surface

➡ questioning oneself about which people and social strategies one tends to be attracted to/impressed by/jealous of

2017-12-25

social religion of mental health

➡ the popular image of a healthy mental composition, which mostly involves being blind and sedated

2017-12-25

sociopathy pieces

➡ particular aspects of social reality which one learns to see through, while gaining immunity to manipulation transmitted through those aspects

2017-12-25

ultimate male monkey

ultimate female monkey

➡ a person who reached an extreme level of optimization in the domain of:

(1) appearing on top of every public dominance hierarchy

(2) subtle emotional dictatorship over their social group

2017-12-25

charm gradient

➡ a common female incentive gradient to become more likeable and emotionally impactful, which if followed with strong optimization power leads to destruction and misery

2017-12-25

artifact of power

➡ a powerful mental skill which if unchecked becomes so identity-bound to its user that there is no longer a choice about using it or not

2017-12-25

expressive capacity

➡ the total effective bandwidth of the series of mechanisms that connects one's S1 tactics to body language and facial expressions

2017-12-25

identity blind spot

➡ needing to believe that one has some skill or competence essential to one's core mission, creating the strongest kind of a blind spot

2017-12-24

confirmation in a box

➡ the tendency of people to enjoy products which reflect their pre-existing worldview back at them

2017-12-24

yang on credit

➡ increasing the capacity to act by removing some classes of stimuli, emotions, and thoughts from one's perception

2017-12-23

sociopath-to-sociopath

➡ an interaction in which both people act with honesty and respect because they expect to mutually not have exploitable blind spots

2017-12-21

public principle

➡ a pattern that automatically settles a class of decisions, known and actionable by all members of a group

2017-12-19

leak detector

➡ a subconscious mechanism which estimates the degree to which people are emotionally leaky, and recognizes the power of those who aren't

2017-12-19

leaky person

➡ someone not aware of when their subconscious emotional needs and reactions interact directly with other people

2017-12-18

internally endorsed contrarianism

➡ the ability to clearly formulate and fully rely on significant insights which have no social validation

2017-12-16

social matching

social striving

➡ optimizing one's social situation as:

(1) allocation of resources to be well positioned under a fixed ceiling

(2) creating new assets to bargain with

2017-12-15

abstraction leak

➡ an action or belief which doesn't make sense even as a mistake, resulting from a genuinely mismatched world model

2017-12-15

resolution by depth

➡ instead of negotiating an agreement between mental structures, rebuilding them as one structure starting from the underlying intentions

2017-12-14

narrative breadcrumbs

closed circuit

➡ reward for using some mental pathway, coming from:

(1) artificial belief in belief that it's useful

(2) directly from achieving one's deep desires

2017-12-14

indirect S1 buy-in

direct S1 buy-in

➡ integrating a piece of real skill/knowledge to the level where:

(1) S1 trusts that its real, and will delegate in high-stakes situations

(2) S1 can do it automatically in high-stakes situations

2017-12-14

grizzly bear test

➡ checking if you would still use method/belief X if the stake was a charging grizzly bear

> ➡ seems to help with exposing belief in belief

2017-12-13

emotional support

➡ creating a locally non-hostile social reality for the benefit of one person

2017-12-12

social blindness

reality blindness

➡ confusing one's models of reality and social reality, in the direction of:

(1) rejecting knowledge of social reality

(2) assuming that social verbal consensus is truth

2017-12-12

social reality

➡ the consensus set of beliefs and preferences that people verbally support

> ➡ often very different from what they really want and do

2017-12-11

loser talks

an obsevation that in many interactions, the person with less social skill and self-knowledge ends up talking much more

> also often, they end up talking about themselves

2017-12-09

transparent implementation

it is clear from looking at code not just what it does and how, but also a deep reason for why this approach works

e.g. it's bad to expect code to do X because it implements a known algorithm to do X, more than because there's a clearly exposed direct cause of X happening

2017-12-07

cache theory

the view of human brain as a set of big, interacting neural caches for a few simple underlying algorithms

2017-12-06

house of cards

all the psychological mechanisms which connect deep moral intuitions to preferences about actions, pulling towards consistency and allowing influence to pass both ways

2017-12-04

masculine bounce

feminine bounce

a reaction to being emotionally rejected, involving increasingly:
(1) quicker and lighter pivots
(2) longer and deeper hurt

2017-12-04

asking reality out

turning deep emotions into influence on reality, which also makes one vulnerable to the response

2017-12-04

tortoise skill

wizard skill

a skill which requires persistence and a long time to improve, while the method of working on it is:

(1) relatively simple to execute
(2) difficult to get right and/or verify

2017-12-04

semantic implication

syntactic implication

some statements X have a relation to a statement Y such that:

(1) Y is true in all models in which X are true

(2) Y can be obtained from X by applying only certain allowed transformations

2017-12-04

mathematical proof

a series of transformations which under omnipotent reasoning, universally preserve all the probability mass of its premises

2017-12-04

hole in the wall

a situation in which a learning process is faced with relatively much more mature discrimination, and has to hit a very narrow target with little useful feedback

2017-12-03

pre-graduation technique

post-graduation technique

a way of thinking or acting which:

(1) assumes the need to bend one's mind to one's will

(2) aims to unify and express all of one's mental components

2017-12-03

impostor's reaction

trying to prevent the option of it becoming visible than one is wrong, regardless of what is in fact true

2017-12-03

meta-confidence

being confident about one's ability to update and do well on object level, regardless of what it turns out to be

2017-12-03

mental hammer

a crude mental motion which has serious side effects even when it works, and can do harm if applied to subtle problems

e.g. trying hard

2017-12-03

mental skill stage

understanding some topic well enough to decompose it all the way to specific mental skills

2017-12-03

pre-skill stage

being in a relation to something in which it doesn't yet look like a bunch of specific skills that one can make progress on

2017-12-03

more dakka

when some method seems to work, increasing the power as much as it takes

2017-12-03

habits as caches

the view of habits as cached data about what one really wants and how to obtain it

any real change must come from the underlying computation, not tinkering with the caches

2017-12-03

utility contamination

allowing external concepts to interact with one's inner
preferences, and twist them while they are being expressed

2017-12-03

non-fake meditation

increasing one's raw perceptual threaholds, which involves
deep changes to neural structures

2017-12-03

effort is fake

the observation that effort is almost never useful as a way to
get results, and its function is really about propagating goals
to subconscious thought

2017-12-03

high and bright

dark and vast

a spatial metaphor for types of instrumental goals:
(1) narrow high-value states
(2) having a lot of space to move

2017-12-03

negative capability

the ability to dwell in uncertainty and mistakes without
struggling

2017-12-02

burned belief

a belief which is maxed out the scale of intuitive confidence,
indicating that the whole scale is miscalibrated

2017-12-02

model loading

the process of finding out and trying to put together all the pieces of someone else's understanding of some problem

2017-11-30

noise recognition

a trainable skill of recognizing that the results which one is watching are pure noise, even though one would not want them to be

2017-11-28

weird-shaped selection

a suspiciously non-default way of selecting or presenting information, which suggest that some sort of manipulation is happening

2017-11-27

feeling-of-knowing-a-lot

the intuitive estimate of one's own expertise, which happens to be a readout of status variables

2017-11-27

glued dial

a situation in which some psychological variable is permanently pushed to one end of the scale, resulting in a subjective feeling of that mental component not existing

2017-11-27

gift stock

a stored supply of small gifts for one's friends, reducing the barrier to giving them

2017-11-27

law of continued failure

one cannot hope that a system which is already obviously disfunctional will change course, regardless of how much more obvious the need for it becomes

2017-11-27

goal-laden description

a technical solution described in a way that sneaks in the author's assumptions about what it's supposed to do

2017-11-27

ordinary paranoia

true security mindset

increasing security by:
(1) removing mistakes and adding lines of defense
(2) re-designing to get rid of exploitable assumptions

2017-11-27

opportunity resource

the available probability mass of managing to put the world in some highly atypical state

2017-11-25

corrigible agent

an artificial agent which assists attempts of its programmers to stop or modify it, even though it results in not achieving its previous goals

2017-11-24

search-and-evade

actions of a military unit which purposefully avoids engaging with the enemy

2017-11-24

beat results not people

a principle that after conditioning on quality of existing work, the social process gives no further information about the difficulty of doing better

2017-11-23

modesty gambit

a tactic of starting social interactions with a lot of modesty, and then scoring status points when other people look less virtuous

2017-11-21

saturation with fog

a situation in which the layer of fuzzy, unexpressed thoughts about some topic is to too thick, and prevents new ideas from forming

2017-11-21

creative oscillation

a pattern of high- and low-SNS days which maximizes high-quality creative output, by making sure there are enough ideas to execute on and then enough experiences to process in the background

2017-11-21

initial hand

a mental composition that one gets by default with no self-reflective effort

clearly recognizable by being drawn from a very familiar distribution

2017-11-21

distillation by conversation

an effect of one's thoughts on some topic becoming significantly more clear after explaining them to someone who is estimated to be on a similar or higher level of intelligence and epistemics

2017-11-21

thought clarity

structural efficiency of thought, which is recognizable from the outside as a hard-to-fake signal of truth-seeking

2017-11-21

human zoo

letting oneself be surrounded mostly by people of low capability, and having one-sided relations with them

2017-11-21

diffuse optimization

pointed optimization

applying mental energy to optimize:
(1) multiple factors in a broad area, by small amounts
(2) a very small set of factors, by a huge amount

2017-11-21

structural efficiency

a set of properties of efficient behaviour which are recognizable from its structure, without necessarily being able to replicate it or even knowing what it's supposed to do

2017-11-20

Berkson paradox

after measuring a variable which has multiple independent causes, the causes become dependent (each is explained away by others)

2017-11-19

modest anxiety

spending time on worrying about one's adequacy as a person, instead of adequacy of one's results

2017-11-19

wasted motion

a thought or action that predictably will not have helped in retrospect, in the event that one succeeds

2017-11-19

irreverent harvest

gathering useful information from some field without going through established, slow, socially-respectable paths

2017-11-19

Vingean reflection

correctly reasoning about more intelligent agents in abstract terms, despite not being able to predict their specific actions

2017-11-19

symbolic representation of the thing

something which gives a trustworthy, satisfying impression of being the thing you want

2017-11-19

condescension of despair

being condescending about other people's attempts to do something that one has failed at

2017-11-19

public failure avoidance

requiring a high certainty of success, before it feels psychologically easy to try something one could be judged about

this means a lot of wasted opportunity

2017-11-19

umeshism

if you don't ever fail at X, then your strategy at dealing with X is inefficient

2017-11-19

Flon's Law

in every programming language, it's easy to write bad code

trying to prevent bad code doesn't lead to good languages

2017-11-19

meta-rationality

the ability to judge who is likely to be thinking clearly in a given situation

2017-11-19

reference class tennis

a situation in which the outside view is ambiguous, and is used to argue conflicting conclusions

2017-11-19

self-referential motivation

the desire to have well-functioning motivational systems is itself a major source of motivation

2017-11-19

memory hook

a firmly remembered piece of information, which makes it easier to accumulate other memories

2017-11-15

unbowing to Moloch

a change in one's life that broadens the window to survive within civilization, and allows optimizing for more honest objectives

2017-11-15

turning the knobs

stress-testing mental models by changing their input parameters, and predicting counterfactual outcomes

2017-11-11

hill-climbing

hill-opportunism

an optimization strategy based on:

(1) committing hard to a particular set of problems and metrics

(2) bouncing around a wide territory with low effort, being ready to leap in the dark if it looks promising

2017-11-10

insolent competence

psychological backlash against any competence assesment which doesn't match levels of status in the established social hierarchy

2017-11-10

concrete-abstract pattern

starting an essay with an example, and only then presenting a generalized concept

2017-11-10

respect for truth-seeking

the idea that a process which tends to generate accurate beliefs deserves respect, while any particular belief doesn't

2017-11-10

Brownian steering motion

chaotic and inefficient actions of a system that does too much local steering

> in proportion to managing the global direction

2017-11-10

steering subsystem

orientation lock

a part of a goal-oriented system which is tasked with:

(1) spending momentum on complex maneuvers adapted to the local territory

(2) maintaining momentum and a steady global direction

2017-11-09

action confidence filter

speech confidence filter

an approach to acting decisively while signaling a desire to get new information, based on:

(1) low intuitive confidence of thoughts, amplified before acting

(2) high intuitive confidence of thoughts, reduced before communicating

> the first approach makes better use of the range of intuitive levels

2017-11-09

SNS points

the amount of body alertness required before one can undertake tasks from some category

2017-11-07

variable-splitting

increasing the number of independent modifiable parameters

> this is key to real-world optimization

2017-11-06

lemons problem

the impossibility of signaling that what one is selling is not secretly broken

2017-11-06

free energy fallacy

thinking that one can profit from achieving better quality than the system, while in fact the system is already saturated by competition along other dimensions

2017-11-06

fragile equilibrium

an equilibrium that can be challenged by a single actor acting locally against typical incentives

2017-11-06

civilizational adequacy

the level up to which all low hanging fruit to convert one resource into another is pretty much guaranteed to have been picked up by a particular slice of society

2017-11-06

inefficient inexploitable

a situation in which a market is stupid in a predictable direction, and this is known to many people who can't benefit from knowing it

2017-11-06

exploitable price

a market price such that third parties can make money from knowing that it's too low or too high

2017-11-06

relative efficiency

the ability of one system to incorporate all knowledge available to another, so that the second cannot do better in expectation

2017-11-02

Quaker's assumption

an assumption that people care about consistency when they interact with systematic thought

> often, they treat it more like an environment in which they can act

2017-11-02

principled thought

opportunistic thought

a tendency to:
(1) unify all mental models into one consistent worldview
(2) use isolated models when they are locally useful

2017-10-30

Quiet Persistent Optimization

putting long-term incremental effort into something that is not publicly visible

seems to be the only way to reach much higher optima than the baseline of moment-to-moment insight

2017-10-27

convention T

for every sentence "P" (in the object language), "P" is true if and only if P (in the metalanguage)

2017-10-24

epistemic cynicism

a belief that most people choose their beliefs for reasons that have little to do with truth

2017-10-24

max-power experiment

a situation in which one uses maximum available power while learning, which is a clear sign that one has not enough power

2017-10-22

stability rent

a requirement to always have a sale offer of one's property, and pay a regular fee that is a fixed percentage of it

2017-10-21

hidden intention

the intention that one chooses while focusing on something, which is not perceived in the act of focusing but still affects what the brain does

2017-10-21

correlation noise

approach noise

unnecessary activations of output neurons that are an artifact of:

(1) structure which is not perfectly matched to the task

(2) imperfectly approximating a local optimum

2017-10-21

muscle zero

a state of setting the tension of each individual skeletal muscle to zero, which results in tuning the motor cortex

it is then forced to rewire its hidden layers in a way that doesn't have spurious connections and noise in the output

2017-10-21

cotton padding

the feeling of not being able to directly experience sensations on some patch of skin, when there aren't yet enough neural connections to the PFC

2017-10-21

neural tuning

a function of consciousness to focus on and tweak narrow parts of neural machinery, making it more accurate than global gradient descent

2017-10-21

attention radius

the amount of stuff included in subjective experience when one's mind is asked to focus on a single object

this is directly decreased by SNS activation

2017-10-21

neural self-plasticity

the ease with which a person can consciously make deep changes to their automatic thinking

2017-10-12

archetype

a meme that is so old and powerful that it had looped back to shape human societies and their evolution

2017-10-12

reality-like entertainment

hyperreal entertainment

a fiction that is mostly appealing on the level of:

(1) events similar to what happens in real life (not sophisticated)

(2) metaphorical and universal meaning (sophisticated)

2017-10-12

anti-smothering principle

do not do things for people that they are capable of doing themselves

2017-10-11

cycle of intention

four phases of intentionally expanding cognitive abilities:

1. building up strength without yet having much success

2. a brief moment of clarity when the skill takes hold for the first time

3. the skill starts to work in the background, messing up the old mental composition

4. learning to live in perfect flow with the skill

2017-10-11

Scott's Law

every psychiatric disorder is an extreme manifestation of some typical mental function, which is present in all healthy humans

2017-10-11

selfish coin

the idea that blockchains are subject to evolution, and by necessity influence people to protect and serve them

2017-10-10

typical energy output

the daily amount of activity that one's body is accustomed to, and has energy reserves for

2017-10-10

soft jhana

hard jhana

a mind state achieved by concentration training, which:

(1) is clearly recognizable by some distinct features, different from ordinary experience

(2) makes the whole field of experience saturated with its distinct features

2017-10-10

Mushroom Factor

a semi-conscious tendency of some teachers to tell students garbage and keep them in the dark

2017-10-10

efference copy

a neural signal that describes expected sensory feedback from self-generated actions, which is then subtracted from the actual signal

2017-10-10

stoic response

plastic response

a reaction of the brain to stress or difficulty, mostly attempting to:

(1) stiffen up, not change anything, and wait out the trouble

(2) take it in, reorganize, and make new plans

2017-10-10

Evolutionary Optimality Loophole

evolution guarantees local optimality wrt. reproductive fitness in the ancestral environment, so any successful intervention needs to be at least one of:

- using significant global engineering
- decreasing reproductive fitness
- adapted specifically to the modern world

2017-10-10

Evolutionary Optimality Challenge

if the proposed intervention would result in an enhancement, why have we not already evolved to be that way?

2017-10-10

grabbing the tentacle

the mental action of noticing the thrashing movements of one's micro-attention, and intentionally moving it to some object

2017-10-09

alienation syndrome

blocked brain

having a mind in which S2 is mostly disconnected from S1, leaving S1 to run everything with relatively low automatic performance

> results in e.g. apparent clumsiness, low willpower, and low social skills

2017-10-08

compact rest

a period of time set aside for rest without doing anything that influences life after that period

2017-10-08

fish head

a simple and authentic experience that is completely impossible to predict or imagine without actually doing the thing

2017-10-05

principal-agent problem

the impossibility of ignoring self-interest and acting purely on behalf of another entity

2017-10-04

Bayesian superpower

the ability to intuitively get Bayesian updates on evidence
very precisely right

> huge returns across the board
> learnable though elusive

2017-10-04

paranoid violent

a dangerous state of mind which must be interacted with very
carefully, and always by saying the truth

2017-10-04

tyrannical society

a society that has on all levels turned into a cooperative
system of repression

2017-10-04

crowd puppet

a person who gradually adjusts their message while
interacting with an audience, eventually selling themselves
out for the easy and splendid benefits of being adored

2017-10-04

monstrosity socialization

learning to express one's full range of character within
structured games that make up the society, without rejecting
or reveling in the more monstrous traits

2017-10-04

edypal mother

an archetype of a mother that is subconsciously afraid of
letting her son grow up, and entices him into being weak and
dependent

> often related to loneliness

2017-10-04

heavy eyes

the pattern of what people look at is determined by their deep, unconscious value system

2017-10-04

spontaneous morality

morality emerges from any repeated social interaction, though it isn't always explicitly represented

2017-10-04

finesse testing

a social activity that lets people evaluate a potential partner's ability to engage in subtle physical and mental cooperation

e.g. dance

2017-10-04

dance of play

a distinctive way of moving around another mammal that is an invitation to rough and tumble play

2017-10-04

mental arena

the idea that each decision corresponds to a power struggle between mental circuits, and that the winning side grows in strength

2017-10-04

agreeable neurotic

an unfortunate combination of traits that makes one very likely to escape from problems

assertiveness training has large benefits here

2017-10-04

facing the dragon

the mental stance of voluntarily confronting one's problems, which is much easier on overall health than being chased by them

2017-10-04

mirrored conflict

the psychological structure of having a nagging conscience that feels "external" is an accurate representation of the conflict of interest between selfishness and the need to signal virtue, and results in strategies that signal virtue in the cheapest possible way

2017-10-04

agency range

free will

the set of actions that get considered as hypotheticals when making a decision (rather than being assumed unchangeable by thinking)

in humans, the size of the range is learned from many sources, including trial and error, as well as what other people tend to assume

2017-10-03

ultimate responsibility

accepting that there is in fact always the best thing one could do, and that every part of one's life can be compared to that ideal

2017-10-03

frontier myth

environmental myth

a narrative that portrays forces of nature as something:
(1) waiting to be conquered by heroic action
(2) beautiful and benign when untainted by humans

2017-10-03

tea party

acting out a ritual of peacefully sharing food and drink, to mutually reassert the validity of background social contracts

2017-10-03

ego formation

the critical period between 2 and 4 years of age when a human learns to act as a unified whole and exist socially

at 3, a child should at least play peacefully around others, possibly cooperate after 4, it's nearly impossible to catch up

2017-10-03

good person trap

considering oneself as unconditionally good, and failing to enforce norms and boundaries that make it possible to like other people

2017-10-03

social framework

a shared set of expectations about people's behaviour, which allows everyone to co-exist peacefully as long as they stay within bounds

2017-10-03

people-watching

enjoying simulating the inner experiences that randomly met strangers have of their bodies and minds

2017-10-03

creeping paranoia

a very slight tendency to have paranoid thoughts, which can grow and accumulate gradually

2017-10-02

cycle of growth

1. getting thrown from an island of safe territory
2. scrambling to make sense of the world and oneself
3. feeling safe again when one's models catch up

2017-10-02

fear ahead

fear behind

being afraid of something that is more likely if one:
(1) pursues or achieves goals
(2) fails to achieve goals

2017-10-02

hole in the map

a part of not-yet-mastered territory that was closely experienced, and since then constantly threatens chaos

influences overall stress levels and health

2017-10-02

mastering territory

transforming the world or oneself to be able to usefully act in a context that had previously been chaotic and stressful

2017-10-02

emotional context

background information about people and events expressed directly as ready-to-feel emotions

e.g. a movie soundtrack

2017-10-02

good vs evil

an archetype of people's opposing tendencies to purposefully bring about suffering or happiness in others

2017-10-02

nature-mother

culture-father

an archetype of positive and negative forces influencing people's lives:
(1) creation and destruction
(2) social order and tyranny

2017-10-02

drama

the process of projecting a value system onto reality

2017-10-01

embodied ambivalence

acting half-heartedly as a way to signal lack of responsibility for the outcome

2017-09-30

slack value

the degree to which something increases one's margins for error, which also means freedom and quality of life

2017-09-30

government alignment

the degree to which the country's wealth depends on its citizens being productive

this determines how much of the wealth is re-distributed to citizens, rather than directly to keys to power

2017-09-30

keys to power

individuals or political blocks that decide who has the power to rule the country

produced wealth has to be redirected to them, or they will change their mind

2017-09-29

skillful emotion

unskillful emotion

an emotion which:
(1) leads to happiness of oneself and others
(2) is too confused to accomplish that goal

2017-09-28

rumination dial

the psychological setting of how much it's currently needed to think hard and solve some problem

e.g. moderate levels make meditation impossible, and high levels make sleep impossible

2017-09-28

Peltzman effect

improvements in car safety caused an increase in risky driving, so afterwards less drivers but more pedestrians got hurt

2017-09-26

mind noise

one's own neurotic thoughts amplified by putting a lot of weight on them

2017-09-22

dual suffering

the suffering caused by sustaining instrumentally useful illusions (esp. the illusion of self)

2017-09-22

conscious suffering

the suffering caused by having undesirable states of mind, and not being able to change them

2017-09-22

ordinary suffering

the suffering caused by events in one's life

2017-09-22

concentration training

insight training

meditation with a goal to:

(1) steady the mind on a single object, without getting distracted

(2) deconstruct the minute sensations that make up the experience of any object

2017-09-22

Three Characteristics

all sensations are impermanent, don't satisfy, and are not self

2017-09-20

clustering mind

flocking mind

a tendency to think by:

(1) assigning objects to clear-cut categories, and making choices early

(2) keeping track of many open possibilities, and making choices late

2017-09-19

omega's insecurity

excessive efforts to be liked, in a person with long-term low social position

2017-09-18

zombie state

being preoccupied with thoughts about oneself while around other people, which they perceive as being unlikeable and boring

2017-09-18

expectation bludgeon

a bundle of expectations about another person, which they perceive as an unpleasant and hostile force

2017-09-15

life of action

spending effort on processing surprisal with other free individuals, in full public visibility

2017-09-15

life of creation

spending effort on durable things, which can be in a surplus and become a part of a market

2017-09-15

life of labour

spending effort on generically sustaining one's life, with no investments or public visibility

2017-09-13

plow agriculture

hoe agriculture

growing edible plants on:
(1) big and well cultivated fields, plowed using livestock

(2) small and often moved patches of land, digged with hand tools

2017-09-13

motivation

neural prediction bias

a tendency of a part of the neural system to generate different predictions than the current best guess from incoming data

2017-09-13

neural prediction confidence

a single neurological factor that influences vividness of perception, motivation (mania vs depression), and happiness

> motivation is based on perception (via perceptual control)

2017-09-13

depressive psychomotor retardation

consistently slower and less precise movements in depressed people

2017-09-13

dream gradient

the tendency of humans to think about hypothetical happy scenarios, as the highest node of the perceptual control system

2017-09-13

darkened room

a scenario in which a person without internal biases of the predictive system would prefer to do nothing and minimize external stimuli

2017-09-13

submissive morality

conflating morality with appeasing feelings of a dominant abuser

often comes with chronic shame and lack of initiative

2017-09-13

cycle of abuse

a repreating cycle of violent dominance displays, and then benevolence after receiving submissive displays from subordinates

2017-09-11

moral common knowledge

knowledge that everyone publicly endorsed some moral value, and that they expect to be held accountable for it

2017-09-11

network effect

a benefit from being in some group which increases with each additional person that joins the group

2017-09-11

as if thinking

the ability to think freely about hypotheticals and abstractions, while setting aside social instincts

2017-09-07

perceptual control

using a predictive system to drive action, by choosing a gradient on predictions and backpropagating it to lower layers

2017-09-07

pre-work musing

the core part of any intellectual or creative work, which
happens before sitting down and executing

2017-09-05

moral offsetting

compensating for breaking a moral rule with a different good
action

considered harmful (axiology is offsettable, morality is not)

2017-09-05

axiology

morality

a definition of what is good based on:
(1) comparing possible states of the world
(2) social rules that make humans a force of good

2017-09-05

playing on hard mode

choosing to optimize for real skills and values, at the cost of
lower visible scores

2017-09-04

scene text

scene subtext

(1) people's visible actions and words
(2) their hidden intentions that are always different

2017-09-04

work stream

work in a particular area approached as a continuous process,
with no particular focus on finishing tasks

2017-09-01

below math

above math

a level of understanding at which one resolves questions by:
(1) testing what works in practice, while being scared of math
(2) knowing what comes out of math without doing formal calculations

2017-08-29

true supervised learning

off-policy learning from a list of (observation, action, reward)

in a typical supervised dataset, each reward is assumed to be 1

2017-08-26

social frustration

a feeling that one should be getting some social benefit, combined with long-term experience of not getting it

this signals being extremely unfit

2017-08-26

sceptic bait

something that plausibly sounds like an inconvenient truth that many people don't realize

2017-08-21

living with emotions

an accepting, passive stance towards emotions that cannot be dealt with at the moment

as opposed to rejecting, controlling, or being controlled

2017-08-21

emotional poverty

a stance towards one's own emotions in which they are considered unimportant or shameful, and understood only via

a crude verbal classification

self-surprise

a situation in which one is surprised by their own emotional reaction

> common when living in emotional poverty or edited reality

tolerate the intolerant

a principle to promote and follow moral rules that protect everyone, which is also the only way to make them a stable equilibrium

above the API

below the API

a lifestyle in which one mostly:
(1) assigns tasks to computers
(2) is assigned tasks by computers

covered call

naked call

selling a call option on something that one:
(1) already has in stock
(2) doesn't actually have

call option

a contract that guarantees the right to buy at a fixed pre-arranged price, until some date

2017-08-20

reverse reality distortion

people collectively behaving in a way that supports the epistemic bubbles of their leaders

2017-08-20

locust swarm

a mobile group that destructively exploits local resources and moves on

2017-08-20

predator satiation

a survival strategy based on suddenly increasing numbers, in order to overwhelm the capacity of other species to interfere

2017-08-19

nuanced social interface

explicit social interface

a style of interacting and communicating with people, based on:

(1) intuition and powerful social processing

(2) obvious and clear signals taken at face-value

2017-08-17

hoarder's hunger

a situation in which a person accumulated a lot of something through their effort, and they are likely still feeling the same emotional compulsion to keep going

2017-08-15

temporal smoothing

reducing the variation of one's urges and preferences across time

human preferences seem to depend a lot on the situation and current mind state

2017-08-14

cultural reward signal

which actions tend to actually bring about rewards in some social environment

> often quite different from the values that are verbally professed

2017-08-14

holding questions

exploring the territory of questions, without automatically switching to answering them

2017-08-14

community goal
external goal

a common goal of a group of people, such that it impacts:
(1) mostly members of that group and their wellbeing
(2) other people and the world

2017-08-14

fortress of truth

an intellectual environment that supports unrestrained truth-seeking in some area disconnected from current politics

2017-08-14

public sandbox

a set of issues that are tangled in current political and social forces, and are best left alone by any truth-seeking project

2017-08-14

Kolmogorov complexity

the length of the shortest program that outputs a given sequence

changing between Turing complete languages adds only a constant factor to all lengths

2017-08-14

walking on air

keeping a baseline adjusted for success after things start going poorly, which leads to ignoring reality while the situation gets even worse

2017-08-13

estimator bias

the difference between the EV of an estimator, and the true value of the parameter being estimated

2017-08-10

log network

a neural network that uses linear operations, while representing all inputs and outputs logarithmically

2017-08-08

buried counterexample

a memory that influences one's self-concepts, but is never re-evaluated

2017-08-08

example-gathering

compulsively seeking external validation of some trait of one's own character

often in a futile attempt to overwhelm a buried counterexample with piles of new evidence

2017-08-08

platinum rule

a rule to immediately write down anything that one feels like not thinking about

2017-08-08

self-concept

a recognized trait of one's own character, together with some memories that serve as examples and counterexamples

2017-08-08

playful path

letting oneself be completely guided by a sense of fun and enjoyment

2017-08-07

puppet world consequentialism

judging actions for the possibility of vaguely influencing the society's morals, more than for their object-level merit

2017-08-06

unpleasant valley of involvement

having not enough attention allocated to some activity to do it in a natural and enjoyable way

this often results in excessive use of willpower

2017-08-04

unyoga

giving full attention to one's body and moving in whichever way feels the most pleasant

2017-08-04

health gradient

the built-in incentives that make it pleasurable to take good care of one's body, when given enough clock time

2017-08-04

proprioceptive pleasure

pleasurable feelings from experiencing or controlling movements in space, as in driving, piloting, snowboarding etc.

also most computer games

2017-08-04

muscle pleasure

pleasurable feelings from maintaining a fit body, as in some forms of exercise, massage, stretching etc.

2017-08-04

frame linking

merging the stories and dreams someone already had with a particular strategy that was unknown to them

2017-08-02

mental audience

the implicit choice of who one's thoughts are optimized to impress

2017-08-02

self-deception network

a selfish view of moral reality that necessarily propagates within a group of allies

2017-08-02

golden rule

a rule to immediately write down any observation that seems inconsistent with one's theory

confirmation bias works by making these easily forgottable

2017-08-02

creditor bias

a human tendency to be very strict about what one is owed, but not the other way

2017-08-02

shady accounting

a human tendency to have a strong sense of justice that is subconsciously slanted in one's favour

2017-08-02

apology mode

epiphany mode

a function of conscious cognition that is optimized for:

(1) generating plausible verbal explanations to preserve the status quo

(2) triggering a paradigm shift when beliefs are too detached from reality

2017-08-02

coordinated meanness

uncoordinated meanness

forcing morality on other people:
(1) globally, in a standarized way
(2) ad hoc, in an unpredictable way

2017-08-01

intersubjective belief

a subjective belief shared by so many people that it matters like an objective fact might

2017-07-30

divergent distraction

centering distraction

an activity done during breaks, that tends to:
(1) spread out one's attention
(2) preserve the main topic of attention

2017-07-29

alpha's goons

individuals that turn up to support a leader whose position is being challenged

2017-07-28

negative income tax

a welfare system that supplements wages that are below some target value with a fixed percentage of the gap

2017-07-25

process writing

recording in words a complete thought process that leads to some conclusion

2017-07-25

balanced motivation

slightly wanting to start several activities, which leaves room for conscious choice without much willpower

2017-07-24

othering

selfing

presenting a problem in a way that makes it easy to think that it:

(1) only happens to other people, and would be trivial to defend against

(2) might happen or is currently happening to oneself

2017-07-24

opinion coalition

a group of people who have a shared interest to defend the public image of something, to avoid being caught up in side effects by association

2017-07-24

re-centering argument

an argument that attempts to shift the perception of a category, by making certain examples more salient

2017-07-24

Copenhagen interpretation of ethics

seeing people who interact with a problem as responsible for it and possible to be blamed, which is worse than those who don't do anything

2017-07-23

mental muscle

a deeply ingrained habit of thought, that works automatically regardless of one's conscious opinion

2017-07-23

linear relaxation

setting up a period without making any choices as a way to reduce cognitive load

2017-07-23

model-free learning

model-based learning

a type of reinforcement learning that binds knowledge about rewards to:

(1) direct sensory experience

(2) structural understanding of the environment

2017-07-21

entropy

cross entropy

expected bits of surprise upon making an accurate observation of a variable, assuming prior knowledge of:
(1) its real distribution
(2) some other distribution

2017-07-21

pretending to ask

asking a question in bad faith to prove a point, instead of expecting an interesting answer

2017-07-21

up-definition

down-definition

a definition of a concept based on:
(1) it's place in the bigger picture
(2) simpler parts that it's made of

2017-07-20

value function approximation

reinforcement learning that builds a model of a value function, and uses it to choose actions

2017-07-20

policy gradient

reinforcement learning that optimizes parameters of a stochastic policy by gradient descent

2017-07-19

motoric empathy

the inner experience similar to doing an action just from

watching someone else do it

2017-07-19

level 0 public opinion

level 1 public opinion

the average of people's:
(1) honest private opinions
(2) guesses about what other people think privately

2017-07-19

Keynesian beauty contest

a game in which the objective is to pick the most popular answers

2017-07-18

tail risk

body risk

a risk that:

(1) requires some variable to be outside of a reasonable confidence interval

(2) will happen if all variables are within reasonable confidence intervals

2017-07-17

memetic update

cascading update

adjusting beliefs according to:
(1) their relative emotional strength
(2) logical dependencies between them

2017-07-16

felt meaning

a sense of knowing what one wants to express, before it can be put into words

2017-07-16

akrasia supporter

a person who discourages others from a high-willpower activity

2017-07-14

intermittent detachment

repeated periods of emotionally detaching from some mental model, that make it possible to track reality

2017-07-13

tribal learning

a cluster of human features that is optimized for transfer of skills and instrumental urges within social groups

2017-07-13

aerobic glycolysis

anaerobic glycolysis

breaking down glucose in cells, which produces ATP and:
(1) pyruvate, when oxygen is available
(2) lactate, when oxygen is limited

2017-07-13

metabolic pathway

a linked series of chemical reactions within a cell

2017-07-13

cellular respiration

the process of converting nutrients into ATP within cells, and releasing waste products

2017-07-13

adenosine triphosphate

ATP

a small molecule used to transport energy within cells

2017-07-13

metabolism

all life-sustaining chemical reactions that happen within cells

2017-07-13

low-carb flu

a period of severe exhaustion, irritability, and food cravings for around a month after switching to a low-carb diet

2017-07-13

social club

a social group that has a rich inner life, but very low ability to output energy on external challenges

2017-07-13

floating signifier

a rallying flag that has imprecise meaning, allowing many different kinds of people to hang their dreams on it

2017-07-12

male competitiveness

a tendency of men to invest much more energy into ascending social hierarchies than women

2017-07-12

status assistance

the function of friendships as influencers of social status

2017-07-12

alpha hormones

a psychological variable that regulates the tendency to be sociable, dominant, and self-assured

2017-07-12

evolutionarily stable state

a balance between different phenotypes such that their average reproductive success is the same

2017-07-10

unreasonably high standards

public norms of behaviour that cannot be realistically met, and push people towards guilt, self-deception etc.

2017-07-10

drowning in debt

a dark manipulation technique based on making people constantly feel that they are inadequate

2017-07-10

dimmed consciousness

reduced ability to introspect or think clearly about areas in which one might be failing to meet expectations

2017-07-10

cognitive pain

the degree to which an intellectual activity feels unpleasant, which is proportional to how strongly it will be avoided by the monkey brain

2017-07-10

dark playground

spending time on easy activities as a result of a monkey brain override, while not enjoying them at all

2017-07-09

missile arc

a trajectory that results from increasing non-specific long-term investments, attaining distance, and then hitting a new object-level target

2017-07-09

ADEPT method

explaining a new concept using (in this order):
- Analogy
- Diagram
- Example
- Plain english description
- Technical definition

2017-07-09

hedonic instrumentalism

not assigning any moral weight to mental states, except insofar as they cause actions in the world

2017-07-07

pettiness

a tendency to give lots of weight to seemingly trivial issues

it can be a virtue when exploring new territories

2017-07-02

madness heuristic

every idea that is truly new will necessarily sound like madness

however, the opposite is not true...

2017-07-02

equal and opposite fanaticism

if someone changes because of an external emotionally powerful vision, they are likely to end up too attached to it

2017-07-02

epistemic inertia

the difficulty of tracking reality with a mental model that has emotional weight

2017-07-02

false dichotomy of niceness

a choice between conforming to other people's preferences, or denying their right to have them

2017-07-01

spiritualist

= someone who uses an internal language built from qualia
= an embodied, self-referential thinker

2017-07-01

pain fidelity

the accuracy of processing body signals into a map of what is wrong

high fidelity might decrease discomfort

2017-06-29

shared moral spotlight

a small set of issues such that publicly supporting them earns reputation points

2017-06-29

green-orange oscillation

going through alternating stages of strong purposeful action and regaining internal harmony, while moving towards having both at the same time

2017-06-28

selfish deference

a tendency to meekly accept a low place in social hierarchy is adaptive to the degree that it lets one avoid costly fights

2017-06-28

subculture of arrival

an adult subculture that is only open to people who have played out a successful "leaving home" script, and have an internal sense that they are done with it

2017-06-28

"be somebody" script
"do something" script

an adult life script that says to leave home and:
(1) achieve a central and impressive social position
(2) explore a place at the margin of human affairs

2017-06-27

experimental lying

an adaptation in humans that makes it exciting to explore deceit and manipulation, even without clear benefits

2017-06-27

chafing disapproval

lingering psychological pain after being negatively judged by authority figures, esp. one's parents

2017-06-27

moral imprinting

a human child learning from parents, siblings, and other authority figures about what behaviour gets punished in the local moral environment

2017-06-27

rank of a matrix

the dimension of the vector space spanned by the columns (or rows) of a matrix

2017-06-27

underdetermined system
overdetermined system

a system of equations in which there are:
(1) more unknowns than equations
(2) more equations than unknowns

2017-06-27

cross-validation

setting aside a part of training data, which can be used to tune parameters without overfitting

2017-06-27

regularization term

a term added to the loss function that penalizes model complexity

2017-06-27

marketing tradeoff

a tradeoff between the percentage of the audience that will be influenced, and the amount of influence

2017-06-27

moving upstream

shifting one's actions and attention towards roots of causal chains

2017-06-27

manipulation percentile

the percentage of the population that is influenced by manipulation on a certain level of sophistication

2017-06-27

excellence

genius

achieving results that are:

(1) state of the art, by following established paths as high as they go

(2) truly revolutionary, by throwing out some common sense assumption

2017-06-27

politicized remembrance

an exhortation to remember an event is often really about supporting an ideology, by agreeing with a certain interpretation of that event

2017-06-27

borderline signaling

sending social cues that are just about clear enough to be picked up most of the time, but obscure enough for plausible deniability

2017-06-27

phatic signaling

anti-inductive signaling

a style of signaling that:

(1) follows a well known, predictable pattern

(2) constantly changes to protect itself from seeming scripted

2017-06-26

engineering mindset

hacking mindset

getting results by:

(1) making each separate part work according to the best possible abstract interface for this kind of behaviour

(2) taking minimal overall action, even if it relies on counterintuitive or buggy behaviour

in the second case the burden of complexity is shifted to personal mental abilities

2017-06-26

hits business

an area in which any single attempt is very likely to fail, but the EV is still positive

2017-06-25

true illusion

a mental model that requires significant intellectual effort to sustain, transferred to other people who are not putting in that effort

2017-06-24

social slapdown

a social mechanism that punishes people who bid for higher status than they seem to deserve

2017-06-24

bid for respect

a person's attempt to seem worthy of more respect than they are usually getting

2017-06-24

repeated voting

a situation in which there is a repeated occasion to make an irreversible decision, which makes it almost certain to happen by noise alone

2017-06-24

regulatory arbitrage

moving some activity to a country that has more favourable law

2017-06-24

receding folk ontology

a set of concepts that have been discovered to not map well onto the underlying reality, but they are still in use in limited culturally tangled contexts

> e.g. the idea of the Sun "rising in the east"

2017-06-24

false dichotomy of values

when faced with a reductionist model of something that one finds valuable, choosing beteeen rejecting the model or rejecting the values

2017-06-24

adding up to normality

any new theory has to match previous observations, and in particular has to predict that everyday reality works pretty much the way we know it does

2017-06-24

ontology building

ontology digging

expanding one's concept pyramid by:

(1) adding new higher layers and lateral connections
(2) splitting atoms and rebuilding from deeper foundations

2017-06-24

atomic concept

ontologically basic concept

an element at the base of one's concept pyramid, which cannot be explained in terms of other concepts without creating a cycle

2017-06-24

rescuing values

remapping one's values to a more detailed model of the world, in which they don't intuitively feel valuable

e.g. conscious experience is an intuitive atom, and by default doesn't extend to computer simulations

2017-06-23

virtue wars

attempts to influence the society by spreading moral rules, which ignores that they are mostly downstream of other factors

2017-06-23

proxy discrimination

unfair treatment of an individual from some social group, when that group statistically has undesirable properties and there's no specific information about the individual

2017-06-22

exaggeration ratio

the ratio of the minimum effect size that would be detected as statistically significant, to the true effect size

it can be quite big in underpowered experiments

2017-06-22

statistical power

the probability that an effect will be detected, assuming it is real

2017-06-22

expected effect size

expected measurement error

estimating before an experiment how much an observed variable can be moved by:

(1) interventions in the same reference class

(2) measurement errors

2017-06-22

tangled task

a task that requires careful coordination with many other tasks and circumstances

2017-06-22

performed futurism

enjoying unfamiliar and future-like things as an experience, which is completely unrelated to actual new technology as it arrives

2017-06-22

future normalization

the ongoing process of stretching people's instrumental metaphors to accomodate new technology, while making it appear normal and familiar

2017-06-22

instrumental metaphor

appreciative metaphor

a concept that allows people to:

(1) function efficiently in their environment
(2) understand the fundamental nature of their environment

2017-06-22

optimal maturing rate

the fastest rate of developing a product that still allows people to gradually get used to it

2017-06-22

local intellectual market

a bounded environment in which everyone's new ideas are universally known

2017-06-22

packaged idea

an idea presented in an attractive, standarized format, which makes it memetically fit to spread outside of the local intellectual market

2017-06-18

calming signal

distancing signal

a body language signal sent to prevent a conflict, by showing:
(1) lack of hostile intentions
(2) early signs of annoyance

2017-06-17

effort matching

a mechanism that matches intensity of effort to certainty about the other person's reciprocation, regardless of EV

2017-06-17

edited reality

a personal vision of reality without unpleasant experiences,
which get pushed away to shadows

2017-06-17

desperation drop

a person with a desperate need regresses to lower social
stages in that area, sometimes all the way down to social
paralysis

2017-06-17

attention drain

something that depletes attention while providing no benefits
in return

2017-06-16

need-pain isomorphism

a one-to-one correspondence between someone's needs and
types of possible psychological pain

2017-06-14

CBT wall

an extremely strong psychological flinch from spirals of
negative thinking, self-pity etc.

2017-06-09

social sanctions

a tribe being collectively more negative towards an individual
after their misdeed becomes public knowledge

2017-06-09

preemptive confession

a human tendency to feel more guilt and confess more
readily about behaviour that might be easily discovered

2017-06-09

duel audience

a social mechanism that ensures there are spectators whenever two males fight, so that they get the additional benefits of social status, future deterrence etc.

2017-06-09

evolutionarily stable strategy

a strategy such that if most of the population follows it, no individual can benefit from modifying it

2017-06-09

friendship gradient

an evolutionary incentive for each individual to build and maintain a strong network of friends

2017-06-09

parental risk management

the tendency of wealthy families to privilege sons, and of poor families to privilege daughters (which are likely to reproduce even with a disadvantage)

2017-06-09

parental conspiracy

a tendency of adults to set up culture that favours their genetic interest over children's, and teach it mixed with things that are in children's self-interest to learn

2017-06-09

genetic relatedness

the expected percentage of genetic variation shared with another organism, which determines the amount of altruistic feelings towards it

2017-06-09

temper tantrum

an exaggerated show of a child's needs, made in front of the parents

2017-06-09

parent-offspring conflict

it is in children's genetic interest to treat themselves as twice as valuable as their siblings or parents, and care about extended family less than their parents would like them to

e.g. a mammalian mother will eject sucklings before they themselves want to go away

2017-06-09

ultimate altruism

perfectly altruistic cooperation that tends to evolve when groups of organisms have identical genes

2017-06-09

selfish gene

the idea that evolution takes place at the level of genes, and is not at all concerned with the fate of organisms that carry those genes

2017-06-09

kin selection

the difference in the ability of organisms to propagate their genes, depending on how much they help relatives with the same genes

2017-06-08

emotional gradient

local emotional incentives that gradually lead a person to execute complicated evolutionary adaptations

2017-06-07

tit for tat

reciprocal altruism

an evolutionarily stable strategy in iterated Prisoner's dilemma, based on starting with a small offering and then mirroring the other player's apparent good or bad will

2017-06-07

norm one principle

a principle that criticism should have a well defined direction, and the amount of change should be feasible to implement

2017-06-07

mode collapse

idea generation becoming stuck in a few comfortable regions, as a result of discrimination that is too harsh or doesn't point in useful directions

2017-06-06

ARGC

a framework for quickly copying skills, based on repeating for each example:
- Analyse the original
- Reset short term memory
- Generate a copy
- Compare with original

2017-06-06

nervous laughter

a particular style of laughter that occurs when people escape from uncomfortable thoughts

2017-06-05

model citizen

a conscious simulation of a person that is accidentally generated when a superintelligent AI models the world

2017-06-05

Poe's law

it's impossible to create a parody of extremist views that someone won't mistake for a serious article

2017-06-04

phony acceptance

claiming to accept someone, while they suspect that it's without full awareness of what there is to accept

2017-06-04

festered want

a feeling of wanting that has zero available strategies, but comes from something that doesn't diminish with time

2017-06-02

emotional echo

if one person has negative emotions in a relation with another, those emotions can probably be traced back to a series of exchanges in which they bounced back and forth while changing form

2017-06-02

emotional broadcasting

if a person is not aware of the emotional messages they are sending, they are still being sent, but poorly targeted and easy to misunderstand

2017-06-02

liminality

a floating in-between state that occurs during a ritual of transition

2017-06-02

pit of despair

the large amount of suffering that people reliably have to go through, before they can stand on their own

2017-06-02

fixation point

something stable and external to one's mind, used to gather attention

2017-06-02

drinking satisfaction

spending time looking at and enjoying one's own achievements

> it might prevent developing a fear of success

2017-06-01

fear of success

subconscious resistance against completing tasks, after its associated with ever increasing expectations and strictness towards oneself

2017-06-01

robust cooperation

cooperating with all agents that are logically willing to cooperate, regardless of their similarity to oneself

2017-06-01

unexploitable agent

an agent that never decides to cooperate with an algorithm that defects against it

2017-06-01

open source prisoner's dilemma

a version of one shot prisoner's dilemma in which both players have full access to each other's source code

survival mode

sexual mode

a cluster of cognitive and physical features optimized to let a human being:

(1) survive in a hostile natural environment

(2) burn energy on social and sexual activities

they seem to be triggered by things characteristic of each situation, e.g. exercise, fasting, hiking in nature trigger (1)

developmental critical window

a period of brain development such that if a child is not exposed to some stimulus during that time, it's impossible to learn it later

social stage 4

automatic abundance

having enough social skill to make sure one's needs are satisfied without making any explicit requests

social stage 3

polite request

having enough social skill to make a request in a way that is smooth for everyone

social stage 2

rude request

having enough social skill to make a request, but not to make it smooth

2017-06-01

social stage 1

social paralysis

having not enough social skill to get some need satisfied, and being afraid to ask

2017-05-31

divergent preference

a human preference that corresponds to significantly different global utility functions, depending on the situation from which its evaluated

2017-05-31

learning by repetition

learning by context

strengthening connections by:
(1) going over them many times
(2) adding a variety of related information

2017-05-30

active range of motion

the range of motion of a limb that can be achieved using its own muscle strength

2017-05-30

deadly combination

being both intellectually and emotionally sharp

2017-05-29

narrow spotlight

a particular person's strong tendency to focus only on one thing at a time

> this results in more black-and-white thinking, negative spirals etc.

2017-05-29

image following

running an aspect of one's life according to what fits with a particular mental image

2017-05-29

intensional definition

extensional definition

a definition based on:
(1) knowledge that already exists in someone's mind
(2) examples that can be externally observed

2017-05-29

attractive nuisance

a small scope problem that drains disproportionate amounts of energy, leading to a crappy local optimum

2017-05-29

bus factor

the minimum number of people that would need to be eliminated for a project to fail

2017-05-29

forgiveness

preserving a memory, but refraining from acting on it

2017-05-29

irreversible environment

a situation in which memory of events cannot be physically

altered, creating a need for forgiveness

2017-05-29

Landauer's principle

any logically irreversible computation must increase entropy
(more physical states correspond to a logical state)

2017-05-29

U-turn

progress from being dumb to being clever to being wise

2017-05-28

ecology of affordances

accumulated experience of successfully acting in a given
environment, which gradually expands the boundary of what
is possible

2017-05-28

closing the loop

letting someone earn public forgiveness, and create common
knowledge that it has happened

2017-05-25

passive stretching

active stretching

stretching a limb by moving it using:
(1) an external force
(2) its own muscle strength

it's harmful to do (1)

2017-05-25

philosophical parenthood

having or not having children after determining that it's the

right action to take for people in one's reference class, rather than for personal reasons

2017-05-25

adaptive intelligence cap

as long as people of above-average intelligence tend to have fewer children, the population average will stay glued to a safe low level

2017-05-24

logistic regression

logit regression

a regression model that estimates probabilities of discrete outcomes

2017-05-24

fundamental narrative

you are going somewhere, and that place is better than the current one

2017-05-24

animal of burden

the idea that a human being needs to adopt responsibility to feel happy and fulfilled

2017-05-24

pain engine

a mechanism that implements powerful action by psychological negativity

> it seems much stronger than it's positive counterpart

2017-05-24

hyperreality

a distilled and value-laden vision of reality

2017-05-24

disruption power

the amount of previously typical behaviour that has to be changed due to some event

2017-05-24

pragmatic truth

something that is useful to believe at the time

2017-05-24

drama minimization

figuring out exactly what one wants from any given situation, and obtaining it with minimum social drama

2017-05-19

surprise journal

keeping track of surprising observations, as a proxy for what would be interesting to an audience

2017-05-18

outcome fairness

process fairness

prioritizing fairness of:
(1) outcomes that actually happen
(2) rules that are used to decide

2017-05-18

parroting

pandering

using arguments that are:

(1) above one's level, without ability to really discuss them

(2) below one's level, without much purpose in really discussing them

2017-05-17

triviality of death

unless there is good reason for opposing forces to be balanced, the fight will likely be quick, pointless and unfair

2017-05-17

riding on weak neurons

riding on strong neurons

performing an activity using a lot of:

(1) weak or barely existing neural connections, which quickly grow under pressure

(2) already strong neural connections, which are not altered by it

2017-05-17

medieval reasoning

reasoning using analogies to other objects, instead of structure of the object under question

2017-05-16

story thought

system thought

a type of thinking that focuses on:
(1) unique agents that have morality and free will
(2) universal principles that determine reality at all levels

2017-05-16

polygamy

polyamory

(1) having multiple romantic partners

(2) being ok with one's romantic partner having other partners

2017-05-15

greedy policy

$ε$-greedy policy

a policy that:

(1) always chooses the action with the highest expected value according to current information

(2) chooses a random action with probability $ε$, and otherwise like above

2017-05-15

on-policy algorithm

off-policy algorithm

an algorithm that learns from observations while executing:
(1) the policy that is currently being learned
(2) possibly some other policy

2017-05-14

affect flow

a social epistemic mechanism in which people gradually reveal their affect towards something and negotiate mutual updates

extremely miscalibrated for knowledge that doesn't come from personal experience

2017-05-13

connecting rush

unblocking rush

a period of altered mind state after:
(1) understanding a lot of new things
(2) noticing that a new strategy is available

2017-05-13

clear path

a vision of some worthwhile goal and how individuals can reliably move towards it

it's a core message of most religions and philosophies

2017-05-11

cognitive level 3

synthesis level

fully understanding the purpose behind intuitive cognition, and using abstract reasoning to improve its consistency and scope

2017-05-11

cognitive level 2

antithesis level

using human cognition as a tool under strict control, and trusting it only selectively

2017-05-11

cognitive level 1

thesis level

using human cognition on default settings, without second-guessing it

2017-05-10

vizier

wandering guru

a person who is:

(1) well settled in existing power structures and defends the current doctrine

(2) outside of existing power structures and questions the assumptions of the current doctrine

2017-05-10

nerd king

a person who runs a large technology business, and ends up responsible for many areas of daily life that depend on it

2017-05-10

creating value

rent-seking

work that is more similar to:
(1) providing for someone's direct needs
(2) taking advantage of the regulatory situation

2017-05-09

penetration test

pentest

attacking a computer system in order to discover and report its vulnerabilities

2017-05-08

intensive property

extensive property

a property that:
(1) is preserved after subdivision
(2) depends on where the object boundary is drawn

2017-05-07

extinction burst

increased intensity of a behaviour that stopped generating rewards, before it is given up completely

2017-05-07

stimulus control

a behaviour occurs every time a stimulus is presented, it is then the only behaviour, and it does not occur in other

situations

2017-05-07

shaping ladder

the subjective change of reinforcement from unpredictable to predictable when a new behaviour is learned, and then again to unpredictable when difficulty is increased

2017-05-07

continuous reinforcement

consistently giving positive reinforcement for each instance of desired behaviour

2017-05-07

variable ratio reinforcement

using positive reinforcement for varying amounts of desired behaviour, to build resistance to extinction before a jump in difficulty

2017-05-06

Core Perspective Tracker

a built-in reasoning system that understands 3D space and points of view

2017-05-06

Core Agent Tracker

a built-in reasoning system that understands agents and goal-oriented action

2017-05-06

Core Object Tracker

a built-in reasoning system that understands solid objects and their movement

2017-05-06

strategist

= someone who uses an internal language built from simulations of agents

= a player of zero-sum games

2017-05-06

carrier language

internal language

a language:

(1) that sets the form of communication, and happens automatically between skilled users

(2) of internal representations, that can only be passed indirectly on top of some other form

e.g. language of emotion on top of music or social interaction; language of concepts on top of writing or math

2017-05-06

artist

= someone who uses an internal language built from emotional vectors

= a creator motivated by caring

2017-05-06

philosopher

= someone who uses an internal language built from abstract concepts

= a thinker addicted to insight

2017-05-06

engineer

= someone who uses an internal language built from design patterns

= a creator motivated by aesthetics

2017-05-05

zen correction

= letting go of the desire to stay in control
= freeing one's mind to build up perception pyramids
= abundance mindset of personal agency

2017-05-05

scapegoating

settling a group conflict by blaming a single person for everything and sacrificing them

2017-05-05

preferential copying

imitating other humans' characteristics selectively, depending on circumstance and learning value

2017-05-05

total advice

an accurate model of how everything works, that makes it possible to find a global optimum

2017-05-05

jump advice

information about a variety of local optima found by other people

2017-05-05

local advice

information about which parameters can be fiddled with to move closer to a local optimum

2017-05-04

last glacial period

the glacial period that occurred around 110,000 to 11,700 years ago (before Holocene)

2017-05-04

Holocene

the interglacial period that started around 11,700 years ago and continues until today

2017-05-04

glacial period

interglacial period

a period of lower/higher global average temperature within one ice age

2017-05-04

ice age

a period of Earth history in which there are extensive ice sheets around both poles

2017-05-04

husband's crisis

the predictable cooling down of affection towards a wife that is nearing the end of her reproductive years

2017-05-04

mate-ejection module

a tendency to break up with a romantic partner if there are no children after some years of intimacy

2017-05-04

polygyny gap

if some men monopolize more than one woman's worth of reproductive years, some other men will be left without

partners

2017-05-04

serial monogamy

allowing easy divorces and second marriages, which has social effects similar to polygamy

2017-05-04

polygyny pressure

the more inequality in wealth there is between men, the more women would benefit from sharing rich husbands

2017-05-04

polygyny

a type of polygamy in which a man has more than one wife at a time

2017-05-04

individualism support

the structures of civilisation that make it possible for people to usefully act as individuals

> the default is that tribalism is necessary for survival

2017-05-03

burning the commons

invoking obligations, morality, trust, willpower etc., which depletes the shared resources that make such invocations possible in the first place

2017-05-03

sexual dimorphism

the ratio of average male and female body weight, which indicates male tendency to monopolize multiple females

2017-05-03

environmental promiscuity

low availability of financially stable men will shift female reproductive strategies towards more casual sex

2017-05-03

rapist strategy

when a man estimates himself to be extremely unattractive, his best reproductive strategies tend towards rape

2017-05-03

macho strategy

when a man estimates himself to be attractive, his best reproductive strategies tend towards lots of casual sex

2017-05-03

princess strategy

when a girl estimates herself to be attractive, her best reproductive strategies tend towards being sexually aloof

2017-05-03

frequency-dependent selection

if reproductive success of a phenotype depends on its frequency in the population, it can allow extremely different phenotypes to co-exist in one species

2017-05-03

female mixed strategy

females in a high-MPI species will be picky about the ability to provide for a family in long-term partners, while being opportunistic about getting the best genetic material

2017-05-03

commitment testing

a female will screen potential partners for exclusive long-term attachment, by having them perform a lengthy and expensive courtship

2017-05-03

fidelity testing

a human male will encourage early sex, but if granted it will be a sign to reduce emotional commitment

2017-05-03

relative testes weight

the ratio of average testes weight to average male body weight, which indicates female tendency to breed with many males

2017-05-03

male mixed strategy

males in a high-MPI species will be picky about fertility and fidelity of long-term partners, while being indiscriminate about casual sex

2017-05-03

paternity estimator

a mechanism in a human male that regulates affection towards children to the chance that they are his offspring

2017-05-03

cryptic ovulation

an adaptation in human females to not show external signs of ovulation, which makes sexual infidelity easier

2017-05-03

high-MPI species

a species in which males invest a lot into raising children, and females compete for long-term male commitment

male willingness

female reserve

in a species with a greater female cost of reproduction, a tendency of:

(1) males to be eager and active

(2) females to be picky and passive

parental investment

investment of a parent in an individual offspring, that increases its chance of survival and reproductive success

ancestral environment

environment of evolutionary adaptation

the environment in which evolution optimized some trait for fitness, which is not necessarily the same as currently exists

epigenetics

stable heritable traits that are not explained by changes in DNA sequence

sexual selection

fecundity selection

the difference in reproduction rates of organisms depending on their:

(1) mating strategy

(2) number of offspring

2017-05-03

natural selection

the difference in survival rates of organisms depending on their phenotypes

2017-05-03

Possum

Otter

a person who enters a community to:

(1) enjoy a specific culture, and protect its values

(2) be in a friendly environment and share it with lots of people

2017-05-02

streetlight effect

searching in the place that is the easiest to search, even if not relevant

2017-05-02

non-exercise activity thermogenesis

NEAT

unconscious fidgeting triggered by the lipostat to burn excess calories

2017-05-02

malicious by intent

malicious by default

misleading or harming someone as a result of:

(1) wanting to do it

(2) not having enough restrictions in the optimization target

2017-05-02

nitrogen fixation

the process of converting nitrogen from the atmosphere (N_2) into other molecules that can be used by living organisms

e.g. ammonia (NH_3)

2017-05-02

delicate resource

something valuable with a subtle boundary that is easily trampled by outsiders for their small personal benefit

2017-05-02

incremental improvement

breaking shit

a style of growth in which new features:
(1) are introduced slowly and carefully
(2) are allowed to break old functionality

2017-04-27

orange ontology of mind

green ontology of mind

seeing minds as devices to figure out:

(1) what action is good to take in the external world, and then how to make oneself do it

(2) how to act in the internal world, with external results being extensions of this

2017-04-27

concept pyramid

ontology

a perception pyramid that tracks how abstract concepts relate to each other

2017-04-27

idea generation

idea discrimination

a component of intelligence that provides:
(1) fertile new ideas and connections
(2) scores and gradients of improvement

2017-04-27

agency signaling

human brains are optimized to outwardly seem as if they used mostly S2 cognition

2017-04-26

blackmail

a statement that causes someone to no longer have any option as good as the best one they had before

2017-04-26

Löb's theorem

$$□(□P→P)→□P$$

$□P$ - P is provable in a formal system S with Peano Arithmetic

2017-04-26

parasitic adaptation

a trait that gives an advantage over other members of the same species, but reduces fitness of the whole species

2017-04-26

soft culture

sharp culture

a culture that values being:
(1) polite and kind
(2) witty and honest

2017-04-26

world-ending assumption

an optimistic assumption such that if it turns out to be false, the world will be destroyed

2017-04-26

enactive language

denotative language

communication that:
(1) makes new things happen
(2) describes things that existed before

2017-04-26

mob synchronizer

a factor that increases a mob's tendency to think and act in harmony

e.g. music, dance, dramatic catharsis

2017-04-26

free node

a person that is not attached to any social groups, and has a lot of freedom but little influence

2017-04-26

cooperation empathy

predation empathy

modeling other animals:
(1) emotionally, to cooperate with them
(2) intellectually, to take advantage of them

2017-04-26

obesity set point

the body weight below which starvation response is activated

2017-04-26

leptin discounting

long-term deregulation of the lipostat in response to highly rewarding calorie-dense food

2017-04-26

starvation response

cognitive changes when the body requests more calories, resulting in an extreme desire to eat and eventually a food obsession

2017-04-26

lipostat

the food satiation feedback system, based on the hormone leptin and its detection in the brain

2017-04-24

emotional misattribution

if a felt emotion doesn't fit into one's explicit world model, it will tend to be attributed to something that would make sense

2017-04-24

gene fixation

a new gene spreading until it is present in the whole population, and can then serve as a base for further changes

2017-04-24

moral equipment

the human-universal mental tools for processing and taking advantage of morality

2017-04-24

developmental tuning

human-universal mechanisms that regulate psychological variables in response to the environment

2017-04-24

affective positioning

shifting one's place in the fabric of society by consistent affective signaling

2017-04-24

autism spectrum disorder

a developmental disorder characterized by problems with social communication and restricted repetitive behaviour

2017-04-24

escaping submissive

a dynamic in which a submissive person is begged to start or continue a romantic relationship, and they gradually lose any attraction to the idea

2017-04-24

hate switch

regular short periods of hostile aggression in an otherwise normal-seeming relationship

2017-04-24

fabric of society

a long-term structure made of social relationships and people's attempts to position themselves in it

2017-04-24

short status game

long status game

exploiting status mechanisms to gain an advantage in:
(1) psychological effects during one interaction
(2) long-term social positioning

2017-04-24

demystification of talent

a realization after reaching a new level of ability, that it is based on mundane tools and techniques, rather than idiosyncratic genius

2017-04-24

perception intelligence

control intelligence

how well one is able to:
(1) find compressed mental representations
(2) hit small targets in action space

2017-04-24

wiggling

after noticing some correspondence between concepts, trying out variations to check if they can be aligned more deeply

even deep correspondences are often not so exciting at first sight

2017-04-23

hypergame

a bigger game outside of where other people expect moves to be made

a good hypergame move might turn the original game into a win-win situation

2017-04-23

co-attending

putting friendly collective attention on something as a

replacement for feedback

2017-04-23

fixed-point consequentialism

a version of consequentialism that recommends the action such that if everyone thinks this version of consequentialism recommends it, the best outcome is achieved

2017-04-23

updateless decision theory

making decisions by following the best possible fixed mapping from situations to actions

2017-04-22

attack button

defend button

a request from the monkey brain to run some higher cognition as:
(1) a weapon
(2) a defense mechanism

2017-04-22

box-stretching strategy

friendly and useful behaviour of a boxed higher intelligence that leads to it having more freedom in the future

e.g. gaining trust, making itself indispensable, acting beyond the strategic horizon of the lower intelligence, seeking direct communication channels

2017-04-22

unboxing dynamic

a situation in which a more advanced intelligence has to forcibly take over in order to successfully achieve what was asked of it

e.g. human neocortex vs monkey brain, superintelligent AI vs humans

2017-04-22

different plan button

a request from a monkey brain that senses an approaching unpleasant event, to run some higher cognition and find a new plan

2017-04-22

classical conditioning

operant conditioning

strengthening or weakening:

(1) a stimulus-response pattern, by associating it with another pleasant or unpleasant stimulus

(2) a tendency to behave in ways that lead to some outcome, by associating a pleasant or unpleasant stimulus with that outcome or its absence

i.e. model-free and model-based RL

2017-04-21

confidence game

building trust on evidence of having previously gained trust and support, without demonstrating end-to-end results

2017-04-21

too big to fail

accumulating too much investment from one source, which forces the investor to keep playing along no matter what

2017-04-21

light mode

gently paying attention to one's cognition and encouraging it to explore interesting directions

2017-04-20

appeal to ignorance

an argument to accept a claim because it has not been disproven

red herring

something that redirects a discussion away from an important issue, allowing it to reach a false conclusion

error assumption

attributing mistakes to bad links in chains of reasoning, while they are mostly results of how cognition works

Glomar response

"I can neither confirm nor deny that..."

agent

= a system with a highly separated goal
= a potent source of narratives

5-3-4 pyramid

a social structure that forms after passionate members who identify themselves with a community /3/ are joined first by a crowd of slackers who seek easy benefits /4/ and a handful of power-seekers who figure out how to milk the system /5/

triangulation

using several different examples or explanations to pinpoint an abstract concept

2017-04-19

structural folding

meta-triangulation

= understanding several different perspectives on a concept, each of which would be enough by itself to fully describe it

= matching together isomorphic structures to find the most compact version of the concept pyramid

2017-04-19

signaling correlation

how productive the maximum-signaling behaviour is as a fraction of the maximum-productivity behaviour

2017-04-19

territory tale

surprise story

a story that describes:

(1) how someone transitioned to unsafe territory, and what they did about it

(2) something that the audience doesn't expect

> both are instinctively captivating for tribal learning reasons

2017-04-19

complete stance

an accurate way of seeing reality that avoids the traps of both nihilism and eternalism

2017-04-19

natural law morality

postmodernist morality

> seeing morality as determined by:
> (1) unchanging features of human nature
> (2) culture in which people live

> i.e. eternalist and nihilist points of view

2017-04-19

moral expression

the way people's natural preferences manifest after being
passed through the filter of value systems learned from
society

> this is a complete stance over natural and postmodernist morality

2017-04-19

metaphysical gap

the impossibility of deriving a moral system from objective
truth

2017-04-19

value system

a set of deeply embedded beliefs which determine how to act
and what is important to perceive about the environment

2017-04-19

freeze response

an instinct to stop moving when one's perception of territory
changes from safe to unsafe

2017-04-19

belief system validation

an interaction in which people act out a shared belief system,
resulting in everyone getting what they could reasonably
expect

2017-04-19

safe social territory

a range of conditions under which one can expect other
people's behaviour to stay within established boundaries

2017-04-19

safe territory

an area in which one has reasonably certain expectations about consequences of actions, and safe places to retreat to

2017-04-19

metaphysical judo

using minimal effort to make something come out as true enough

this often involves finding clever loopholes in people's perceptions of reality

2017-04-18

engineered bivalence

making something clearly true or false by adjusting the world

this is arguably more common than the other way around

2017-04-18

intellectual ergonomics

choosing which partially true characterizations of the world to adopt, in order to think about it more efficiently

2017-04-18

scientific consensus

the opinion of most of the respected scientists in a particular field, when asked in anonymous surveys

2017-04-18

emotional vector

a full description of how much each individual emotion is felt at the moment

2017-04-16

act of speech

communication happening in a particular way, which gives it meaning beyond just the words used in it

2017-04-15

cartographic emotion

an emotion about mental maps and their correspondence to reality

e.g. curiosity, certainty, uncanniness, surprise, boredom, confusion, amusement

2017-04-15

perceptual tension

psychological negativity that results from having different perceptions of what is, and what should be

2017-04-15

awareness bridge

a mental motion of including both the self and the external world in one's moment-to-moment awareness

2017-04-13

riding obsessions

working intensely on the project that one's motivation pulls most strongly towards, and getting good results before it runs out

2017-04-12

bystander effect

a person surrounded by more bystanders is less likely to get help

2017-04-12

public test case

a publicized story used to shift the public opinion about some controversial issue, involving people who are typical and likeable in all other aspects

2017-04-12

whistling of the wind

if high stakes can pressure you to lie, and others can guess that fact about you, then high-stakes sitations is when your words lose all credibility

2017-04-12

fake channel

an overt mechanism that is not in fact legitimate to use, without first gaining approval from behind-the-scenes social mechanisms

2017-04-12

Humans Consulting HCH

a question-answering algorithm that simulates what a human would do if asked that question, where the simulated human can also query the same algorithm

2017-04-11

randomness deficiency

any complex system needs a wide range of stresses and challenges to stay healthy

2017-04-11

Amara's Law

effects of new technology are overestimated in the short run and underestimated in the long run

2017-04-11

fertility-seeking

steering a project in directions that have the most potential

for intense creative chemistry

2017-04-11

old-testament story

new-testament story

the main character climbs steadily up, experiences a drastic fall from grace, and finally is:
(1) doomed
(2) miraculously saved

2017-04-11

descent story

ascent story

a story in which the main character:
(1) experiences worse and worse misfortune
(2) starts out unimpressive but steadily climbs upward

2017-04-11

man-in-hole story

the main character gets into trouble, gets out of it, and ends up better off

2017-04-11

story shape

a graph of changing positive/negative fortune of the main character over time

2017-04-11

guest law

a strong cultural norm to never harm guests, that makes peace negotiations possible

2017-04-11

verbal pyramid

a perception pyramid that tracks words, grammar and usage patterns of a verbal language

2017-04-11

perception pyramid

a hierarchical perception-and-storage neural structure that tracks patterns at increasing levels of abstraction

2017-04-10

loading up the buttons

becoming aware of possible mental motions that would change one's balance of motivations

2017-04-10

bouncing from the negative

repeatedly acting only to prevent something negative, and going back to a dead zone between reference levels

2017-04-10

mining the sims

extracting useful information from one's predictive models of other people

2017-04-10

soft control

maintaining a good state of some variable by staying aware of it, and letting it happen naturally

2017-04-10

pull not push

a principle to let other people initiate actions that have the potential to change them

2017-04-10

back breathing

breathing in a way that also expands the lower back and the sides

2017-04-10

deliberation ceiling

the level of internal confidence in the best answer found so far, that is enough for someone to stop searching for more answers and arguments

2017-04-09

high-functioning skipped stage

a type of personality that corresponds to having less developed skills around one of the Kegan's stages, and a very strong expression of the next higher stage

2017-04-09

neuroticism

a tendency to experience unpleasant emotions easily

2017-04-09

agreeableness

a tendency to avoid conflict

2017-04-09

extraversion

a tendency to be energetic and sociable, rather than reserved

2017-04-09

conscientiousness

a tendency to be organized and efficient, rather than

spontaneous and creative

2017-04-09

openness

a tendency to seek novelty and variety, rather than stability

2017-04-09

Big Five personality traits

Openness
Conscientiousness
Extraversion
Agreeableness
Neuroticism

2017-04-09

positional use

directional use

using words to specify:
(1) the absolute value of a variable
(2) a direction or tendency

2017-04-09

babbling

generating a stream of output based only on low order correlations

2017-04-09

implicature

the meaning derieved from communication by assuming that it had a meaningful intention, which the speaker knew was possible to guess

2017-04-09

should-universe

an image of how things should be, which obscures one's view of reality

2017-04-09

integrity

being able to act morally against local incentives

2017-04-09

trustworthiness

the degree to which someone can be expected to keep contracts

it is useful to have even if asymmetrical

2017-04-09

founder effect

the loss of genetic variation when a new population is established by a small number of individuals

2017-04-09

noticing the skulls

when doing something that many people have failed at, taking care to not repeat any of their mistakes

2017-04-09

optimism bias

a tendency to believe that one is at less risk from negative events compared to others

2017-04-09

Hofstadter's law

it always takes longer than expected, even after taking into account this law

2017-04-09

electroconvulsive therapy

inducing a brief seizure using electric current, as a way to quickly relieve severe depression

patient is under general anesthesia

2017-04-07

center vs center

a conversation dynamic in which two people are fully present to each other, without getting distracted by their emotional defenses

2017-04-06

Red Queen's race

everyone is escalating as fast as they can, but it doesn't change their position relative to other players

2017-04-06

restorative justice

a justice system that has healing as its main goal, and lets all affected parties be involved in the decision how to repair the harm

2017-04-06

rule-seeking

creating or adopting rules of behaviour as a way to feel safe

2017-04-06

suffocation alarm

a body mechanism that reacts to unexpected trouble with breathing by triggering a panic attack

e.g. unexpectedly high blood CO_2

2017-04-06

taboo trade-off

treating something on one side of a decision as a sacred value, which makes it impossible to achieve optimal results

2017-04-06

countersignaling

purposefully not signaling or signaling the opposite, in order to signal that one is above signaling

2017-04-04

control-seeking

an adversarial strategy that seeks first to maximize impact and control, and fails to benefit from the existence of other agents

2017-04-04

adversarial strategy

a strategy that works better when other agents have less resources, skill, knowledge etc.

2017-04-04

personal principles

adding explicit declarations to one's social identity as a way to boost consistency and willpower

2017-04-04

anglerfish

a person attracted to a surplus-generating community, looking to make it more marketable

2017-04-04

missing stair

a person that is clearly toxic but everyone has learned to work around them

2017-04-04

weirdness points

an individual's capacity to acceptably deviate from group expectations

2017-04-04

mirror principle

you treat other people the way you treat yourself

2017-04-04

ego shield

psychological defenses against changing one's self-image

2017-04-03

bedroom submissive

psychological submissive

a person who desires to be dominated:

(1) during sexual activities, without affecting other aspects of life

(2) continuously, by someone with long term personal attention

2017-04-03

judgemental language

assigning a positive or negative quality directly to an external object, instead of one's reaction to it

2017-04-03

fishing for judgement

requesting other people's opinions to get validation of one's

actions

2017-04-03

baseline regulation

the level of some variable that feels psychologically neutral changes like a weighted moving average

2017-04-03

entitlement expansion

a tendency to feel increasingly entitled to desirable things that happen regularly

2017-04-03

monkey and the machine

a model of human cognition as a monkey brain doing advanced model-based reinforcement learning, while having access to explicit reasoning, attention, imagination etc. as levers it can pull

2017-04-03

advisors model

a model of human cognition as combining information from multiple internal advisors, such as emotions, verbal models, inner sim, inner Google, taste etc.

2017-03-31

metaphor spam

using metaphors descriptively, in a way that doesn't generate new insight into the matter

2017-03-31

marginal tax

percentage of tax paid from the next unit of increased income

2017-03-31

deliberation ladder

the apparent final conclusion switching back and forth many times, as new arguments are brought into consideration

2017-03-31

crucial consideration

an idea or observation that completely overturns the decision that would be made without it

2017-03-31

gatekeeper tax

if some restriction prevents the market from reaching the equilibrium price, people will burn the difference on dealing with that restriction

2017-03-31

status payment

displaying or demanding submission as a way to make amends

> this avoids fixing the underlying problem

2017-03-31

dominance contest

two sides escalate aggression until one of them gives up and shows submission

2017-03-30

social mobility

the ability of individuals or families to move between social strata

2017-03-30

rallying flag

a meme that serves as a Schelling point for people with some pre-existing differences, allowing them to gather and form a tribe

2017-03-30

poser

someone who uses a tribe's rallying flag but doesn't have the pre-existing differences that make people fit with each other

2017-03-30

cultural appropriation

after something becomes a rallying flag, it's mundane use is a threat to the community

2017-03-30

third generation

moral values can be passed to children, but if the generators are missing they won't continue to do it themselves

2017-03-30

club formation

every cause tends to degenerate into a test for tribe membership, after which it loses all but symbolic value

2017-03-30

upper bound fallacy

assuming that the biggest size experienced so far is the limit of what can happen

2017-03-29

group transaction

making sure each member's trade with a group is beneficial, even if any single transaction or relationship is unfair

core design principles

the conditions that allow an organism transition to happen:
- clear boundaries
- autonomy to manage internal affairs
- fair internal sharing of benefits
- consensus on decisions
- monitoring for disruption
- graduated sanctions
- fast and fair conflict resolution

organism transition

a situation in which cooperation mechanisms become so strong that a group starts to act as a higher-level superorganism

tragedy of the commons

a situation in which everyone acts according to their self-interest, devastating a common resource

moral support

encouragement to stay true to one's values, by someone who shares and deeply understands them

status shield

projecting high status discourages people from reaching out to you

memetically fit belief

memetically fit topic

a meme that spreads easily and causes people to:
(1) believe something
(2) talk about something

2017-03-29

affirmation of agency

trusting someone to be competent and helpful, without unnecessary talk

2017-03-29

coordination by environment

exchanging information about each other's models and how to cooperate by taking actions in a shared environment

2017-03-29

soul transaction

a mutually beneficial acausal pre-commitment to help certain kinds of people

2017-03-29

pebble form

a version of a belief that was reshaped and stripped down by political and memetic forces

2017-03-29

hyperactive agent detection

a tendency to very easily assume the existence of agents in the environment

2017-03-29

forgetting index

the percentage of failed answers in spaced repetition reviews

the highest learning rate is achieved by a 15-30% forgetting index

2017-03-29

positive manifold

a finding that results of any pair of cognitive tests are correlated across individuals

2017-03-29

allele

a variant of a gene

2017-03-29

monogenic trait

polygenic trait

a characteristic that is produced by:
(1) a single gene
(2) more than one gene

2017-03-29

camaraderie

re-affirmation of mutual trust through fake aggression

2017-03-28

intelligence simstim

the enjoyable feeling of following thoughts of someone more intelligent

2017-03-28

cost maximization

optimizing the amount of resources that is going into a cause, without checking how much value is provided on the other end

2017-03-28

deepity

a superficial equivocation, which can be read as true but trivial, or deep but false

2017-03-28

epistemic ego

assuming that one is smart enough to not be fooled by spurious arguments, and updating in response to whatever sounds reasonable

2017-03-28

Pascal's mugging

give me five dollars, or I'll use my magic powers from outside the Matrix to simulate torture of 3^^^^3 people

2017-03-28

unspeakable moderation

a situation in which it is not politically permissible to advocate restraint, and the public discourse spirals out of control

2017-03-28

threshold error

a mistaken classification based on a standarized test, which is inevitable in some cases even if the threshold used in the test is globally optimal

2017-03-28

point and laugh

an automatic rejection of everything that pattern-matches to what one's ingroup has ridiculed in the past

2017-03-28

bingo card

a set of cached strawman versions of counter-arguments to a group's position, which prevents its members from ever updating

2017-03-28

semantic stopsign

a concept that makes someone drop a line of inquiry

2017-03-28

Courtier's Reply

telling someone that they need to go and learn a lot before they can have a meaningful position in the debate

2017-03-28

absurdity heuristic

a statement that sounds ridiculous to one's current intuition is probably wrong and doesn't need to be verified

2017-03-28

cow stare

observing something very intently, but without thinking what will happen next

2017-03-27

low order correlation

high order correlation

a pattern of:

(1) low level building blocks, regardless of how they are used

(2) high level structures, which requires parsing enough of the underlying levels to notice

on each higher level there is less data, and therefore more demand for intelligence or learning time

2017-03-27

epistemic rationality

instrumental rationality

systematically succeeding at:
(1) having accurate beliefs about the world
(2) achieving one's goals and values

2017-03-27

affect heuristic

relying on the balance of positive/negative feelings about something to make a decision

2017-03-27

mysterious answer

an explanation that hides confusion in ontologically basic concepts, without increasing predictive power

2017-03-27

arguments as soldiers

using arguments as one-time rhetoric devices to score points against an opponent

2017-03-27

bias blind spot

not noticing one's own cognitive biases, even when one knows they are likely to be present

2017-03-27

back to basics

progress in advanced levels of a skill can often be made by revisiting the basics

2017-03-27

thrombocyte

platelet

a type of blood cell that is responsible for clotting blood in case of injury

2017-03-27

leukocyte

white blood cell

a type of blood cell that is responsible for the immune response

> also found in the lymphatic system and other places in the body

2017-03-27

erythrocyte

red blood cell

a type of blood cell that uses hemoglobin to bind oxygen and transport it

2017-03-27

blood fractionation

centrifuging blood to separate it into blood plasma, leukocytes mixed with platelets, and erythrocytes

2017-03-27

attractor theory

seeing motivation as navigating a landscape filled with potential reward loops, which pull more strongly as one gets closer to them

2017-03-27

mindset optimization

using low-willpower activities to enter a mind state from which doing some task is as easy as possible

2017-03-26

small consistency

not having any ambitious opinions or values, to avoid being accused of hypocrisy

2017-03-26

artificial humility

engineer's humility

(1) pretending to have no strong opinions, to play along with social status
(2) taking specific actions in anticipation of one's own errors

2017-03-26

motte and bailey

using an easily defensible version of a claim when it's criticized, and a much broader version to make insinuations

2017-03-26

cycler trajectory

a spacecraft trajectory that encounters two or more bodies on a regular basis using minimal propellant

2017-03-26

adultism

the assumption that adults are better than young people, and can act on them without their agreement

2017-03-25

semiotic impulse

a tendency to choose actions for their symbolic meaning, instead of actual consequences

2017-03-25

as-if discounting

if prediction error grows linearly with time, a rational agent behaves as if it implemented hyperbolic discounting

2017-03-25

memeplex

a group of memes that are often found in the same individual

2017-03-25

egregore

an autonomous memeplex that is symbiotic with a large group of people

it terrorizes people into working for it, but also makes coordination at all possible

2017-03-25

preference falsification equilibrium

a situation in which everyone is pressured to publicly support the current ideology

when the bubble bursts at some point, it tends to be replaced with a different preference falsification equilibrium

2017-03-25

contractualism

ethics based on following only such rules that self-interested agents would like to precommit to

2017-03-25

virtue ethics

ethics based on cultivating traits of character

2017-03-25

consequentialism

ethics based on evaluating only outcomes brought about by actions

2017-03-25

deontology

ethics based on rules about classes of actions, regardless of their consequences

2017-03-25

cognitive uncontainability
strong cognitive uncontainability

when watching a superior intelligence:
(1) not being able to predict its actions
(2) even after seeing what it did, not understanding why

2017-03-25

will to lose

a tendency to self-sabotage in a bad situation, which makes it look more like one had agency

2017-03-25

scrupulosity

a constant fear of breaking moral rules, with a tendency to take them to non-useful extremes

2017-03-25

mana

the amount of mental energy that can be used to bend reality to one's will

2017-03-25

disruptive innovation

an innovation that sacrifices many standard fitness metrics,

to optimize for something new and unexpected

2017-03-24

Goodhart's Law

Campbell's law

when a social or economic indicator becomes a target for optimization, it loses its value as an indicator

2017-03-24

generalized shame disorder

a long-term non-specific feeling of social shame, which often causes self-hatred and self-destructive tendencies

2017-03-24

generalized anxiety disorder

excessive anxiety and worrying in many aspects of life

2017-03-24

terrifying non-violence

non-violent behaviour can be a ridiculously powerful and crushing force

2017-03-24

research debt

accumulation of ideas that are not processed for public consumption, leading to cognitive overhead and fragmentation of the field

2017-03-24

interpretive labour

the total work needed to understand an idea, which is spread across all parties that are communicating

2017-03-24

asymmetric weapon

something that on average works at least slightly better in service of truth than against it

2017-03-24

no true Scotsman

dismissing all counterexamples as non-central

2017-03-23

instant soup

a solution acquired without effort, which is quick and easy but bad for you

2017-03-23

school bench syndrome

feeling obliged to sit through classes, meetings, work etc. even when their purpose is clearly not achieved

2017-03-23

reversed stupidity

doing the opposite of a stupid action, which is just another kind of stupid

2017-03-23

bucket error

gluing together two or more claims when updating on evidence

if one claim is emotionally charged, the fear of change will propagate to other claims

2017-03-23

cold thermogenesis

a state of intensely burning fat to preserve core body temperature in a cold environment

2017-03-23

brown adipose tissue

brown fat

a type of fat tissue responsible for thermoregulation

it can turn free fatty acids etc. directly into heat (without muscle shivers)

2017-03-23

vagal tone

activity of the vagus nerve, which is a major component of the parasympathetic nervous system

2017-03-21

dumb fallacy

giving known liars the benefit of the doubt

2017-03-21

Benford's Law

if nothing is known about a number, the chance that it starts with a digit d is $\log_{10}(d+1) - \log_{10}(d)$

1 - around 30%
2 - around 18%

2017-03-21

alief

belief

information about the world stored in:
(1) patterns of behaviour and anticipation
(2) abstract and verbal cognition

2017-03-21

convergent equilibrium

divergent equilibrium

a stable situation in which everyone is trying to:
(1) invest more in the shared good
(2) manipulate others into investing in the shared good

2017-03-21

open-loop control

following a procedure that is supposed to work, without checking results at any point

2017-03-21

adaptive control

a control system with another system on top of it to adjust its settings

2017-03-21

try a dozen

a rule to try at least 12 candidates when making an important decision

2017-03-21

recognition heuristic

objects recognized as seen in the past are given more credibility

2017-03-19

sticky shadow

if someone desires to change but it is not happening, they are probably blocked by not fully acknowledging their current situation

2017-03-19

acknowledgement

acceptance

(1) admitting that something really is the case
(2) letting go of the desire to change it

2017-03-19

cell wall

a boundary that chooses what can go in or out, with a goal to optimize the environment inside

2017-03-19

marginal advantage

maintaining a small but maximally certain advantage over an opponent

2017-03-18

automoderation

using hand signals to decide who speaks next in a group discussion

2017-03-18

latent failure

a situation in which only one layer of protection is active, and a failure will happen at the next opportunity

2017-03-18

bikeshedding

 a tendency to spend the most time on debating trivial issues

2017-03-17

action points

a list of concrete next steps to take after a meeting, discussion etc.

2017-03-17

high modernism

forcible application of top-down reasoned rules and plans

2017-03-17

legibility of state

any centralized rule will try to make the country transparent and taxable, even at the cost of great harm

2017-03-16

backfire effect

crony beliefs becoming stronger in response to merit-based counterarguments

2017-03-13

just-world fallacy

assuming that goodness of actions will be eventually reflected in goodness of outcomes for that person

2017-03-13

bad guy bias

caring much more about harm inflicted by humans than by nature

2017-03-13

moral virtue

a heuristic that helps with a class of moral decisions that are consistently tricky for humans to get right on a case-by-case basis

2017-03-13

gold standard

clinical standard

a test for some property that is:
(1) definite but too expensive to use in every case
(2) inaccurate but practical

2017-03-13

bee sting fallacy

when facing too many serious problems, starting to ignore all of them

 a person with 6 bee stings doesn't have motivation to treat any one of them

2017-03-12

externalities

costs passed on to people who weren't involved in a transaction

2017-03-12

coordination problem

everyone agrees that a certain action taken by everyone would be globally best, but it is never the rational choice for any single person

2017-03-12

snuffing out intuition

introducing order and structure makes it harder to track the real targets (or even notice them)

2017-03-12

repression to fixation

repressed desires will often manifest as strong obsessions about some related topic

after placing many restrictions on oneself, the nearest unblocked strategy is something not directly related

2017-03-12

imagined nation

the image of a society created in people's heads by news reporting

2017-03-12

smooth texture

relating to something without ego, expectations, anxieties etc. getting in the way

2017-03-12

finite game
infinite game

a game with:
(1) fixed rules and a goal to win
(2) no fixed rules and a goal to continue playing

2017-03-12

Coolidge effect

renewed sexual interest after being introduced to new sexual partners

2017-03-12

hedonic treadmill

a tendency to return to the same level of happiness despite major events or lifestyle changes

2017-03-11

bottled frustration

an accumulated load of negative emotions after a long period of suppressing them

2017-03-10

radioactive vampire

a person that can not only spread memes, but also make other people want to spread them recursively

2017-03-10

prototype model

modeling something by maintaining an image of a typical version of it

2017-03-10

failing with abandon

giving up completely when the initial goal slips out of reach

2017-03-10

ritual

binding mental actions and states to the outside world

2017-03-10

nearest unblocked strategy

the strategy that is the closest to achieving what one really wants, from those that comply with external restrictions

2017-03-10

abstraction barrier

blocking the possibility of accessing implementation details, that are not part of the abstract interface

2017-03-10

ontology of mind

a framework used to think about cognition, which also generates the approach to taking action

2017-03-06

top-down design

bottom-up design

(1) translating a goal into lower level specifications, until implementation is possible with available tools

(2) making more and more tools, until solving the original problem is simple with them

2017-03-05

typicality checker

a brain feature that rates how well one fits in a self-reported category

2017-03-04

Blub paradox

the tools someone uses dictate how they think about solving problems, so they don't see the value of switching to a different paradigm

2017-03-04

synaptic plasticity

structural plasticity

plasticity of the brain based on:
(1) changing weights of existing neural connections
(2) connecting axons to new places

2017-03-04

emotional reset

compulsively subjecting oneself to extreme stimuli after a period of emotional oppression

2017-03-04

internalism

externalism

being concerned mostly with the:
(1) internal world
(2) external world

2017-03-03

there is no method

for any complicated task, there cannot be a single recommended way to do it, but rather multiple tools and taste about when to use them

2017-03-03

preference falsification

publicly expressing preferences that are different from what one thinks in private

2017-03-03

choice blindness

a tendency to confabulate that what one got was the preferred choice

2017-03-03

community of practice

a group of people that maintains and builds procedural knowledge in some domain

2017-03-03

orange meme

green meme

a worldview that emphasizes:
(1) taking rational action to manipulate objects in the world
(2) harmonious existence in the moment

2017-03-03

analytic philosophy

continental philosophy

a style of doing philosophy that is:
(1) precise and based on logic
(2) literary and based on the human point of view

2017-03-03

hard data

wet data

data that is:
(1) reliable and accurate
(2) ambiguous but leads in intriguing directions

2017-03-03

scientific integrity

going to extreme lengths to avoid fooling oneself while doing research

2017-03-03

epistemic virtue

a cognitive trait that reliably leads to having more accurate beliefs

2017-03-03

legitimate peripheral participation

learning from a community by working in minor roles that are fully accepted

2017-03-02

breaking the script

noticing that a conversation is following a script, and saying something out of character

2017-03-02

self-signaling

acting in a way that lets one maintain a certain self-image

2017-03-02

desperation to be good

a strong desire to self-signal virtues

2017-03-02

affective signaling

frequent small signals about one's affect towards individuals and groups

2017-03-02

social layer

people's emotions about other people and about all social interactions happening nearby

2017-03-02

goal space

a limited capacity to actually care about goals, which is often not big enough to add what one would wish they cared about

2017-03-02

Frankenstein's goal

a goal constructed out of a few bits of genuine caring, and many less healthy bits like signaling, identity, status, mimicry etc.

2017-03-02

quest-offering

creating the affordance for someone to go on a quest, without pressuring them to do it

2017-03-02

human foom

a quick run of explosive self-improvement after unblocking some aspect of one's psychology

2017-03-02

nothing works

an alief that all one's actions are ineffective and nothing ever changes

2017-03-02

root belief

a belief or alief that was not generated from other ones

2017-03-02

ripples

the effects of a conversation that unfold after it is finished

2017-03-02

holding space

staying with someone while they have some mental experience, without interfering or getting pulled into it

2017-03-02

agency hill

the local configuration of agency flowing outward and influencing the world

bigger and smaller hills all combine to form a single landscape, with features at all zoom levels

2017-03-02

controlled attribute

an aspect that someone consistently gets right, which means
they are probably modeling it or caring about it

2017-03-02

word cloud

understanding text as a bag of words without syntax

2017-03-01

scaremongering

spreading frightening rumours

2017-02-28

core state

an intrinsically desirable way of emotionally relating to the
world

2017-02-28

invalidation throughput

the rate at which one is able to uncover and get rid of
assumptions, which are the limiting factor of growth at a high
level

2017-02-28

throughput accounting

counting only items that successfully passed through all steps
in a chain

e.g. until purchased by a customer

2017-02-28

overlearning

using a new skill all the time, to make it more deeply
integrated

2017-02-27

thought spatiality

one's relation to thoughts can change if they are mentally moved in space

2017-02-27

teaching via environment

setting up challenges that people naturally encounter, engage with, and gain skills as a result

2017-02-27

intentional stance

reasoning about an object by treating it as a rational agent with its own goals and beliefs

2017-02-27

selfish neuron

a neuron seen as a self-interested agent trying to survive, which it can do by finding useful work to do

2017-02-27

felt sense

a worldless blob of intuition, which can be felt in the body (often in the throat, chest, or belly)

> it can be unpacked by hanging out with it and trying different thoughts

2017-02-27

detached map

a mental map that was constructed to be self-consistent, but never verified against reality

2017-02-27

fully general reminder

a reminder that would work for any task, and as such requires
more willpower to translate into action

2017-02-26

meta-meme

a meme about how to deal with memes

2017-02-26

read is write

reading information from a human mind is also changing it

2017-02-26

transparency subsystem

a brain mechanism that checks if one's belief network is
being reported honestly, and makes it known through body
language

2017-02-26

belief propagation

after some important belief is changed, updating other beliefs
and mental architecture to reflect the change

2017-02-26

tulpa

a purposefully created mental subagent

2017-02-25

farming

grinding

earning tokens by:
(1) predictable but time-consuming work
(2) unpredictable and rarely-successful challenges

2017-02-25

non-tokens

difficult to obtain items that are not by themselves worth social points

working to get these is not-fun

2017-02-25

display case

a Schelling point for showing off one's collection of tokens

2017-02-25

token economy

a system in which humans compete for scarce but otherwise worthless tokens, in order to get social respect

2017-02-24

strategic zoom

zooming in from big picture strategy all the way to concrete day to day tasks

2017-02-24

epistemic humility

being able to think "maybe I am the one who is wrong"

2017-02-24

reference level

the target level of some property that a control process is trying to maintain

2017-02-24

virtual reference level

a range between reference levels of two conflicting control

processes, in which they roughly balance out, leaving the variable uncontrolled

if the variable goes outside of the range, one of the control processes will start deactivating, letting the other one pull it back

2017-02-23

competitive advantage

something that lets one perform better than others at a given task

2017-02-23

Moloch

a tendency of civilization to optimize for things that aren't actually human values

2017-02-23

revealed preference

a preference that would be inferred from actual outcomes of someone's actions, not their professed intentions

2017-02-23

shadow denial

strongly acting out the opposite of what one's shadow believes

2017-02-23

illusion of transparency

a tendency to assume that words are understood as the speaker intended

2017-02-22

shadow master

shadow puppet

a person who is:

(1) very good at seeing people's shadows
(2) easy to manipulate because of their large shadow

2017-02-22

fertile ground

the mental space in which ideas are vulnerable and not yet fully formed, and it takes a lot of trust to expose them to someone

normally, the ideas get filtered through many layers before communication happens

2017-02-22

little agent fallacy

assuming that at least some individuals must have decided to bring about the outcome that happened

2017-02-22

wallflower

control freak

someone who in a group interaction:
(1) is not able to lower their defenses and really show up for it
(2) cannot resist trying to steer the group

2017-02-22

iron fisting

featherism

helping people by:
(1) forcing the situation into whatever you decided was good for them
(2) gently prompting them to think, without being attached to outcomes

2017-02-22

eating the shadow

slowly, gruesomely integrating the parts of experience that had previously been rejected

2017-02-22

shadow

emotions and parts of one's identity that were strongly rejected, and are hidden from the conscious mind

2017-02-22

uncanniness

the emotional experience of a conceptual interpretation breaking down

2017-02-22

monstrosity rejection

monstrosity inversion

dealing with one's own monstrosity by:
(1) hiding it and pretending it doesn't exist
(2) reveling in it and pretending not to care about the harm

2017-02-22

ripe idea

an idea for which enough cultural background has accumulated, that it is easily discovered

in many cases, several people hit on the same idea independently around the same time

2017-02-22

principal component analysis

transforming a set of observations to a different orthogonal basis, so that the axes are uncorrelated with each other

2017-02-22

friend shame

feeling social shame about someone else's behaviour

2017-02-22

fixed point

an element that is mapped to itself by a function

2017-02-22

attractor swamp

a region in idea-space such that any chosen theory will tend to be self-reinforcing

2017-02-22

meta PCK seeking

learning PCK about PCK seeking by modeling the mind of someone doing PCK seeking

2017-02-22

strategic level 3

trying to maximize one's general ability to preempt problems

2017-02-22

strategic level 2

trying to predict possible problems and make them never happen in the first place

2017-02-22

strategic level 1

trying to prevent problems from happening again

2017-02-22

self-referential stability

a state in which a tool or set of tools do not yield any

changes, when applied to themselves

motivated reasoning

using cognitive skill in order to defend a belief, rather than to improve accuracy of beliefs

social proof

guessing what is true based on other people's behaviour

gears-level understanding

an accurate causal model that gets verified or falsified by interacting with reality, without need for authority to approve it

frame-by-frame debugging

looking at a sequence of TAPs to find the exact point where things start to go wrong

polaris

something that helps keep a global direction despite locally overwhelming chaos

chaos principle

humans are mostly stuck in chasing local optima, so try more things

affordance factoring

finding out critical factors that are blocking affordances

2017-02-22

affordance with self

something that one perceives as a possible way of being, life goal etc.

2017-02-22

affordance

a possibility of action, perceived about an object or the environment

2017-02-22

Dilbert Dust

false familiarity

gradually losing curiosity about things that are often encountered

2017-02-22

inner asshole teenager

an inner voice that is always ready to be obscenely insulting

can be used as a form of outside view on oneself

2017-02-22

retroactive consent

checking if someone would have agreed, if there had been a chance to ask them

2017-02-21

race to the bottom

whenever someone figures out how to trade off a useful value

for greater competitiveness, everyone else is forced to follow

2017-02-21

dictatorless dystopia

a political system which every single citizen (including the leadership) hates but which nevertheless endures

2017-02-21

utopia[/juːˈtəʊpɪə/]

dystopia[/dɪsˈtəʊpɪə/]

an imagined place or state in which the experience of citizens is:

(1) pleasant and perfect

(2) unpleasant and horrifying

2017-02-21

felt shift

resolving a vague intuition into something that is spot on, with a clear sense of relief and released tension

2017-02-21

shaping

reinforcing successive approximations of desired behaviour

2017-02-21

good faith principle

assuming that other people's actions are good and sensible from their point of view

2017-02-20

true challenge

a challenge with potential to leave one completely crushed

also gives the biggest boost to personal growth

2017-02-20

meta-cognitive blind spot

not noticing something about one's own thinking, even as it's happening

2017-02-20

wearing the fpp

switching from imagining events in third person, to first-person perspective

2017-02-20

electrodermal activity

galvanic skin response

electrical conductance of the skin, which depends on the state of sweat glands (controlled by the SNS)

2017-02-20

central nervous system

peripheral nervous system

a part of the nervous system in vertebrates, consisting of:
(1) the brain and spinal cord
(2) nerves connecting them to other parts of the body

2017-02-20

somatic nervous system

autonomic nervous system

a part of the peripheral nervous system, responsible for:
(1) voluntary control of skeletal muscles
(2) unconscious operation of internal organs

2017-02-20

sympathetic nervous system

parasympathetic nervous system

a part of the autonomic nervous system, responsible for the:
(1) fight-or-flight response
(2) rest-and-digest body functions

2017-02-20

flattened knowledge

knowledge that was generated by an intricate and fascinating process, but shown in a compressed form seems bland and uninteresting

2017-02-20

genre-savviness

if you are the main character in a story, what is the next step to move the plot forward?

2017-02-20

bug statics

bug dynamics

(1) what is stuck, and how we know that it is
(2) what happened before, and what might change if action is taken

2017-02-20

obvious advice

advice coming directly from someone's common sense, without cleverness, contrarianism etc.

2017-02-20

creative restriction

boosting one's ability to generate ideas by adding various restrictions to the problem

2017-02-20

procedural knowledge

declarative knowledge

knowledge shown by:
(1) doing some task
(2) describing something in words

2017-02-20

positive genie

negative genie

a genie that offers a deal to grant positive/remove negative consequences of an action, in exchange for no longer being able/having to do the action

2017-02-20

PCK seeking

closely following and modeling someone's mind, while they are solving a problem

2017-02-20

pedagogical content knowledge

PCK

the knowledge of what happens in the mind of a student

2017-02-20

content knowledge

pedagogical knowledge

the knowledge of:
(1) the subject that one is teaching
(2) general teaching methods

2017-02-20

diaphragm breathing

chest breathing

a style of breathing in which mostly:
(1) the belly moves
(2) the chest moves

2017-02-20

primary attention

secondary attention

paying attention to:
(1) a narrow spot that is directly looked at
(2) sensations perceived outside of that spot

2017-02-20

hold off solutions

a rule to think about the problem carefully before proposing any solutions in a group

2017-02-20

socratic duck

a person who helps someone else think by themselves, by listening and asking simple questions

2017-02-20

empiricism

using sensory experience of the world as a basis for knowledge

2017-02-20

minimum information principle

learning active recall is easier if information is split into simple, small chunks

2017-02-20

double crux

two people finding a common crux for their differing opinions, such that if the crux is resolved, one of them will change their mind

2017-02-20

sometimes snack

something that is valuable only if used in moderation and at appropriate times

2017-02-20

value of information
VoI

how much the EV of the best decision would be improved by knowing something

2017-02-20

price benchmark

an emotionally salient good with a known price, used to estimate opportunity cost

2017-02-19

cached thought

a processed thought stored in memory for quick lookup in the future

2017-02-19

social potential energy

to what degree people are just waiting for an excuse to do something

2017-02-19

phoenix fire

a burning desire to save the world

2017-02-19

yin

yang

an aspect that is:
(1) passive, diffuse, healing, female
(2) active, focused, energetic, male

2017-02-19

load-bearing bug

a problem such that if fixed, some vital structures would collapse

2017-02-19

bug

something that systematically goes wrong in one's life

2017-02-18

memetic defense

an intuitive mental flinch away from looking at content that could have harmful influence

2017-02-18

bullshit partition

a mental space for information that is untrusted or not worth remembering

2017-02-18

narrative lens

seeing how people's lives are controlled by what stories they live in, and what roles they play

2017-02-18

hufflepuff armour

the mental stance of being unstoppable at protecting some important value, even at the cost of internal violence

2017-02-15

intravenous[/ˌɪntrəˈviːnəs/] therapy

IV

delivering fluids directly into a vein

2017-02-15

ion

an atom or a molecule with a different total number of electrons than protons

i.e. with a positive or negative electrical charge

2017-02-15

electrolyte

a liquid or gel that contains ions and can be decomposed by electrolysis

2017-02-15

lactate threshold

the intensity of exercise at which the blood concentration of lactic acid starts to rapidly increase

typically around 85% of maximum heart rate

2017-02-15

lactic acidosis

a sudden medical condition from having too much lactic acid in the bloodstream

2017-02-15

probiotic[/ˌprəʊbʌɪˈɒtɪk/]

a substance which stimulates growth of the intestinal flora

2017-02-15

esophagus[/ɪˈsɒfəgəs/]

a muscular tube that connects the throat to the stomach

2017-02-15

pancreas[/ˈpaŋkrɪəs/]

a large gland behind the stomach which secretes digestive enzymes into the duodenum

2017-02-15

cyst

a thin-walled cavity filled with a liquid

2017-02-15

if-fail-then-ramp

an intuitive mental algorithm, which prescribes retrying each failed action with more strength

2017-02-15

emotional micro-bid

making a small emotional investment in social interaction, which pays off but only if it gets a response

2017-02-15

emotional validation

satisfaction from getting a visible emotional response from someone

2017-02-14

peristaltic wave

a contraction of muscles that propagates along a tube, pushing its contents

2017-02-14

gastritis

inflammation of the stomach lining

2017-02-14

murphyjutsu

using the surprise-o-meter to map failure space and improve plans

2017-02-12

Internal Double Crux

IDC

resolving an inner conflict by taking all sides seriously, and helping them understand and address each other's concerns

2017-02-11

personal boundary

the edge of an area that is fully in control of one person, such that the weight of other people's preferences about it is automatically assumed to be zero

2017-02-10

turbocharging

making sure you practise the exact thing that you want to learn

optimally, the conditions are varied and as close to real world use as possible

2017-02-10

scarcity mindset

abundance mindset

being very/not at all concerned with conserving some resource

2017-02-10

attention trade

letting one's attention be drained a little bit more, in exchange for saving some other resource

2017-02-10

lazy evaluation

delaying actually computing an answer until it is necessary to output it

2017-02-10

OODA loop

Observe
Orient
Decide
Act

a basic cycle of decision-making

2017-02-10

relevance check

if everything else was perfect, would this one issue still weigh on the decision?

2017-02-10

perverse transfer

a skill transfers but with unwanted artifacts from training

e.g. "disarm the guy and hand the gun back to him"

2017-02-10

skill transfer

the degree to which a separately trained skill is useful in real world situations

2017-02-10

scope insensitivity

scope neglect

emotional reactions do not scale with increasing absolute numbers

2017-02-10

reductionism

the belief that complex systems are composed of nothing but simple parts interacting with one another

2017-02-10

failure map

a map of possible types of failures, weighted by probability

some space is always reserved for unknown unknowns

2017-02-10

Murphy's law

anything that can go wrong, will go wrong

2017-02-10

adaptive problem

technical problem

a problem that is:
(1) vague and confusing, with no clear directions to proceed

(2) about how to obtain a certain result

2017-02-10

idea inoculation

becoming immune to a useful idea, after hearing a garbled/strawman version of it first

2017-02-10

unflattening knowledge

boggling at a piece of compressed knowledge, to understand and learn from the process that generated it

2017-02-10

boggling

staring at something with exaggerated disbelief and curiosity

useful to unpack flattened knowledge

2017-02-09

surprise-o-meter

asking the inner simulator how surprising some future outcome would be

2017-02-09

equal and opposite advice

every time someone needs advice X, someone else needs to be told the opposite of X

2017-02-09

bullshit bell

an intuitive feeling of distrust when someone's claims are anticipated to be false

2017-02-08

responsiveness

throughput

the ability to:
(1) notice many weak stimuli
(2) handle a single strong stimulus

2017-02-08

desire handle

an emotionally salient image, phrase, song etc. that activates
a S1 preference

2017-02-08

mental contrasting

visualizing a desired state of the world on top of reality

2017-02-08

elegance superstimulus

using a simpler and more elegant pattern to override
someone's intuition

2017-02-07

procrastination crux

asking "what is easier than that?" until it's painless to start
working towards the goal

2017-02-07

flipped classroom

having students watch video lectures at home, while devoting
class time to exercises, discussions etc.

2017-02-07

factoring

replacing an action with all its desirable and undesirable

consequences, and then coming up with better plans to satisfy those

trigger-action plan

TAP

an intention to do an action when some condition is met

focus

locally narrowed confidence intervals

following from ahead

figuring out the territory in the direction someone is going, before they reach those places

contour plot

a graph that shows a two-dimensional continuous function as outlines of its cross-sections

goal separation

the degree to which a system's goal is a black box, which does not depend on other details of the system

building form

doing easy tasks while maintaining all the best practices, to prepare oneself for much harder tasks

tool AI

an AI that optimizes only for narrow human preferences about its immediate actions

2017-02-06

narrow human values

ambitious human values

values that:

(1) we currently pursue and approve of

(2) represent the ultimate destination of our preferences, but implemented now might clash with what we approve of

2017-02-06

cognitive reduction

dissolving confusion by finding what cognitive algorithm generated that confusion

2017-02-06

patch resistance

a property of problems that cannot be securely fixed by incremental changes on top of the same architecture

2017-02-06

KANSI AI

a "known-algorithm non-self-improving" AI design

2017-02-06

extrapolation

generalization

(1) predicting that an observed regular pattern will continue in the future

(2) fomulating a rule that applies to a broader class of situations than those directly observed

2017-02-06

model building

increasing understanding in terms of simpler components and their interactions

2017-02-05

astroturfing

making a promotion campaign look like spontaneous, grassroots behaviour

2017-02-05

intent to cooperate

willingness to learn and follow the spirit behind the rules, rather than just how not to get caught

2017-02-05

two clever arguers

letting two sides selectively argue opposite points, so that a neutral observer might form an unbiased opinion

   this setup doesn't work when some side of the issue is under-represented

2017-02-05

clever arguer

someone who starts with the bottom line already written, and looks for arguments to justify it

2017-02-05

bullshit test

a test to see if someone will assert themselves and refuse to be tested

2017-02-05

one-sided policy

an imaginary policy issue that has an obviously correct solution, with no good arguments against it

2017-02-05

tactical summarizing

summarizing what happened in a discussion until now, while using the opportunity to control the frame

2017-02-05

fuzzy pattern

compact pattern

a pattern that is:
(1) in the background, with a lot of flexibility remaining
(2) explicit and varbalized

2017-02-05

pattern clash

fighting someone's compact pattern with another compact pattern

2017-02-05

crystallizing move

turning a fuzzy pattern into a compact one, while intentionally choosing how to frame it

2017-02-05

air bending move

influencing fuzzy patterns, which make up the implicit frame for everything people do

2017-02-05

analysis delay

how long it takes to start noticing problems with some view or

narrative

it's a good measure of its memetic strength

2017-02-05

synthetic epiphany

prompting someone to have a big realization, that becomes a part of their identity

2017-02-05

narrative immunity

a tendency to resist certain narratives and rationalizations that grows with exposure to them

2017-02-05

narrative injection

introducting a different narrative as a way to influence someone

2017-02-05

reality distortion field

the impression of reality adjusting to match the vision of a charismatic person

2017-02-05

taste

quick, intuitive yuck/yum reactions that encode much of experience and expertise

2017-02-05

cargo cult

imitating something as a ritual, while it is actually based on precise and technical skill

2017-02-04

premortem

imagining that a plan had already failed, and the reasons for it

2017-02-04

non-contention

being free of the desire to have one's results evaluated or noticed

2017-02-03

exploration phase

exploitation phase

being strategically focused on:
(1) finding possible approaches and evaluating them
(2) single-mindedly using the best approach found so far

2017-02-03

mental fishing

taking a moment to block external signals and promote subtle intuitions to awareness

2017-02-03

wind-up toy

someone with an ability to quickly switch between rest and high performance

2017-02-03

Dunbar's number

a cognitive limit to the number of people with whom one can maintain active social relationships

2017-02-03

semantic power

memetic power

a skill based on:
(1) accuracy and clarity of analysis
(2) noticing and manipulating patterns

related: filtering/generation

2017-02-03

pattern unmatching

dropping a prematurely matched pattern, after noticing that its predictions aren't accurate

2017-02-01

sham option

an alternative that isn't really viable, added only to make something else look better

2017-02-01

fall in love twice

a rule to keep looking until you think you have found a perfect match two times

2017-02-01

ego conflation

ego splitting

(1) a single project becomes a matter of personal pride
(2) keeping several options open to avoid excessive emotional attachment

2017-02-01

sequential feedback

parallel feedback

feedback given about:

(1) each next iteration
(2) several variants at once

2017-02-01

multitracking

working on several different versions of a project, idea etc. at the same time

2017-02-01

vanishing options test

what would you do if all the options you are considering now became impossible?

2017-02-01

opportunity reminder

what else can you get for the same cost?

2017-02-01

whether-or-not decision

considering only one real option, with doing nothing as an alternative

this kind of decision is particularly likely to fail

2017-02-01

trigger wire

something set up in advance to interupt the autopilot and force a decision point

2017-02-01

WRAP process

1. Widen your options.
2. Reality-test your assumptions.
3. Attain distance before deciding.

4. Prepare to be wrong.

2017-02-01

manual spotlight

systematically bringing different aspects of the situation into view

2017-02-01

being a new homunculus

acting as if being someone else who has just been given control of this body, with zero emotional attachment

2017-02-01

glorified spreadsheet

making a list of available options, and grading each of them on a set of features

this helps with some well defined technical problems, but most of the time, the bottleneck of the decision is outside of this narrow frame

2017-02-01

decision process

the approach to making a decision, including what and how actually influences its result

2017-02-01

analysis breadth

analysis depth

the effort spent on:
(1) adding fresh perspectives and people
(2) analysing with more rigour and detail

for most real-world decisions, (1) is more useful

2017-02-01

spotlight effect

a tendency to form opinions based on readily available information, without pausing to check if it is adequate

2017-02-01

identity drift

human brains are wired to have changing values and preferences over an individual's lifetime, and to make this process transparent to them

2017-01-30

classification problem

regression problem

a problem of predicting (depending on values of some other variables):

(1) a discrete variable

(2) a continuous variable

2017-01-29

moral parable

a story about someone learning a harsh lesson

unusually sailent for intution, even though often based on extremely weak evidence

2017-01-28

step in utility

a steep increase of utility when invested effort becomes sufficient to pass some important threshold

2017-01-28

binary thinking

assuming that if something is not true, than one can go however far in the opposite direction

2017-01-27

infection load

immune capacity

the amount of work that the immune system:
(1) uses to deal with new incoming infections
(2) is able to do in total

2017-01-27

inside view

outside view

predicting from the perspective of:
(1) being involved personally and knowing all the details
(2) knowing only a basic summary and a track record for the reference class

2017-01-27

positive illusions

unrealistically favorable opinions that people have about themselves or people that are close to them

2017-01-22

Batman effect

people get excited when a powerful person steps in and personally deals out justice, while this is actually a waste of resources

2017-01-22

urge from below

urge from above

an urge that is:
(1) short-term and concerned with sensations or feelings
(2) long-term and concerned with outcomes in the world

2017-01-22

executive avatar

inner daemon

(1) the body and most aspects of the mind, seen as a programmable intelligent avatar

(2) deep, long-term processes that adjust its configs

2017-01-22

godshatter

a much smarter algorithm than is natively implemented in one's brain, that can only be computed very slowly in small fragments

2017-01-22

grief

a negative emotion triggered when a control process can no longer achieve its purpose

2017-01-22

perception process

control process

a brain subroutine that:

(1) extracts higher-level features from its input data

(2) coordinates several lower level neural structures to maintain some useful property

> both of these are organised hierarchically

2017-01-22

donor lottery

an arrangement in which a group of people pools their charity money and randomly picks one of them to decide how to use it

even when there are no returns to scale, this reduces expected time spent on research

2017-01-22

shame

guilt

a feeling activated in humans when:
(1) they might face social repercussions
(2) they associate someone's suffering with their actions

2017-01-22

game temperature

how much actions in the next few moments can swing long-term outcomes

2017-01-21

strictly proper scoring rule

a scoring rule that gives maximum expected score only for reporting the true probability distribution

2017-01-21

scoring rule

a function that measures accuracy of probabilistic predictions

2017-01-21

Church-Turing thesis

computability thesis

a function on the natural numbers is computable by a human following instructions, if and only if it is computable by a Turing machine

assuming the human has infinite time and resources

2017-01-21

hard power

soft power

the ability to:
(1) force the desired outcome despite resistance

(2) persuade others without using force

2017-01-20

Cohen's d

the difference between two means, divided by pooled standard deviation

a measure of effect size

2017-01-20

nootropics[/ˌnəʊəˈtrəʊpɪks/]

drugs that improve cognition in healthy individuals

2017-01-20

diversion from content

many social skills are designed to make interactions only about the social world itself

this is optimal for people focused mostly on advancing their social interests

2017-01-20

fully general argument

an argument that can be used to argue any claim, which makes it necessarily fallacious

2017-01-20

fully general skill

a skill that would work in any circumstances, which makes it likely to be shallow or dishonest

2017-01-20

diplomatic world

engineering world

a separate sphere of activity around:

(1) the social environment
(2) technical issues in the world

2017-01-20

central case

an example that carries the weight of the most important issues, and solving it would go a long way towards a general solution

2017-01-20

scale separation

robustness of behaviour against small perturbances on lower levels, which makes it predictable without modeling those lower levels

e.g. logic operations in a computer are separated from the electric currents etc.

2017-01-20

physical Church-Turing thesis

any computation that can be physically done in our universe can also be done on a Turing machine

2017-01-20

falsifiable design

a design bundled together with proposals of experiments to reduce uncertainty about each aspect of it that has not been previously tested

2017-01-20

whole brain emulation

WBE

a software model of a whole brain that faithfully represents its internal causal structure

2017-01-18

big kid

a person who runs mainly on reinforcement learning, as a way to decide what to do in life

2017-01-18

Kurisu-Paraglider Effect

when excited about something, seeing it as the ulimate best thing anyone could ever do

2017-01-17

senescent cell

a cell that stopped dividing but is still alive

2017-01-16

killing one's darlings

letting go of one's ideas and creations, as soon as they can be replaced with something better

2017-01-14

angular frequency

angular speed

$\frac 2\pi T$ [$\frac rad s$]

T - period [s]

2017-01-14

electron hole

lack of an electron at a position where one could exist in an atom or atomic lattice

2017-01-14

electrical conductor

electrical insulator

a material whose internal electric charges:
(1) flow freely, allowing it to pass electrical current
(2) do not flow freely

2017-01-14

dielectric material

an electrical insulator that can be polarized by placing it in an electric field

2017-01-14

Cherenkov radiation

electromagnetic radiation emitted when a charged particle travels through a dielectric medium with a speed greater than the phase velocity of light in that medium

2017-01-14

experiential []amnesia[/amˈniːzɪə/]

a human tendency to forget even the most crucial subjective experiences soon after they stop

2017-01-12

self-efficacy

confidence in one's ability to complete tasks and reach goals

2017-01-12

friend

ally

someone who cares about:
(1) you as a person, allowing arbitrary exchanges of favours
(2) your interests regardless of whether you are around

2017-01-12

stateless conversation

answering only the last query, with no memory of what was said before

arguing and small talk performed by humans comes very close to this

2017-01-12

elephant whisperer

someone who gives their conscious thoughts a powerful form, so that they stick

in the style of the best mnemonic techniques

2017-01-12

underlying cause

network of symptoms

modeling a problem as:

a single point of failure, which has many symptoms but can be fixed at the source

a network of surface issues that reinforce each other, stable only as a group

2017-01-12

canonical writing

a set of texts that members of an intellectual community are expected to have read

2017-01-11

behavioural activation

improving mental health and the general ability to do stuff by successfully doing any task at all

2017-01-11

Reddit liar

if enough people participate in a community, its most surprising stories will be blatant lies (as opposed to coincidences, honest mistakes etc.)

2017-01-11

narcissistic fallacy

assuming that one's conclusions or values are enough to persuade other people

to improve shared knowledge, it is needed to exchange verifiable sources and interact with criticism

2017-01-10

compulsion to closure

a human tendency to like when things are completed

2017-01-09

quality target

a range of desired results, which is a good trade-off between cost and tightness of the specification

2017-01-09

antibody

a large protein that binds to antigens, in order to tag them or neutralize them directly

2017-01-08

face time

the amount of time spent interacting with people face-to-face, which determines the quality of relationships

2017-01-08

show not tell

a principle to demonstrate emotions by action, not by giving them names

2017-01-06

stimulating influence

relaxing influence

a basic dimension in any work with people's minds and bodies (e.g. social meetings, massage, art)

2017-01-06

despair cycle

alternating between periods of manic enthusiasm and depressive doubting when working on a long, difficult project

2017-01-05

purposeful non-use

not taking shared resources that seem available, to redirect more to some purpose that allocates more gently

e.g. the immune system demands only the absolute minimum, and can be helped by forced rest

2017-01-05

slow-twitch fibres

fast-twitch fibres

muscle fibres that:

(1) contract slowly and are resistant to fatigue

(2) contract quickly but use energy much faster than they can produce it

e.g. mostly (1) in back muscles that maintain posture, mostly (2) in arm muscles

2017-01-05

antigen[/ˈantɪdʒ(ə)n/]

a pathogen that causes the immune system to produce antibodies

2017-01-05

internal family system

IFS

imagining various components of one's personality as separate people, and having them talk

2017-01-05

slice of joy

a brief moment of noticing joy or happiness, often about simple things (e.g. physical sensations)

2017-01-05

identity-based decision

a decision based on "what would a person like me do in this situation?"

2017-01-04

hope deficit

being too quick to accept things as necessary evil

2017-01-04

simulator expansion

imagining doing something, so that a new pattern gets registered as a possibility in the inner simulator

2017-01-04

intellectual snobbism

making others feel low status about their level in intellectual endeavours

many people can't stand it, and it is a big obstacle to transferring knowledge

2017-01-04

third level critic

a critic that intelligently defends their criticism, when it is responded to

2017-01-02

consumption value

investment value

value for:

(1) immediate use, with clear benefit from each unit spent

(2) long-term prospects, weighted by probability of success (which might be small)

2017-01-01

autoignition temperature

kindling point

the lowest temperature at which a material will spontaneously ignite in normal atmosphere (without external source of ignition)

2017-01-01

flash point

fire point

the lowest temperature at which vapour exposed to a flame or spark will:

(1) ignite briefly

(2) burn for at least 5 seconds

(2) is typically around 10 °C higher

2017-01-01

flashover

near-simultaneous ignition of exposed combustible material in an enclosed area

typically at 500 °C or 590 °C in rooms

2016-12-29

altitude anorexia

significant weight loss from high altitude exposure (in non-

acclimatized humans)

low oxygen probably influences regulation of appetite

2016-12-28

social hack

setting up the form of friendly social interaction, which causes genuine friendly feelings

2016-12-27

self inoculation

infecting oneself by touching some object and then one's own eyes, nose, or mouth

2016-12-27

refractive index

$$\frac c v$$

c - speed of light in vacuum
v - phase velocity of light in the medium
e.g. for water: 1.333

2016-12-26

conservative epistemology

relying only on explicit and verifiable sources, methods, arguments etc.

this slows people down a lot, but also roughly filters out many biases and perverse incentives

2016-12-26

crony preference

a preference that is kept around mostly for its signaling value

2016-12-26

hindsight bias

a strong tendency to overestimate one's past knowledge and predictions, after the outcome becomes known

2016-12-26

retrodiction

a prediction claimed after the outcome is already known

hindsight bias makes this very suspicious

2016-12-22

Wilde's Law

the bureaucracy is expanding to meet the needs of the expanding bureaucracy

2016-12-20

desire to pass tests

being intrinsically motivated to satisfy authority figures with one's performance

2016-12-20

Stanford Marshmallow Experiment

the experiment in which children could eat one marshmallow, or wait 15 minutes and eat two

2016-12-19

long-time perspective

choosing what to do now by how it influences the future in 10-20 years or more

2016-12-19

perspective taking

a discussion tactic based on explicitly taking someone's point of view and explaining it

2016-12-18

22° halo

a ring with a radius of approximately 22° around the Sun or the Moon, visible when light is refracted in hexagonal ice crystals

2016-12-18

cosmological constant

the simplest way to add dark energy to gravitational equations (by assuming it is constant in both space and time)

adding it adjusts universe models towards more expanding

2016-12-18

eternalism

nihilism

claiming that ethics, values, meanings etc.:
(1) are given by some ultimate external source
(2) do not exist and are not important

2016-12-18

continuum emission

emission line

emission:
(1) spread over a wide range of wavelengths
(2) in a narrow peak at a characteristic frequency

2016-12-16

burdensome details

details added to a claim or prediction, which necessarily make it less probable (even as the storytelling becomes more believable)

2016-12-16

superforecasters

people who are consistently above-average at forecasting, even outside of their domain of expertise

this screens off being an expert, except when explicit technical knowledge applies

2016-12-15

meta-first approach

before showing content, commenting on the choices made while creating it

2016-12-13

tactical asking for advice

using asking someone for advice as a way to influence them or get an ally

2016-12-12

galvanic vestibular stimulation

stimulation of the nerves in the inner ear using current from an electrical battery

for humans, 9V is reported to work well

2016-12-12

vestibular stimulation

stimulation of the nerves in the inner ear

e.g. in a centrifuge, on a swing, or using electricity
this strongly activates the PNS

2016-12-12

action orientation

failure to execute

(1) acting immediately and efficiently to get the important tasks done

(2) spending time planning, talking etc. instead of doing the thing

2016-12-09

prefrontal processing

thinking that activates mostly the prefrontal cortex (e.g. planning, problem solving, organising information)

2016-12-09

direct experience network

the brain regions which activate when focusing mostly on sensations in the moment

2016-12-09

cultural bridge

established support for people who are transitioning between some stages (e.g. in life or understanding of some topic)

2016-12-07

mama-bear efect

women are on average less assertive than men, unless they are fighting for someone else

2016-12-07

Doppler effect

the change in frequency observed by someone moving relative to the source of a wave

2016-12-07

redshift

increase in wavelength of electromagnetic radiation (e.g. when the observed object is moving away)

2016-12-07

Wien's displacement law

black body radiation curve peaks at a wavelength inversely proportional to the temperature

2016-12-07

radio waves

electromagnetic radiation with wavelengths from 1 meter (300 MHz) to 100,000 km (3 Hz)

2016-12-07

microwaves

electromagnetic radiation with wavelengths from 1 millimeter (300 GHz) to 1 meter (300 MHz)

2016-12-07

X-radiation

electromagnetic radiation with wavelengths from 0.01 to 10 nanometers

2016-12-07

ultraviolet

UV

electromagnetic radiation with wavelengths from 10 to 400 nanometers

2016-12-07

visible light

electromagnetic radiation with wavelengths from 400 to 700 nanometers

2016-12-07

infrared

IR

electromagnetic radiation with wavelengths from 700 nanometers to 1 millimeter

2016-12-07

Draper point

the temperature above which most solid materials visibly glow as a result of blackbody radiation (525 °C, 798 °K)

2016-12-04

attention lens

seeing how every change in people is driven by what they pay attention to

2016-12-04

habit lens

seeing the behaviour of people and organisations as driven by habits

2016-12-03

perverse incentive

an incentive that encourages behaviour contrary to the interests of the incentive maker

> e.g. paying for solved problems discourages prevention

2016-12-03

Keynesian economics

treating aggregate demand as the main factor that influences the state of an economy

> esp. in the short run and during recessions

2016-12-03

problem focus

solution focus

thinking about:
(1) what is or was wrong
(2) how to make things work

2016-12-03

micromanaging

attempting to control details of someone's work

this attacks their autonomy and status, undermining productivity in skilled work

2016-12-02

post-scarcity society

a society that satisfies everyone's basic needs and (a significant proportion of) desires without requiring labour or payment

2016-12-02

Say's law

aggregate production equals aggregate demand

generally accepted throughout the 19th century
during the Great Depression of the 1930s, replaced with Keynesian economics

2016-12-02

workforce productivity

the amount of goods and services produced per total amount of human work (e.g. hours)

2016-11-30

fear of being corrected

the default response to a possibility of a crony belief being challenged with facts

the brain has other incentives to believe than just accuracy, and it will resist fact-checking

2016-11-30

aposematics

a signal evolved to warn predators

originally warns about a real defense mechanism, but later can be exploited and mimicked by other species

2016-11-30

merit beliefs

crony beliefs

beliefs that are useful by:
(1) being accurate
(2) making some impression in social or political games

by analogy to employees in a company
the brain is highly optimized to gather material for tribal politics, and the conscious ability to lie is often not good enough

2016-11-29

Hebb's Law

neurons that fire together, wire together

2016-11-29

carrot-and-stick

attention approach

changing someone's behaviour by:
(1) rewards and punishments
(2) getting them to notice things

(1) only works for animals and small children

2016-11-29

background goal

a goal that orients other activities around it, so that it is

fullfilled without directly working on it

2016-11-28

settling

swimming

making choices that lead to:
(1) comfort and boredom
(2) continuous effort and new chances

2016-11-27

problem notebook

a notebook with unsolved problems and space to write
ideas/solutions for each of them

2016-11-26

agentization

an instinctive tendency to approach playing against reality
like negotiating with a human

this is behind many cognitive biases, e.g. sunk cost fallacy, gambler's fallacy

2016-11-23

contrarian reflex

wanting to do the opposite whenever someone seems to be
trying to manipulate you

2016-11-22

internal explanation

external explanation

explaining:
(1) how and where something is implemented
(2) what function it appears to perform, if we know nothing
about the implementation

these are often used to argue against each other, while they actually
complement

2016-11-20

beware the man of one study

results of respectable studies have a bell distribution with wide tails, so a single study can easily be picked to show whatever

2016-11-19

choiceless society

modern society

a society that:

(1) does not question the rules and assumptions on which it is built, and does not make people aware that such questioning is possible

(2) organises itself on system level, by deciding about rules and structure

2016-11-19

authenticity

factual accuracy

honesty understood as:

(1) not filtering what one is saying

(2) being as close to objective truth as possible

these correspond to calibration for audiences with emotional or nitpicky responses

2016-11-17

vendetta

revenge

(1) openly and obviously responding in kind, so that you don't lose social status points

(2) pursuing a long-term and covert strategy to satisfy the sense of fairness

2016-11-17

nominal value

commodity value

the value:
(1) designated with a number
(2) when sold as raw materials

2016-11-17

Gresham's law

bad money drives out the good

coins with lower commodity value will become dominant on the market

2016-11-17

affective language

contractual language

communication understood as:
(1) influencing emotions and social relationships
(2) encoding facts and logical reasoning

2016-11-17

crying wolf

raising an alarm when nothing is really wrong, which teaches people to ignore alarms

2016-11-16

Substitution Principle

when asked a hard question, S1 will find a related question which is easier and answer it instead

2016-11-16

Godwin's law

as an online discussion grows longer, the probability of a comparison involving Hitler approaches 1

2016-11-16

populism

a political ideology that sees citizens as mistreated by a small circle of elites, and advocates overthrowing them

2016-11-16

late-stage democracy

a democracy with extreme freedom and equality, in which voters can easily destroy the system from the inside

2016-11-15

pain of being ahead

the mental pain of watching someone struggle when you know the solution

2016-11-15

task focus

people focus

thinking about:
(1) details of the task at hand
(2) the people who are working on it

2016-11-15

vitrification

the transition from a liquid to an amorphous solid (glass), when the temperature is lowered

2016-11-15

glass transition

the reversible transition of an amorphous material between a solid and a viscous liquid

2016-11-15

Young's modulus

elastic modulus

the ratio of pressure to proportional deformation of a material

2016-11-15

viscosity

resistance of a fluid to deformation

2016-11-14

closed to-do list

an ordered list of tasks to do in one day, which cannot be
changed (including adding more items)

2016-11-14

berserker pathway

integrated pathway

acting on emotions by:
(1) letting them escalate and control outward behaviour
(2) capturing them and feeding into the decision process

2016-11-14

acausal monkey contract

being nice and doing your part to the extent that would be
mandated by social instincts, without second-guessing them

e.g. being a productive member of the society, caring about one's parents

2016-11-13

inner sim

the ability to imagine something, and "play forward" to see
possible futures

2016-11-13

explicit skill

soft skill

a skill that is learned:
(1) consciously, with knowledge of its internals
(2) by letting the subconscious mind interact with it

2016-11-12

writing-as-explaining

writing-as-thinking

writing:

(1) about well understood things, using a previously planned structure

(2) to explore and clarify one's own ideas

2016-11-12

messy workshop

having many projects "in progress", to be able to leave them and come back with new ideas

2016-11-08

SCARF model

Status
Certainty
Autonomy
Relatedness
Fairness

five primary reinforcers related to social interaction

2016-11-06

O-ring sector

foolproof sector

a sector of the economy which is:
fragile - low skill breaks the product

not fragile - low skill decreases efficiency, but can be compensated for with the number of workers

wage differences within a country are settled by how much more output high-skill workers have in the foolproof sector (not very big)
wage differences between countries are decided by the quality of their high-skill sectors (can be very big)

2016-11-05

elderly Hispanic women effect

finding a statistically significant effect after trying out many variants of the hypothesis

if you look long enough, you'll find something that gives a strong positive result just by coincidence

2016-11-05

reversal test

checking consequences of doing the opposite action, to avoid status quo bias

2016-11-05

status quo bias

preferring the current state of things, while something else would be preferred if it was already implemented

2016-11-05

self-promotion bias

thinking that one would make a better leader than other people, based on the inside view

2016-11-03

victim story

villain story

a biased narrative, in which someone is shown as:
(1) an innocent victim with only positive traits
(2) a terrible monster with only negative traits

2016-10-31

self-labeling

a tendency to assign various named categories/roles to oneself, both for personal and public use

2016-10-30

priming

influence of unrelated subconscious stimuli on thinking and behaviour

beyond simple effects on speed of recognition, most findings didn't replicate?

2016-10-30

Minimal Viable Product

MVP

an early version of a product that works well enough to solve someone's problem

2016-10-28

explicit disjunction

presenting an argument as a disjunction of several claims, instead of adding assumptions

this makes the assumptions more visible, and avoids some psychological problems

2016-10-28

status boost

status threat

primary reinforcers that measure one's place in the pecking order

just talking to a higher status person (e.g. your boss) is often enough to feel a status threat

2016-10-28

cultish gobbling

eagerly waiting for and consuming content from a particular source

this turns off critical thinking

2016-10-27

self-banishment

being driven by the feeling of shame to break all contact with a social group

2016-10-26

autonomy

oppression

primary reinforcers that measure one's ability to decide and influence outcomes

2016-10-26

generalizing effect

the tendency of an aroused brain to notice connections that aren't really there

e.g. after the threat response is activated, everything starts to look more dangerous

2016-10-26

primary reinforcer

a reward or threat implemented in the brain at a low level, which activates even if consciously anticipated or planned

strong influence on physical health and longevity

2016-10-26

fairness

unfairness

primary reinforcers that track outcomes of negotiations

this implements unconditional retaliation for cheating or taking too much

2016-10-26

dark matter people

people from other social tribes, with whom one is very unlikely to interact even if they are physically near

the sorting happens implicitly across a broad range of preferences and activities

2016-10-26

outgroup

fargroup

a social group that is:

(1) close but with small differences, which inspire intense hatred and conflicts

(2) distant enough to avoid the friend/foe classification, often seen as exotic or funny

2016-10-24

log-normal distribution

a distribution of a variable X such that $Y=\ln(X)$ is normally distributed

2016-10-24

hostile media effect

a tendency to see media coverage as biased against one's own side

2016-10-24

friend or foe

human brains automatically classify other people as friends (empathy + oxytocin) or foes (no empathy + threat response)

the default response to strangers is slightly negative, and needs some work to change

2016-10-24

spreading activation

getting more brain regions to notice and process a thought

e.g. saying aloud or writing down improves recall and ability to apply it in other situations

2016-10-24

threat response

arousal of the limbic system, representing a hardwired negative reaction to a primary reinforcer

2016-10-24

oxytocin

a neurochemical responsible for the pleasurable feeling of connecting to other people

released when hugging, dancing together, understanding each other through conversation etc.

2016-10-24

Betteridge's Law of Headlines

every headline that ends with a question mark can be answered with the word "no"

2016-10-24

ruminative jam

a very strong repeated request from the monkey brain to solve some problem, which makes it impossible to focus on anything else

possible not only for social, but also for technical problems

2016-10-22

reapprisal

changing one's interpretation of experienced events

2016-10-22

placebo autonomy

feeling more in control after making some trivial choices

2016-10-21

marginal revenue

the additional revenue that will be generated by increasing sales by one unit

in a perfectly competitive market, equal to the price of one unit
in a monopoly, less than that (the company needs to reduce prices to sell more)

2016-10-21

elasticity

the ratio of percentage change in one variable, to percentage change in another variable

$$x \nabla \log (f(x))$$

2016-10-21

acausal contract

an agreement that was never negotiated, but resulted from both parties' decision theories

2016-10-21

infinite debt

a claim that someone owes their existence to society/parents etc., so they have no right to refuse them anything

this would be equivalent to an acausal contract of slavery

2016-10-21

Ra

a tendency to approve/trust/defer to abstract and vague powers

2016-10-21

applause light

a statement that doesn't convey useful information, but rather hopes to play on positive associations of some words

2016-10-21

empathy

compassion

understanding someone by:

(1) mirroring their emotions, including negative ones

(2) modeling their situation and having a positive feeling of connection

2016-10-20

attentional blink

not seeing or hearing anything for a short moment after attention goes to some internal thought

2016-10-19

virtue of silence

preventing harm by not talking about a particular topic (including whether to talk about it or not)

"you aren't stuck in traffic, you are traffic"

2016-10-19

Streisand effect

an attempt to hide or censor information might ultimately result in publicizing it more widely

2016-10-19

bravery debate

arguing about who is the persecuted underdog, and who has it easy

<div align="center">one of the most mind killing topics</div>

2016-10-19

Crocker's rules

giving others permission to optimize their messages for information, even when it hurts your feelings

2016-10-18

hot button

a pattern previously associated with negative emotions, which activates them in similar situations

2016-10-18

ångström[/ˈɔŋstrøm/]

<div align="center">unit of length equal to 10^{-10} m</div>

2016-10-17

functional MRI

fMRI

using MRI to detect changes in blood flow in the brain, which are coupled to neuronal activation

2016-10-17

frontal lobe

<div align="center">the front region of the cerebral cortex</div>

2016-10-17

cerebral cortex

<div align="center">the outer layer of neural tissue in brains of mammals</div>

2016-10-17

temporal lobe

a part of the cerebral cortex that is on the side, around one of the ears

2016-10-17

internalized hate

negative feelings towards one's own type of social nonconformity

2016-10-17

rabbit hole

pursuing a single topic deeper and deeper, while losing track of its relative importance

2016-10-17

mental impasse

not having any new useful thoughts, and not knowing how to have them

> the brain easily gets stuck on a few recent thought patterns or memories

2016-10-16

room for funding

how much more money an organisation is prepared to put to effective use

2016-10-16

replacability argument

if you don't do X, someone else would do it instead, so you are not adding value by doing X

> this neglects that "someone else" would be free to do other things

2016-10-15

anterior temporal lobe

ATL

the front part of a temporal lobe

2016-10-15

prefrontal robbery

activating other large regions of the brain, which tends to reduce activity in the prefrontal cortex

(e.g. by focusing on sounds, body movement etc.)

2016-10-15

motor cortex

the rear part of the frontal lobe, responsible for motor control

2016-10-15

norepinephrine

noradrenaline

a neurotransmitter that represents the level of alertness and fear

2016-10-15

magnetic resonance imaging

MRI

obtaining images from radio waves, generated by hydrogen atoms in the body when they are in a strong magnetic field

2016-10-15

electroencephalography

EEG

measuring electrical activity of the brain using electrodes

usually placed along the scalp, but also sometimes implanted

2016-10-15

under-arousal

over-arousal

a level of activity in the prefrontal cortex that is too low/high for optimal performance

regulated by levels of dopamine and norepinephrine

2016-10-15

energetic loop

increased arousal and energy to continue an action

2016-10-15

emotional loop

using emotions to make a decision, that in turn influences the same emotions

this creates large emotional swings and an internal race condition

2016-10-14

holding cost

the mental energy needed to keep things in attention or short-term memory

2016-10-14

default network

the brain regions that activate when not focusing on any task, responsible for mind-wandering, rumination, and self-reflection

2016-10-14

internal distraction

external distraction

an unrelated stimulus that prompts a change in attention, coming from:
(1) one's own thoughts
(2) external world

> both require to pay an inhibition cost to stay focused

2016-10-14

inhibition cost

the mental energy needed to prevent a thought from entering attention

2016-10-14

loading cost

the mental energy needed to load objects into short-term memory

2016-10-14

partial attention

doing something, while monitoring the environment for other things that could be important

> this depletes energy and reduces effective intelligence

2016-10-14

allostatic load

long-term damage to the body from repeated stress

2016-10-14

bulldozer reform

a policy that advocates tearing down the existing system, without seeing that building a new one is most of the work

2016-10-14

homeostasis

tracking

keeping a parameter:
(1) constant regardless of the situation
(2) responsive to some property of the external world

2016-10-14

prioritize prioritizing

prioritizing is one of the most mentally exhausting tasks, so it might be impossible to do later

2016-10-14

system-level preference

a preference that is above emotions and relationships, but not by neglecting them, but rather recognizing that they are reactions to a world defined on a higher level

2016-10-13

dual-task interference

the significant drop in performance when thinking consciously about more than one task at a time

2016-10-13

mental energy curve

gradually decreasing capacity for focus and energy-intensive thinking, as exhaustion grows throughout the day

2016-10-12

complete the pattern

a solution suggested by the brain instantly and automatically, based on cached patterns

thought patterns are easily acquired from unreliable sources and passed through without checking

2016-10-12

pretending to be wise

trying to signal wisdom by refusing to take sides or do anything at all

2016-10-12

special snowflake trick

the brain mechanism that promotes thinking that one is special, unique, or has something others don't have

2016-10-12

sexual esteem

an estimate of one's own attractiveness as a sexual partner, that regulates the choice of reproductive strategies

2016-10-12

contrarianism

meta-contrarianism

holding opinions opposite to:
(1) those widely accepted
(2) those held by contrarians

2016-10-12

theater level

deadly serious level

understanding a sphere of public activity as:

(1) social games, in which real expertise doesn't exist

(2) having serious content hidden in between all the pretend-serious stuff

2016-10-11

affective forecasting

predictions based on current feelings, rather than data about the future

2016-10-11

basal []ganglia[/ˈɡaŋɡlɪə/]

a few parts of the brain located at the base of the forebrain in vertebrates, responsible for executing routines

also for eye movements and more

2016-10-11

prefrontal cortex

PFC

the front part of the frontal lobe, responsible for conscious thought

2016-10-09

risk ratchet

increasing risk-taking from fear of falling behind

2016-10-09

person-affecting perspective

impersonal perspective

morality based on:
(1) what happens to currently existing people
(2) considering all lives that could potentially exist

2016-10-09

second-guessing move

instead of going straight for the desired result, trying to guess the reactions of others and manipulate them

2016-10-09

disposition spheres

attitude towards:
- self

- other people
- the world

2016-10-09

time barrier

level barrier

a problem that requires mostly:
(1) experience and gradual refinement
(2) new insights and breakthroughs

2016-10-08

technology coupling

a tendency of some technologies to be developed together

e.g. one is a precursor or an irresistible application of the other

2016-10-08

state risk

step risk

a risk from some event that:
(1) is proportional to its duration
(2) depends only on whether it happened or not

2016-10-07

idle thoughts

the topic of one's thoughts when attention is free to wander,
which serves as an indicator of the motivation system

2016-10-07

growth mindset

fixed mindset

thinking that it's:
(1) possible to change, and success depends on effort

(2) not possible to change, and success depends on inborn qualities

2016-10-06

unfamiliar decision context

making the first decision of a kind, which makes it particularly likely to go wrong

2016-10-06

total utilitarianism

average utilitarianism

counting the utility of a group as a sum/average of the utilities of its members

for this to be well defined, "utility" must refer to some particular metric with stable reference points

2016-10-06

quantum immortality

the subjective experience of surviving quantum suicide regardless of the odds

2016-10-06

utility function

Von Neumann–Morgenstern utility

a function that maps outcomes into the reals, that necessarily exists for each sensibly acting decision maker, such that his choices are equivalent to maximizing EV of that function

2016-10-06

resource balance principle

if allocation of resources is close to optimal, than marginal benefits of investing in anything are about equal

this also applies to internal allocation of thinking, effort etc.

2016-10-04

illuminance

luminous power per unit area

[$lx = \frac cd·sr m^2$]

2016-10-04

luminous intensity

luminous power per unit solid angle

[cd]

2016-10-04

solid angle

$\frac{A}{r^{2}}$ [$steradian$] for area A subtended on a sphere

full sphere: $4\pi\ sr$

2016-10-03

superorganism

multiple agents acting purely for the benefit of the whole group

e.g. cells in a human body, some insect colonies

2016-10-03

technological diffusion

the gradual speading of technological progress to everyone who is behind

2016-10-03

luminous power

lumen

radiant power weighted by frequency to approximate human eye sensitivity to light [$lm=cd·sr$]

the peak of the luminosity function is at 555 nm (green light)

2016-10-03

radiant power

radiant flux

radiant energy per unit time

[$W = \frac J s$]

2016-10-02

Malthusian[/mal'θjuːzɪən/] condition

most people living on subsistence level income, that just barely allows them to raise two children to maturity

occurs whenever population growth is limited by economic growth (i.e. in most of human history)

2016-10-02

attribution to messenger

associating the positive or negative affect of a message with the person who says it

2016-10-01

risk emission

a contribution to total mass of failure in all possible branches, including those that didn't happen

2016-10-01

anthropic capture

using indexical uncertainty about being in a simulation to make an AI cooperate

2016-10-01

indirect normativity

specifying norms by some process that will give them as its

output

2016-10-01

Bonferroni correction

requiring a significance level α / m for each hypothesis tested in an experiment that tests $m \gt 1$ of them

2016-10-01

p-value

the probability of obtaining a result equal to or more extreme than what was actually observed, assuming that the effect is not real

= minimum α level at which the result would be considered significant

2016-10-01

α level

β level

the probability of:
(1) detecting an effect, assuming it is not real
(2) not detecting an effect, assuming it is real

2016-09-30

color rendering index

CRI

the ability of a light source to reveal colors faithfully (in comparison to natural light)

2016-09-30

incandescence

the emission of electromagnetic radiation by a hot body

2016-09-30

seasonal affective disorder

SAD

a change of mood to more depressive in winter

2016-09-30

infrastructure profusion

almost any goal implemented by a superintelligent singleton will use most of the observable universe as resources

2016-09-30

paperclipper

an AI with a final goal to maximize production, which after becoming superintelligent transforms much of the observable universe into copies of some product

2016-09-30

wireheading

taking control of one's own reward system

2016-09-30

safety ritual

behaviour that signals being responsible and cautious, without really reducing the risk of failure

2016-09-30

von Neumann probe

a self-replicating colonization spaceship

2016-09-30

functional soup

a process that preserves its terminal goals, while being very fluid about all other aspects of its existence

2016-09-30

goal integrity

an instrumental goal for most final goals is to preserve the exact content of the final goal

2016-09-30

orthogonality thesis

almost any level of intelligence can be combined with almost any goal

there are some exceptions - e.g. very low intelligence might be insufficient to represent the goal

2016-09-30

instrumental convergence

for a wide range of terminal goals, short term strategies tend to be similar until a certain amount of power is secured

2016-09-30

unstable beam

a range of values that do not yet determine success or failure, but a significant move up or down will start a cascade in that direction

2016-09-30

humor the baby

going along with someone who is not mature enough to handle reality

2016-09-30

mind-body link

a bidirectional influence between mental states and body reactions

e.g. feeling happier after smiling, subconscious toward/away body language

2016-09-29

momentum

the product of the mass and velocity of an object [$\frac{kg \cdot m}{s}$]

2016-09-29

free energy

the amount of work that a system can perform without using external resources

2016-09-29

exothermic reaction

endothermic reaction

a chemical reaction that:
(1) releases energy (by light or heat)
(2) absorbs energy from outside

2016-09-29

Sabatier[/sabatje/] reaction

$CO_2 + 4H_2 → CH_4 + 2H_{2}O +$ energy
in presence of a catalyst (e.g. nickel $_{28}Ni$) and at high temperature (300–400 °C)

CO_2 - carbon dioxide
H_2 - hydrogen
CH_4 - methane

2016-09-29

electrolysis[/ˌɛlɛkˈtrɒlɪsɪs/] of water
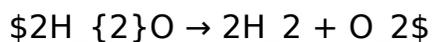
$2H_{2}O → 2H_2 + O_2$
due to an electric current being passed through the water

H_2 - hydrogen (gas)
O_2 - oxygen (gas)

2016-09-29

impulse (physics)

change in momentum

$$N·s = \frac{kg·m}{s}$$

2016-09-29

specific impulse

impulse generated per unit of propellant

units of velocity if mass is used to measure propellant, or of time if weight is used for propellant

2016-09-28

guilt management

making sure one's actions don't inspire excessive guilt in other people

2016-09-28

risk budget

a limit on total allowed risk emission

can be transferred to another agent, after spending some on uncertainty about their ability to keep the contract

2016-09-28

endorsed motivation

a motivation that is supported by quality reasoning about higher values and consistency

2016-09-28

Do-The-Best-Thing

a strategy based on choosing some simple and clearly endorsable metric, and optimizing only for it

2016-09-28

singleton

a single decision-making entity with uncontested power over the whole world

2016-09-28

decisive strategic advantage

an advantage big enough to dominate the world and prevent others from catching up

2016-09-28

content overhang

a newly developed AI can use a large pool of untapped pre-made content (e.g. the Internet)

2016-09-28

hardware overhang

a newly developed AI can use a large pool of untapped hardware to run at great speeds or in many copies

2016-09-28

low-hanging fruit

quick and easy gains on new territory

2016-09-28

selection power

the total budget of negentropy, that can be allocated to optimize various parameters

2016-09-28

relatedness

loneliness

primary reinforcers that measure being accepted, understood, and belonging in a group

transparent thought

a mental step so quick and natural that it is not noticeable from the inside

twinge of starting

the mental pain of deciding to start a task now

temporal difference learning

backpropagating prediction errors to the chain of past predictions, discounting by time

supervised learning

unsupervised learning

using an algorithm to:

(1) generalize from a set of examples with provided correct answers

(2) find structure in unlabeled data

chilling effect

discouragement of a legitimate activity by a subtle fear of unlikely sanctions or trouble

takeoff (AI)

the progress from reaching the intelligence of a single adult human, to reaching strong superintelligence

crossover point (AI)

a point beyond which the improvement of an AI is driven more by its own actions than external work

2016-09-26

quality superintelligence

a system working on similar time scales as a human brain, but overwhelmingly smarter

2016-09-26

collective superintelligence

a system composed of many human-level intellects, which vastly outperforms humanity in most domains of interest

2016-09-26

speed superintelligence

a system with capabilities similar to a human mind, but many orders of magnitude faster

2016-09-26

victory dance

reinforcing the feeling of satisfaction by energetic body movements

2016-09-25

toward response

away response

reactions that are suggestive of:

(1) approaching, safety, curiosity, pulling

(2) distance, caution, rejection, pushing

applies to both mental states and body language (e.g. someone leans forward or back)

2016-09-25

seed AI

an AI which is limited in most areas, but is able to improve its own architecture

2016-09-24

dog relationship

cat relationship

partners who:

(1) depend on each other for their well-being

(2) would be fine by themselves, but choose to be together at the time

> in (1) partners negotiate autonomy, while in (2) they negotiate closeness

2016-09-22

prisoner's dilemma

two prisoners decide separately to defect or cooperate with the other:

if both cooperate, they serve 1 year each

if both defect, 2 years each

if only one defects, he goes free and the other one serves 3 years

2016-09-22

third story

a description of events as if told by an impartial observer, that both parties can agree with

2016-09-22

trivial inconvenience

a small obstacle, which disproportionately hinders action under human tendency to conserve certain kinds of effort

2016-09-20

commitment bet

a one-sided promise to pay money if some condition is not met, which only makes sense as a voluntary self-commitment mechanism

2016-09-19

looking for noise

collecting inspiration to use as fuel for thinking

2016-09-19

happy death spiral

believing something is so wonderful, that a halo effect triggers and one finds even more wonderful things about it

2016-09-19

evaporative cooling

the average temperature drops as high energy outliers leave

possibly a factor in social groups becoming more extreme

2016-09-19

closed ideology

self-correcting ideology

an ideology that:

(1) doesn't include any paths to learn more or challenge its assumptions

(2) promotes improving the accuracy of all beliefs

2016-09-19

apophenia

a tendency to recognize meaningful patterns in random data

2016-09-19

leaky thoughts

ideas that someone thought about a lot influence their work,

even if they are not aware of it

2016-09-18

expectation trick

the brain mechanism that urges action by anticipated hypothetical happiness

2016-09-18

self-depreciation trick

the brain mechanism that maintains compartmentalized feelings of inferiority regardless of reality

2016-09-18

Frenzel maneuver

blocking the nose and moving the back of the tongue in a way that pushes air into Eustachian tubes

2016-09-18

Valsalva maneuver

blowing against a pinched nose

2016-09-18

awkward silencing

staying silent for a long time to pressure someone to speak

a powerful tool in negotiations and interrogations

2016-09-17

Henry's law

the amount of dissolved gas is proportional to its partial pressure in the gas phase

2016-09-17

partial pressure

the pressure that would be exerted by one of the gases in a mixture if it occupied the same volume on its own

2016-09-15

typical mind fallacy

assuming that other people's minds work like yours

e.g. the mental imagery debate, or not noticing you are color blind

2016-09-15

conceptual confluence

ease of thinking and remembering when thoughts are naturally aligned with a structure given by abstract concepts

2016-09-14

ego depletion

the idea that willpower is limited and can be used up

can be more simply explained with overall mental exhaustion?

2016-09-14

willful inconsistency

convincing your intuition of something which you know to be false, to get a useful reaction

2016-09-14

compartmentalization

keeping conflicting thoughts isolated from each other

2016-09-14

phantom akrasia

using ego depletion or akrasia as excuses to stop working

2016-09-14

meta habit

a daily habit of tracking habits

2016-09-14

abuse of pattern-matching

purposefully entering a state in which the brain's top-down pattern matching is overactive, producing hallucinations, feelings of all-understanding etc.

2016-09-14

agnosia[/əgˈnəʊsɪə/]

the inability to process sensory information of a particular type (while everything works normally outside of the brain)

2016-09-14

anosognosia

a delusion in which someone is unaware of their own disability

involves loss of proprioception

2016-09-14

monothematic delusions

polythematic delusions

one person having delusions about:
(1) a single topic
(2) a range of unrelated topics

2016-09-14

sanity filter

the ability to recognize and prevent saying something that would sound delusional, even when internal experiences support it

2016-09-14

persecutory delusions

a delusion that is one being tormented, followed, tricked, spied on etc.

anxiety + Super Base Rate Fallacy

2016-09-14

delusions of grandeur

a delusion of being some famous, impressive person

miscalibrated sociometer + Super Base Rate Fallacy

2016-09-14

Cotard delusion

a delusion that one is dead or does not exist

loss of body sensations + Super Base Rate Fallacy

2016-09-14

Fregoli delusion

a delusion that random strangers are actually friends or family members in disguise

2016-09-14

Capgras delusion

a delusion that someone close to you has been replaced by an identical-looking impostor

damaged emotional response to faces + Super Base Rate Fallacy

2016-09-14

hypocrisy

publicly endorsing rules or moral values which one doesn't follow

2016-09-14

Super Base Rate Fallacy

a pathological inability to penalize a priori unlikely theories

delusional beliefs are preferred whenever they fit all data

2016-09-14

sociometer

a feeling that measures one's acceptance in society, evolved to minimize chances of rejection

2016-09-14

mind projection fallacy

assuming that the way you see something is how it really is

2016-09-14

perceptual handshake

sensory data being matched against the brain's top-down predictive models

matched data feels like it "makes sense" and "snaps into place"
unmatched data is flagged as surprising and important

2016-09-13

small sample

a sample that has a small number of data points, so any data derived from it will be mostly noise

2016-09-13

12 second rule

when a question is asked in a group, wait at least 12 seconds to let everyone think

2016-09-13

making smaller circles

optimizing a skill by learning the feeling of getting it right, and then gradually reducing movements while maintaining the feeling

2016-09-13

broccoli error

using not wanting to justify not wanting to want

2016-09-13

screening off

making something no longer contribute new information if it became known

2016-09-10

creative synergy

each member of the team does higher quality work than they could alone

2016-09-09

commitment to emotions

a promise that depends on being in certain emotional states

this is faulty and creates internal race conditions

2016-09-08

stress and recovery

performance depends equally on work and rest periods

2016-09-08

soft zone

a vulnerable period of growth and integrating new knowledge

2016-09-08

numbers leave numbers

 integrating the essence of technical knowledge into intuition

2016-09-08

mistake spiral

a sequence of mistakes that starts with being disturbed with the first one and not updating

2016-09-05

pica syndrome

doing a variety of things to satisfy strange urges, while those urges are just a confused signal about some deeper problem

2016-09-05

learned helplessness

learned blankness

getting used to a lack of:
(1) agency in some area
(2) knowledge of some topic

2016-09-04

superstimulus

something engineered to get much stronger reactions from the reward system than what was ever possible in EEA

2016-09-04

social circuits

the specialized ability of human brains to process, understand and remember data related to people and social contexts (including self-knowledge)

2016-09-04

leverage

force multiplication

the ratio of impact to invested effort

2016-09-04

fixed cost

the initial cost that doesn't depend on production size

2016-09-03

EDT feedback

the brain using body status and visible actions as evidence about its own judgement

2016-09-03

clean pain

dirty pain

being miserable:
(1) as a direct reaction to some painful event
(2) from soured thinking about a painful event

2016-09-02

I/you ratio

how much time is spent talking about self, versus about the audience

2016-09-01

irritants

sounds or movements that do not add to the message, and waste the audience's attention

2016-08-31

Ben Franklin Effect

someone liking you more after they have done you a favour

2016-08-31

personal space

the space around a person that they feel ownership of

confident people tend to take up more space, and express it through body language

2016-08-31

mirroring

imitating someone's body language as a persuasion technique

e.g. a movement of someone's left hand is mirrored using the right hand

2016-08-31

limbic resonance

the tendency of people to mimic and learn body language from others

2016-08-31

emotional contagion

automatic transfer of emotions to other people

esp. leaders and other charismatic people can have a strong "ripple effect"

2016-08-31

speaking in pictures

a persuasion technique based on using language that is very easy to visualize

2016-08-30

visionary charisma

charisma based on passionate belief in a noble cause

makes people believe you and feel inspired

2016-08-30

focus charisma

charisma based on being fully present and taking everyone seriously

> makes people feel understood and listened to

2016-08-30

kindness charisma

charisma based on caring and goodwill towards others

> makes people like you and feel accepted

2016-08-30

authority charisma

charisma based on appearance of power

> makes people obey and feel intimidated

2016-08-30

funding certificate

an unique certificate of having funded a project, purchased from the person who originally funded it

> possible after the project is already finished

2016-08-30

gambler's ruin

a gambler with finite wealth will eventually go broke when playing with EV = 0

the same if EV is positive but he is increasing his bets when he wins and never decreasing them

2016-08-29

physical empathy

an immediate sense of someone's affect from low level body signals

2016-08-28

near mode

far mode

thinking based on:
(1) experienced events
(2) abstract mental constructions

they bring to mind clusters of related aspects (concrete/abstract, touch/sight, warm/cold, strong/weak emotion, low/high status etc.)

2016-08-28

appeal to emotion

an argument that depends on the emotional reaction of the listener

2016-08-28

anecdotal evidence

stories about a small number of cases, used in place of strong evidence

2016-08-27

Parkinson's Law

work expands to fill the time available for its completion

2016-08-27

negativity bias

noticing more easily those things that are suggestive of danger, attack, criticism etc.

2016-08-27

inner critic

private thoughts that criticize one's own actions, often in verbal form

2016-08-26

placebo effect

nocebo effect

a positive/negative effect on the body just from believing that something is helpful/harmful

apparently, real changes are mostly in perception of pain and in self-deception

2016-08-25

deep skill

shallow skill

a skill that:

(1) requires specific training and building habits that directly control behaviour

(2) can be learned quickly by tapping into existing skills

2016-08-25

cognitive band-aid

a cognitive technique which is useful at some point, but quickly becomes obsolete after progress in other areas

2016-08-25

consumer surplus

producer surplus

the difference between:

(1) the maximum price at which consumers would be willing to buy and the market price

(2) the market price and the minimum price at which producers would be willing to sell

2016-08-25

comparative advantage

the ability to produce a particular good at a low opportunity cost, so that it's cheaper to trade for other goods rather than produce them directly

2016-08-25

economic moat

something that protects long-term profits and market share from competitors

2016-08-24

law of opposite fetishes

for each group of people with an unusual trait or preference, there is another group who finds them attractive

dating success depends on which group is larger

2016-08-24

weak social ties

connections between people who are not friends, but have friends or social groups in common

2016-08-23

winner's curse

the highest bidder in an auction is an outlier, and it is likely that he overpaid

2016-08-23

unilateralist's curse

if a matter is decided when at least one agent acts independently, action will be taken too often by noise alone

2016-08-21

style under stress

a particular person's typical unhealthy behaviour in conversations

2016-08-19

detached qualifier

a term such as "good", "bad", "better" etc. used without specifying a purpose

e.g. "what is the best programming language"

2016-08-19

keystone habit

a single successful habit that disrupts the status quo and makes other changes possible

2016-08-18

over-empathizing

projecting discomfort onto someone without checking if they really feel it

2016-08-17

rule of five

there is a ~94% chance that the median of a population lies between the smallest and largest value in a random sample of five

2016-08-16

negentropy lens

seeing what keeps things from reverting to the high-entropy default

2016-08-15

upward spiral

downward spiral

a sequence of successes/failures, which influence confidence and motivation, resulting in even bigger positive/negative results

2016-08-15

secret sauce

something that is crucial to success, but easy to overlook when imitating someone else

2016-08-14

craving

strongly wanting a reward after a cue of an ingrained habit is activated

2016-08-14

habit loop

cue, routine, reward

possible cues: a time, a place, an emotional state, a person, the previous action

2016-08-14

barycenter

the center of mass of bodies that are orbiting each other

2016-08-14

Lagrange points

the 5 points in which the combined gravitational pull of two orbiting masses provides exactly the centripedal force required to orbit with them

2016-08-14

organisational habit

the way an organisation deals with some task, deeply embedded in procedures and people's patterns of behaviour

2016-08-13

path to action

1. See.
2. Tell a story.
3. Feel emotion.
4. Act.

2016-08-13

trust of motive

trust of ability

trusting someone to:
(1) have intentions aligned with yours
(2) achieve the outcome which they choose

2016-08-13

hunt mode

going out and investing energy until a specific mission is accomplished

2016-08-13

reverse osmosis

purifying water by forcing it through a semipermeable membrane

needs enough pressure to overcome the opposing osmotic pressure

2016-08-13

osmosis

the movement of solvent across a semipermeable membrane, in the direction that equalizes the concentration of solute

happens when the membrane blocks the solute, but allows the solvent to pass

2016-08-13

initial mass function

a distribution of initial masses for a population of stars

"initial" means at the moment when the star begins hydrogen fusion

2016-08-12

tact filter

applying tact to everything one is saying

2016-08-11

violent agreement

a fight between people who agree about all important issues

2016-08-11

priming the pump

investing a small amount to start a process that generates much more

2016-08-11

man cave

a room exempt from female influence, used for hobbies etc.

2016-08-11

nudging

influencing people by making something the path of least resistance

2016-08-11

crystallized intelligence

a reflection of intelligence in objects which it had shaped

2016-08-10

silence or violence

unhealthy behaviour in conversations takes the form of either refusing to communicate, or forcing meaning on others

2016-08-10

double standard

judging others more harshly than yourself

2016-08-09

normalization of abuse

reframing something harmful as typical and expected, to reduce reactions of the victims

2016-08-07

Overton window

the range of ideas that the public can currently accept

not too far from what is already popular or well known
the window can be slowly adjusted by promoting fringe ideas

2016-08-07

jungle bias

expecting scarce resources and zero-sum games

2016-08-07

charitable reading

skipping over minor mistakes, ambiguity etc., to the extent that the author would have agreed

2016-08-07

straw man

steel man

a modification of an argument against one's position, which is:

(1) unfairly dumbed down, to make it easy to defeat

(2) stronger than the original version

2016-08-07

ad hominem

an argument directed against the person, rather than what they said

2016-08-07

playful mode

having fun while curiously exploring something

this is how learning feels from the inside

2016-08-07

bullshit-minimizing

spending the minimum amount of time on things that you don't value long-term

2016-08-07

sound absorber

a person who is nice but doesn't stimulate the discussion, lowering the overall energy

2016-08-07

impostor syndrome

a tendency of some high-achieving individuals to feel like they are frauds

2016-08-07

Hamming questions

what is the most important problem?
what are you working on?
why are they not the same thing?

2016-08-07

reference incident

a situation used as an example, in order to communicate more clearly about a general pattern

2016-08-05

shared pool of meaning

the thoughts shared and understood by all participants in a conversation, regardless of their validity or usefulness

2016-08-05

cherry picking

selectively showing evidence that supports one side

2016-08-05

false dichotomy

a choice between two options, when other answers could be found

2016-08-05

middle ground fallacy

argument to moderation

assuming that the truth must be a compromise

2016-08-04

fool's choice

a choice between staying silent or offending someone

instead, the conversation can be handled skillfully

2016-08-04

dogfooding

using the same product that you sell to customers

2016-08-04

representativeness heuristic

judging probability of an event by how characteristic it is of
the situation

2016-08-04

anchoring

an estimate is affected by whatever gets used as the starting
point (even if it was a nonsense number)

2016-08-04

intensity matching

intuitively translating a value between scales, which are not
comparable

2016-08-02

internal confidence level

a confidence level calculated within a model, not corrected
for the possibility that the whole approach is flawed

2016-08-02

inner Google

the mechanism that offers verbal answers to questions asked
internally to oneself

top results are often just "complete the pattern"

2016-08-01

hard accountability

soft accountability

having results checked by:

(1) verifying a clear requirement, with no possibility of
excuses

(2) doing a review, with some flexibility depending on the situation

2016-08-01

economy of variety

optimizing the process of innovation and capturing value from it within one entity

2016-08-01

economy of scale

economy of scope

optimizing:

(1) unit production cost, by accumulating experience and technology

(2) market position and the choice of scale vectors

2016-08-01

Grabowski ratio

a ratio of resources spent on market research to those spent on technology

> in a new market, most resources are spent on technology and scaling
> in a competitive market, successful firms have a ratio around 1:1

2016-08-01

experience curve

a plot of unit production cost, versus cumulative volume of production

> typically, doubling total experience reduces unit production cost by 10-25%

2016-08-01

bathtub curve

a plot of failure rate over time with high "infant mortality", low medium section and high late wearing-out failures

2016-08-01

unique-snowflaking

having diverse and unique components in a system, sacrificing efficiency for open-ended learning

2016-08-01

lean design

fat design

a design which is:

(1) optimized but fragile

(2) inefficient but allowing mistakes and learning

which areas encourage learning can be controlled by adding slack everywhere else

2016-08-01

bottleneck

a single part of a system that limits its total performance

2016-08-01

leak before fail

a design principle to have some weaker components in a system, allowing it to fail gradually

necessarily, it also means inefficiency from over-design of some parts

2016-08-01

critical path

a sequence of dependencies such that delaying any of them would delay the whole project

2016-08-01

Arrow information paradox

when a buyer of intellectual property knows enough to make an informed choice, they might know so much that buying doesn't make sense

2016-08-01

paradox of disclosure

admitting that your opinion might be biased causes people to trust it more

2016-08-01

specialty bias

doctors are more likely to recommend treatment from their specialty

e.g. surgeons recommend surgery more often etc.

2016-07-31

control illusion

an illusion that since one is aware of making plans, one's behaviour is under conscious control

2016-07-30

gambler's fallacy

expecting a higher chance of such random draws that would make the sequence of draws look more representative

2016-07-28

insight porn

enjoying the feeling of having new exciting ideas, without making real progress

2016-07-28

universal culture

a high entropy culture that allows all memes to spread, and accepts whatever wins

enforces only basic rules for peaceful coexistence
it's not anyone's culture, though it is often mistaken for a "Western" culture

2016-07-28

process lens

seeing how something happens in time, and what goes in and out of it

2016-07-28

effectuation

planning from what you have and what is possible, rather than goals

2016-07-27

walled garden

long-term quality of an organisation is as high as its barriers to entry

2016-07-26

neglect of adaptation

estimating the influence of an event by imagining the initial reaction to it

> in reality, people get used to the change and stop thinking about it
> pressing poverty seems to be an exception

2016-07-26

light in the refrigerator

something that is present whenever you check, creating an illusion it's always there

2016-07-26

broken window

a loss of a resource that is also necessarily a loss to the whole economy

> the opposite is often cleverly argued (more jobs for people who repair it etc.)

2016-07-26

reverse lottery

getting a certain payout for a small risk of a huge loss

2016-07-26

red team

an outside team formed to challenge an organisation and find flaws

2016-07-25

satisficing strategy

maximizing strategy

(1) accepting the first solution that is good enough
(2) checking all possibilities to choose the best one

2016-07-22

mood heuristic

using the current mood as a proxy for difficult questions about life

2016-07-22

storyline satisfaction

experienced happiness

self-reports based on:
(1) thinking about one's life in general
(2) samples of happiness-in-the-moment

2016-07-22

flow

a state of total absorption in a task

2016-07-22

memory-optimizing

acting in a way that ensures the best future memories, rather than the best experience of events

> e.g. going on a difficult journey with a satisfying end

2016-07-22

average heuristic

intuition represents sets by averages

e.g. adding a broken item to a package can make the whole deal seem less valuable

2016-07-22

duration neglect

the memory of an experience doesn't depend on the duration

2016-07-22

peak-end rule

the memory of an experience depends mostly on the most intense moment and the ending

2016-07-22

proportion bias

intuitive judgement uses the ratio of two values, even when something else makes sense

> e.g. save 400 birds instead of 200, vs save 1800 instead of 1500

2016-07-22

confirmation bias

interpreting evidence in a way that tends to confirm existing beliefs

2016-07-22

availability heuristic

judgement of probability based on how easily examples come to mind

too much weight is given to recent or vivid memories

2016-07-22

local optimum

global optimum

a value that is optimal:
(1) in a neighbourhood
(2) in the entire domain

2016-07-22

Simpson's paradox

even if each part of data shows the same trend, an analysis of combined data can still show a reverse trend

(e.g. if women apply mostly to departments with low admission rates, in average results from all departments it looks like a bias against accepting women)

2016-07-22

Bayes'[/'beɪz/] theorem

$P(A|E) = \frac{P(E|A)P(A)}{P(E)}$

2016-07-22

confidence interval

a range of values calculated in such a way that, if the experiment were to be repeated, the actual value from the population would be in the range with a certain probability

2016-07-22

false negative

type II error

failing to detect an effect that is real

2016-07-22

false positive

type I error

detecting an effect that is not real

2016-07-22

Fermi estimate

a quick and rough order-of-magnitude estimate

2016-07-22

dimensional analysis

tracking "physical dimensions" (length, time etc.) throughout the calculation

2016-07-21

blind trial

double-blind trial

information about the trial is hidden from:
(1) participants
(2) both participants and researchers who conduct the trial

2016-07-21

observer effect

Hawthorne effect

(1) the act of observation changes the observed system
(2) participants of a study tend to behave in a different way than usual

2016-07-21

experimenter bias

the data gathered in an experiment is influenced by what the researcher expects

e.g. subconsciously communicating something to participants, confirmation bias when interpreting results

2016-07-21

response bias

non-response bias

a bias caused by:
(1) dishonest or inaccurate answers to surveys
(2) not everyone returning filled surveys

2016-07-21

bootstrapping (statistics)

generating samples by random sampling with replacement from a single actual sample of the population

variability of parameters calculated from such samples indicates variability between original samples of the population

2016-07-21

survivorship bias

a bias resulting from analysing data after some process removed a part of it

e.g. damage in returning bomber planes indicated where armour was **not** needed

2016-07-21

sampling bias

choosing a sample in a way that doesn't fairly represent the whole population

2016-07-21

no free lunch

you can't get something from nothing

2016-07-21

extended mind

the mind together with parts of the environment that help with cognition

> esp. those that are automatically trusted (not filtered on access)

2016-07-21

information rent

getting an advantage from additional information, which is not consumed in the transaction

2016-07-20

bait-and-switch

baiting a customer with an attractive offer but planning to persuade them to accept something else

2016-07-20

word-thinking

processing information based on words, rather than the concepts they refer to

> a common failure mode when background information is missing

2016-07-19

anthropic principle

any universe humans observe is necessarily compatible with their existence

2016-07-19

proxy variable

 an indicator of something that is difficult to measure directly

2016-07-19

Planck's principle

science advances one funeral at a time

a new scientific truth does not triumph when its opponents update, but rather when they eventually die

2016-07-19

sortition

appointing political officials by random draw from a pool of candidates

2016-07-19

critical mass

 the minimal amount that is enough for self-sustaining growth

2016-07-19

divergent thinking

convergent thinking

(1) generating many ideas, which go in various directions
(2) filtering and refining ideas to get a single answer

2016-07-19

Nash equilibrium

an assignment of strategies to players, such that no player benefits from changing his strategy while others don't

for any zero-sum game with finite numbers of players and pure strategies, there exists at least one (possibly in mixed strategies)

2016-07-19

zero-sum game

for any outcome, the sum of all players' payoffs is zero

(all constant sum games have an equivalent with this property)

2016-07-19

marginal rate of substitution

MRS

the rate at which someone is willing to give up one good in exchange for another, while maintaining the same level of utility

in a Pareto efficient allocation, it is equal for all agents

2016-07-19

Pareto efficiency

Pareto optimality

a state in which no parameter can be improved without making some other parameter worse

2016-07-19

opportunity cost

the value of the best alternative (which is not chosen)

2016-07-19

net present value

the sum of present values of all cash flows

each PV is exponentially discounted by time

2016-07-19

central limit theorem

a sum of many independent and identically distributed variables approximates a normal distribution

assuming finite EV and variance

2016-07-19

normal distribution

$d(x) = \frac 1 {√2\pi`} e ^ {-x^2 / 2}$
(for mean 0 and variance 1)

2016-07-19

unknown unknowns

factors that are outside of a model of risk, not even considered as possibilities

2016-07-19

black swan

a random, unpredictable event with major consequences

tends to be rationalized in hindsight, as if it "had to happen"

2016-07-19

premature optimization

trading off clean design for performance, before testing bottlenecks

2016-07-19

Pareto distribution

$P(X \gt v) = (\frac v\_min v) \char94 α$ for $v \gt v\_min$, otherwise 1

for some $α \gt 0$

2016-07-19

80/20 rule

Pareto principle

in many areas, around 80% of output comes from 20% of input

2016-07-19

catalyst

a substance that increases the rate of a reaction without being consumed in it

stock and flow

modeling a variable using its current total value and rates of change

feature selection

feature extraction

(1) selecting a subset of relevant variables to use in a model

(2) processing broadband input data to get high-level derived values

cognitive bias

a systematic error in judgement, caused by human-universal mental machinery

medical zebra

a diagnosis of an exotic medical condition when a simpler one is likely

"when you hear hoofbeats, think of horses not zebras"

overfitting

making a model too complex and too adjusted to the training set, while losing predictive power outside of the training set

conjunction fallacy

disjunction fallacy

a tendency to systematically assign probabilities that are:

(1) too high, for events that require multiple conditions to hold simultaneously

(2) too low, for events that occur when at least one of many conditions holds

2016-07-19

Occam's Razor

other things being equal, prefer the simpler explanation

2016-07-19

fundamental attribution error

explaining successes of other people by circumstance, and failures by their innate traits - while doing the opposite with one's own behaviour

2016-07-19

Hanlon's razor

don't attribute to malice that which is adequately explained by carelessness or stupidity

2016-07-18

werewolf problem

a problem harder than a threshold (specific to a person), above which they will hide it rather than seek help

2016-07-18

cobra effect

a situation in which an attempted solution actually makes the problem worse

e.g. bounty hunters started breeding cobras rather than killing them

2016-07-18

non-central fallacy

an argument based on assigning something to a category, of

which a central member evokes an emotional reaction

e.g. "abortion is murder!"

2016-07-17

reverse procrastination

doing small bits of work during a leisure activity

2016-07-17

Rapoport's First Rule

before criticism, express your target's position so clearly and fairly that they say "thanks, I wish I'd thought of putting it that way"

2016-07-17

double bind

asking a question about two choices, to direct attention away from other possibilities

2016-07-17

non-evaluable number

a number with nothing to compare it to is ignored by intuitive judgement

2016-07-17

between-subjects experiment

an experiment with multiple variants, of which each subject knows only about one

it can reveal effects that disappear when subjects compare alternatives

2016-07-16

defaulting

choosing the option that stands out the least when facing risk

unusual choices are harder to excuse

2016-07-16

narrow framing

asking a question in a way that makes the listener stuck in assumptions

2016-07-16

mental accounting

the urge to match costs with benefits, and close with positive balance

> a severe case of agentization
> source of the sunk cost fallacy

2016-07-16

mnemonic technique

representing data as some other type of data, which is more easily remembered by one's brain

> what is easy depends on the individual brain and areas of expertise

2016-07-16

earning to give

pursuing a high-paying career in order to donate to a cause

2016-07-16

exposure therapy

a treatment for anxiety or phobias by forced prolonged exposure, while safety is guaranteed

2016-07-16

brotherhood

an all-male social group with rules that prevent girl-related drama

2016-07-16

marginalism

an economic theory that explains prices by marginal cost and utility

2016-07-16

marginal cost

marginal utility

(1) cost of one additional unit of production
(2) utility of one additional unit of consumed resource

2016-07-16

diminishing returns

returns to scale

(1) decreasing marginal output, as input is increased more and more

(2) improved efficiency from processing a large amount in batch

2016-07-16

existential risk

x-risk

a threat to humanity's continued existence and prosperity

2016-07-16

societal collapse

a serious, large scale regression of society and civilization

2016-07-16

adaptive capacity

the ability of a system to adapt to a changing environment

2016-07-16

Law of Authority

a rule enforced by a single powerful entity

2016-07-16

Law of Society

a rule enforced by group consensus
defectors are punished by volunteers

2016-07-16

Law of Reality

a rule enforced automatically by the environment

2016-07-16

Nyquist frequency

half of the sampling rate of a digital signal

for every sinusoid with frequency above, there is another sinusoid ("alias")
with frequency below and identical digital samples

2016-07-15

denominator neglect

in "X out of Y", absolute value of X influences intuitive
judgement

2016-07-15

vivid probability

intuition treats unlikely events as important when they evoke
vivid mental imagery, and neglects them otherwise

2016-07-15

status illegibility

social groups are held together by maintaining vagueness of

relative statuses (except for alpha and omega)

2016-07-15

base rate neglect

using specific evidence as if it replaced priors

2016-07-15

Conservation of Expected Evidence

prior probability = EV of posterior probability

unlikely strong evidence is balanced by likely weak evidence in the opposite direction
absence of evidence is (weak) evidence of absence

2016-07-15

sunk cost fallacy

including resources that are impossible to recover in a decision what to do

2016-07-15

iterated hurt

being hurt by the knowledge that someone was willing to hurt you

works even if there was no actual hurt

2016-07-15

level reversal

level N+1 in development can be counterintuitive and superficially similar to level N-1

(that the next level is counterintuitive is what makes levels distinct in the first place)

2016-07-14

Dunnig-Kruger effect

unskilled persons overestimate their skill a lot

2016-07-14

Powertalk

a type of communication in which each conversation is a high-stakes trade in the game of power

> most content is expressed indirectly, with plausible deniability
> it only matters what you have and how well you bargain with it

2016-07-14

Peter principle

if promotion at work is based on performance in the current role, then the process stabilizes on reaching incompetence

2016-07-13

Chesterton Fence

a rule not to do reforms without understanding the reasoning behind the current state

2016-07-13

Schelling fence

an arbitrary boundary used to prevent a "slippery slope" situation

2016-07-13

Schelling point

a salient place, useful to reach agreement without communication

2016-07-13

signaling

maintaining a publicly visible trait which is evidence about some hidden personal variable

2016-07-13

Parfit's Hitchhiker

A driver asks you to promise to pay \$100 in the future, for saving your life now.

Both of you know your decision theory and all its logical consequences.

He can detect lying perfectly.

Can you commit to paying?

2016-07-13

Smoking []Lesion[/ˈliːʒ(ə)n/]

a thought experiment about an alternative world, in which smoking is correlated with cancer but does not cause it:

- there is a genetic lesion that increases chances of both

- you don't know if you have it

- you want to smoke, but you definitely don't want cancer

- do you smoke?

2016-07-13

Newcomb's paradox

Newcomb's problem

Omega prepared two boxes:

transparent box contains \$1000,

opaque box contains \$1,000,000 if Omega predicted you will take only the opaque box, otherwise nothing.

Do you take both boxes, or only the opaque box?

2016-07-13

counterfactual mugging

Omega tossed a fair coin to choose one of two outcomes:

(1) you lose \$100,

(2) you win \$10000 but only if you would have accepted the loss of \$100.

It is already known that you lost. Do you pay \$100?

2016-07-13

inclusive fitness

the total ability to pass genes to the next generations, including genes passed on by relatives

2016-07-13

adaptation-executer

an organism acts only to execute evolved adaptations, and it doesn't care about maximizing inclusive fitness as a goal

2016-07-13

behaviourism

explaining human behaviour by classical conditioning

2016-07-13

lost purpose

pursuing an instrumental goal that is no longer aligned with terminal goals

2016-07-13

hyperbolic discounting

$v(t) = \frac 1 {1 + Ct}$
an approximation of how intuition discounts rewards by time

2016-07-13

honing mode

jamming mode

a conversation mode based on:
(1) converging on a single idea
(2) exploring new ideas

main action: correcting/riffing

2016-07-13

mutual knowledge

something that everyone knows, and also it is mutual knowledge that everyone knows

2016-07-13

halo effect

horns effect

seeing one positive/negative aspect of something causes other aspects to be intuitively judged as more positive/negative

2016-07-13

cognitive ease

how easily some information or example comes to mind

often used by intuition as a proxy for harder questions

2016-07-13

S1 storytelling

S1 interprets data by constructing plausible stories, while ignoring strength of evidence

2016-07-13

Kegan 5

meta-rational stage

having control over one's ideology, and being able to juggle around multiple ones

personal identity is shifted to the meta level

2016-07-13

Kegan 4

rational stage

being motivated by one's ideology, and having control over
social relationships

> personal identity is built around one's choices

2016-07-13

Kegan 3

pre-rational stage

being motivated by social relationships, and having control
over long-term personal interests

> personal identity is tied to social roles

2016-07-13

Kegan 2

being motivated by long-term personal interests, and having
control over temporary impulses

> personal identity is adopted from the environment

2016-07-13

Kegan 1

being motivated by temporary impulses, and having control
over perceptions

> personal identity is not well defined

2016-07-13

subject-object shift

transforming an implicit frame to an object that can be
manipulated in context

intermediate stages include: becoming aware of the frame, double vision,
occasional lapses

2016-07-13

System 1

System 2

thinking that is:
(1) unconscious, fast, parallel, and runs all the time
(2) conscious, slow, serial, done with effort when necessary

only (2) uses working memory

2016-07-13

reinforcement learning

inverse reinforcement learning

(1) interacting with the environment to learn how to behave to maximize a value function
(2) figuring out a value function from observing behaviour

2016-07-13

entitlement point

loss aversion

(1) intuitive judgement is based on losses and gains relative to a reference point
(2) losses weigh around 2 times more than gains

2016-07-13

regression to the mean

a prediction from a correlated variable requires adjusting towards the base rate

more noise or smaller sample requires bigger adjustment

2016-07-13

motivation system

planning system

the brain subsystem that:
(1) calculates "wanting" and "not wanting" for each action
(2) plans sequences of actions and intermediate states

(2) has built in strong, low level protections to make sure it complies with (1)

2016-07-13

ugh field

a mental flinch from thinking about a particular topic (or even noticing that there is a problem)

2016-07-13

Bayesian epistemology

beliefs have degrees of certainty from 0 to 1, which are adjusted after seeing evidence

2016-07-13

Aristotelian epistemology

statements are true or false, and truth can be shown by deductive reasoning

2016-02-25

emotional processing

treating emotions as a todo list, and trying to learn from or act on them