Intelligence Metrics

Measuring the Degree of Intelligence in Design

BillDembski.com

For the CELS (= Conference on Engineering in Living Systems) meeting in Denton, Texas, June 1–3, 2023.

1

Intelligent Quotients for ID?

- ID has tended to focus on design as an all-or-nothing proposition—it's designed *or* it's not designed (*or* we can't tell that it's designed).
- ID could as well focus on design as a matter of degree, gauging the degree of intelligence involved in the design.
- This would mean coming up with something like IQ or intelligent quotient metrics for design. (Note the plural.)

2

Percentile Intelligence Metrics: IQ Scores

- 130 and above: 98th percentile
  (Cutoff for “gifted” programs)
- 120-129: 91st-97th percentile
- 110-119: 75th-90th percentile
- 100-109: 50th-74th percentile
- 90-99: 25th-49th percentile
- 80-89: 9th-24th percentile
- 70-79: 2nd-8th percentile
- 69 and below: 2nd percentile or below

3

Percentile Intelligence Metrics: ACT Scores

36: 99th percentile	27: 85th percentile	18: 38th percentile
35: 99th percentile	26: 82nd percentile	17: 31st percentile
34: 99th percentile	25: 78th percentile	16: 26th percentile
33: 98th percentile	24: 74th percentile	15: 20th percentile
32: 97th percentile	23: 69th percentile	14: 15th percentile
31: 95th percentile	22: 63rd percentile	13: 11th percentile
30: 93rd percentile	21: 57th percentile	12: 7th percentile
29: 91st percentile	20: 51st percentile	11: 5th percentile
28: 88th percentile	19: 44th percentile	

4

Percentile Intelligence Metrics: SAT Scores

- 1600: 99+ percentile
- 1500: 98th percentile
- 1400: 94th percentile
- 1300: 86th percentile
- 1200: 74th percentile
- 1100: 58th percentile
- 1000: 40th percentile
- 900: 25th percentile
- 800: 10th percentile

5

Competitive Intelligence Metrics: Golf

Official World Golf Ranking (OWGR) system awards points based on players' performances in recognized worldwide tournaments over a two-year "rolling" period, with the points awarded for each tournament calculated based on the strength of its field (i.e., the ranking of the participating players). The most recent 13 weeks of play are given the most weight in the calculation.

6

World Men's Golf Ranking, May 28, 2023

1	SCOTTIE SCHEFFLER	11.1367	567.9694
2	JON RAHM	10.638	478.7093
3	RORY MCILROY	8.3769	368.5841
4	PATRICK CANTLAY	7.322	292.8806
5	XANDER SCHAUFFELE	6.5476	294.6436
6	MAX HOMA	5.6012	263.2546
7	VIKTOR HOVLAND	5.57	289.6417
8	MATT FITZPATRICK	5.3128	270.9553
9	CAMERON SMITH	5.3127	212.5078
10	WILL ZALATORIS	4.8859	195.4369

SOURCE: https://www.owgr.com/current-world-ranking

7

Competitive Intelligence Metrics: Chess

- The ranking of chess players worldwide is based on the Elo rating system, which was developed by Arpad Elo. The system calculates the relative skill levels of players in two-player games such as chess. (Compare golf, in which multiple players compete among themselves, rather than head to head.)
- Each player starts with an initial rating, and then the player's rating increases or decreases based on game outcomes (win-lose-draw). The amount of change depends on the rating of the opponent.
- Beating a higher-rated opponent leads to a greater rating increase than beating a lower-rated player, while losing to a lower-rated player leads to a greater rating decrease than losing to a higher-rated player.
- Drawing a comparably rated opponent doesn't change your rating, but drawing to a weaker/stronger opponent respectively lowers/raises your rating.

8

World Chess Ranking, May 2023

1	Carlsen, Magnus	NOR	2853
2	Nepomniachtchi, Ian	RUS	2794
3	Ding, Liren	CHN	2789
4	Firouzja, Alireza	FRA	2785
5	Nakamura, Hikaru	USA	2775
6	Giri, Anish	NED	2768
7	Caruana, Fabiano	USA	2764
8	So, Wesley	USA	2760
9	Anand, Viswanathan	IND	2754
10	Radjabov, Teimour	AZE	2747

SOURCE: https://ratings.fide.com/top.phtml

9

Measuring intelligence via information

- Information at its most general is about the elimination of possibilities.
- The more possibilities eliminated, the greater the information (compare “royal flush” vs “two of a kind” in poker).
- The total possibilities eliminated gives a measure of information.
- We will denote the possibility space by Ω.
- The information in a subset T of Ω will thus compare $|T|$ with $|\Omega|$.
- If there’s a probability measure P on Ω, then the information in T will by definition be $I(T) = -\log_2 P(T)$.

10

The etymology of intelligence and information

- Intelligence: Latin *inter+lego* = to choose between
- Key activity of intelligence: Decision
- Decision: Latin *de+caedere* = to cut off or kill (compare homicide)
- Information: Latin *in+formare* = to put shape into (put one shape and therefore not another)

11

The intelligence of my neighbor dog Mac

- My neighbor dog Mac can open doors with lever handles.
- I can open doors with lever and knob handles as well as sliding doors.
- Locksmiths can open still more doors.
- These differences in demonstrated ability signify a difference in intelligence (for the task at hand, which in this case is opening doors).

12

The intelligence of Mac and others

- Consider the doors that "Mac" can't open, CD_{Mac}
- Consider the doors that I can't open, CD_{me}
- Consider the doors that locksmiths can't open, $CD_{locksmiths}$
- Consider the doors that God can't open, CD_{God}
- $\emptyset = CD_{God} \subset CD_{locksmiths} \subset CD_{me} \subset CD_{Mac}$
- Therefore, God is more intelligent at opening doors than locksmiths, who are more intelligent than me at opening doors, who is more intelligent than Mac at opening doors.

13

The scale of intelligence — *scala intelligentiae*

- Given an information-based intelligence metric over the space of closed-door possibilities, it follows that:
  - $\infty = I(CD_{God}) > \ldots$
  - $I(CD_{locksmiths}) > I(CD_{me}) > I(CD_{Mac}) > I(CD_{average\ dog})$

14

Important proviso: Intelligence is task specific

- When it comes to opening doors, I have the advantage over Mac, who has the advantage over most other dogs.
- When it comes to sniffing around to find food, Mac has the advantage over me and locksmiths. Mac is more intelligent at such tasks.
- That said, we can always expand the range of tasks to include subsidiary tasks. Thus, we might imagine an inclusive task that combines opening doors and sniffing for food.
- But with such inclusive tasks, neither Mac nor I would exhibit a "dominating" intelligence—more on this shortly.
- $\Omega = \Omega_1 \oplus \Omega_2 \oplus \cdots \oplus \Omega_k$.

15

Unrestricted Intelligence Metrics

- The information-based intelligence metric we just considered looks at where an agent's ability at a task breaks down.
- It therefore assigns an information measure to those possibilities that demonstrate inability.
- The smaller the range of possibilities where (a given) ability breaks down, the greater the information and thus the higher the intelligence.
- Such an intelligence metric applies to the entire possibility space. It is unrestricted.

16

Relativized Intelligence Metrics

- We could also consider the relative intelligence between two agents.
- For instance, we might not care about the door-opening ability of a rabbit or a slug, but we might want to know the relative door-opening ability (and thus intelligence) between Mac and me.
- Thus, we might want to know *how much more intelligent* I am at opening doors than Mac (confined just to this task or activity).
- In information-theory, such a relativized measure would be denoted by $I(CD_{me}|CD_{Mac})$. Note that $I(CD_{Mac}|CD_{me}) = 0$.
- If there's an underlying probability measure, $I(CD_{me}|CD_{Mac}) = \log_2 P(CD_{me}|CD_{Mac})$.

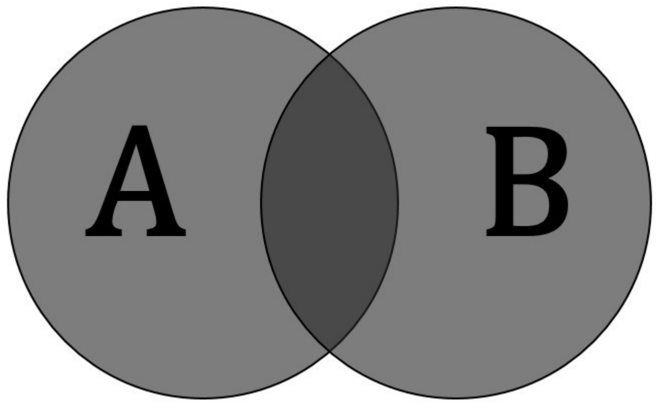
17

Dominating vs. Non-Dominating Intelligence

- In the previous example, I could open all doors that Mac could open, and locksmiths could open all doors that I could open. My intelligence in opening doors therefore *dominated* Mac's …
- But what if we have a task where one agent is better at some aspects of it and another agent is better at other aspects? In that case, intelligence would be non-dominating.
- If A represents the inability set of one agent and B represents the inability set of another agent, then the relative intelligence of the two agents is, very roughly, represented by a Venn diagram such as the following.

18

Relative Non-Dominating Intelligence



19

Relative Non-Dominating Intelligence

- In such cases, we do well to look at $I(A)$ and $I(B)$ and compare them. It may, for instance, be that $I(A)$ is much larger than $I(B)$, suggesting that A is, on balance, much more intelligent than B in the task at hand, but that for certain outliers or anomalies, B might have the advantage.
- In such cases, because neither A nor B is included in the other, $I(A|B)$ and $I(B|A)$ will signify relative intelligence advantages of A over B and vice versa respectively.

20

When ability is not all-or-nothing

- Intelligence comes in degrees, but so far we've treated the ability of agents to succeed or fail at a task as all-or-nothing.
- But what if performance on a task is itself a matter of degree?
- Imagine, for instance, two golfers, A and B. We want to know the intelligence of the two golfers at sinking 18 balls in an 18-hole course.
- In this case, A's intelligence compared to B's will be probabilistic. On rare occasions, B might have a lower (better) score than A, but in general A will have the lower (better) score.

21

The example of golf

- Continuing from the previous slide, let's therefore imagine that the possibility space Ω consists of {18, 19, 20, 21, 22, …}.
- Let f denote A's probability density over Ω and g denote B's probability density. f will tend to concentrate probability closer to 18, g closer to ∞.
- Let n be the largest integer such that for $18 \leq k \leq n$, $f(k) > g(k)$.
- In that case, $\sum_{18 \leq k \leq n} \left[\frac{f(k)}{g(k)}\right]^r$ for $r > 0$ measures the superiority of A's intelligence at sinking golf balls over B's. Putting a log to the base 2 in front of this and setting $r = 2$ overlaps with standard information measures (need to do some normalization).

22

The Rényi Information Divergence (or Rényi Entropy or just Rényi Information)

- Alfred Rényi, in 1961, generalized information measures for sets to information measures for densities.
- For a random variable X defined on Ω and density f induced by a random variable X on the real numbers $\mathbb{R}$, Rényi defined the quantity

$$h_r(X) =_{def} \frac{1}{1-r} \log_2 \int_{\mathbb{R}} f(x)^r \, dx$$

- for $0 < r < \infty$ and $r \neq 1$.

—Rényi, Alfred, "On Measures of Information and Entropy," in J. Neyman, ed., *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1 (Berkeley, Calif.: University of California Press): 547—561.

23

Variational Information

- In 2004 (unpublished typescript on my website), I introduced a (partial) generalization of the Rényi information.
- For probability measures μ_1 and μ_2 such that $\mu_2 \ll \mu_1$ (the former being "absolutely continuous" with respect to the latter), the Rényi information readily generalizes to

$$h_r(\mu_2|\mu_1) =_{def} \frac{1}{r-1} \log_2 \int_{\Omega} \left(\frac{d\mu_2}{d\mu_1}\right)^r d\mu_1$$

- The case of $r = 2$ provides the most natural extension of conventional information measures, with the integral then corresponding to the variance plus 1 of $\left(\frac{d\mu_2}{d\mu_1}\right)$. This last term is the Radon-Nikodym derivative. Information divergence is basically a variance.

24

Variational Information

$$I(\mu_2|\mu_1) =_{def} \log_2 \int_\Omega \left(\frac{d\mu_2}{d\mu_1}\right)^2 d\mu_1$$

$$= \log_2(\mathrm{Var}_{\mu_1}(\frac{d\mu_2}{d\mu_1})+1)$$

Think of this metric as measuring how much the second intelligence diverges from the first.

25

Intelligence vs. Performance in General

- If someone can deadlift 600 lbs, that indicates a high level of performance, but not intelligence.
- If someone can prove a complicated mathematical theorem, that indicates both a high level of performance and intelligence.
- What is the difference?
- Promising answer: search.

26

Golf again

- This is why we think of golf not just as a sport, but as a game of intelligence.
- The golfer is, as it were, searching for the right way to connect with the ball to place it where it will have the greatest probability of immediately, or mediately, getting into the hole.
- The weight lifter, by contrast, is performing an act of brute strength. Granted, the weight lifter needs good form, and a lot intelligence will go into the training regimen ("searching for the right training techniques"). But in the actual performance, intelligence in the sense of search is less/minimally the issue.

27

Two approaches to measuring intelligence

- Measuring the degree of intelligence exhibited by a designing agent on the basis of the agent's real-time performance.
- Measuring the degree of intelligence exhibited in the designed object without necessarily having access to the designing agent.
- The first involves cause-to-effect reasoning; the second effect-to-cause reasoning (as typical in design inferences).

28

Specified complexity the key to search

- To gauge intelligence in search, it's not enough just to have a small probability target: $T \subset \Omega, P(T)$ small, so that $I(T) = -\log_2 P(T)$ is large, suggesting high intelligence.
- We also need a specification (i.e., for T to be specified), so that T is not just some arbitrary subset of Ω, but a subset of the sort that, if found, would suggest that it was found through intelligence.
- The underlying intuition here: The smaller the target, the more improbable to find it by chance, and thus the greater the intelligence needed to overcome chance (provided that the target is specified—unspecified events of small probability can happen by chance).

29

Definition of specified complexity

- Here, then, is the definition of specified complexity. This can serve as an intelligence metric:

$$SC(T) = I(T) - D(T)$$

- $I(T) = -\log_2 P(T)$ — the information in T.
- $D(T) = \min_{W \text{ where } W^* = T} |W|$ — the minimum description length of T.
- W^* is the event in Ω that the description W describes.
- Assumption about underlying language: binary, prefix free, Turing complete.

30

Examples to clarify specified complexity

- Cars in succession
- Coin tosses: 11111…, 00000…, 101010…,010101…, 00110011…, etc.
- In *The Empire Strikes Back*, Darth Vader tells Luke Skywalker, “No, I am your father.”
- In *Spaceballs*, Dark Helmet (Rick Moranis) tells Lone Starr (Bill Pullman), “I am your father’s brother’s nephew’s cousin’s former roommate.”

31

HELMET Before you die, there is something you should know about us, Lone Starr.

LONE STARR What?

HELMET I am your father's brother's nephew's cousin's former room-mate.

LONE STARR What's that make us?

HELMET Absolutely nothing. Which is what you are about to become. Prepare to die.

32

Important fact about specified complexity

- Consider the union of all events/targets T whose specified complexity with respect to an underlying chance hypothesis is at least σ, i.e.,

$$R_\sigma = \bigcup \{T \mid SC(T) \geq \sigma\}\,.$$

- Then the probability of R_σ is bounded above by $2^{-\sigma}$.
- Search for and finding such targets T therefore trigger design inferences (since specified complexity is a marker of design).

33

Why do engineers produce high SC/CSI artifacts?

- A designing engineer has a simply stated functional goal (which requires only a short description and thus produces a specification).
- But achieving that goal is difficult (hence the low probability.)
- Put those together, and you have specified complexity or complex specified information.
- Specified complexity comes in degrees. We often use cutoffs to decide whether enough intelligence is evident to preclude chance.
- 500 bits corresponds to an improbability of roughly 10^{-150}, which in TDI2 is argued to constitute a universal probability bound.

34

SC/CSI as an intelligence metric

- The greater the SC, the more intelligence was inputted.
- How much SC can material mechanisms actually generate [versus simply shifting around existing SC]?
- How much SC can the Darwinian mechanism generate?
- What do we make of mechanisms/technologies such as ChatGPT? Does ChatGPT create SC? No, it repackages existing SC.
- Compare an embossed sign that falls in a snow storm and inputs SC into the snow (e.g., "Eat at Joe's").

35

Intelligence via IBE vs. Intelligence via SC

- ID critic Jason Rosenhouse argues that SC adds nothing to ID, which he regards as essentially an IBE argument.
- For instance, he sees Mike Behe's argument from IC (irreducible complexity) as an IBE argument for the improbability of the chance formation of such systems/intelligent design of such systems.
- SC therefore adds nothing, as far as Rosenhouse is concerned, to our probabilistic understanding. WRONG
- IBE arguments come in degrees, but typically not in quantifiable degrees. SC/CSI arguments assign precise numbers.
- SC also adds rigor that tends to be absent from IBE arguments. Consider the case of Sisyphus. Historical vs. Analytic Probabilities.

36

Measuring intelligence from cause to effect vs. effect to cause.

- A conventional (probability based) information measure seems fine for a cause-to-effect intelligence metric.
  - That's because we know we're dealing with intelligent agents and the aim is to assess probabilistic advantages.
- Specified complexity or something like it is needed for an effect-to-cause intelligence metric.
  - That's because we're unsure we're dealing with an actual intelligent agent and thus need more than probability or improbability. In particular, we also need specification.

37

Intelligence Metric Based on Conservation of Information

- Baseline or null search has probability p of success [p is very small—needle-in-the-haystack small] — in information-theoretic terms I_p.
- Enhanced or alternate search has probability q of success [q much larger than p and giving high probability of success; q could equal 1] — in information-theoretic terms I_q.
- Active information $-\log_2 \frac{p}{q} = \log_2 \frac{q}{p} = I_p - I_q = I_+$ gives the amount of information, in bits, that needs to be added to the baseline search to improve its performance to that of the enhanced search.
- Other terms used respectively have been endogenous, exogenous, and added information.

38

Key Theoretical Publications on CoI

- William A. Dembski and Robert J. Marks II, "Conservation of Information in Search: Measuring the Cost of Success," *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, Vol. 39(5):1051-1061 (September, 2009).
- William A. Dembski and Robert J. Marks II, "The Search for a Search: Measuring the Information Cost of Higher Level Search," *Journal of Advanced Computational Intelligence and Intelligent Informatics*, Vol. 14 (5):475-486 (2010).
- William A. Dembski, Winston Ewert, and Robert J. Marks II, "A General Theory of Information Cost Incurred by Successful Search," in Marks et al., eds., *Biological Information: New Perspectives* (Singapore: World Scientific, 2013).

39

Active information as an intelligence metric

- The active information $-\log_2 \frac{p}{q} = \log_2 \frac{q}{p}$ gauges the information inputted to improve a search from the baseline search characterized by p to the enhanced search characterized by q.
- Active information constitutes a promising intelligence metric. As with the other intelligence metrics considered, the greater it is, the more intelligence required.
- Shannon wrote to Vannevar Bush at MIT in 1939, "I have been working on an analysis of some of the fundamental properties of general systems for the transmission of intelligence." [Quoted from James Gleick, *The Information: A History, a Theory, a Flood*.]

40

Information-density intelligence metrics

- In the first decade of the 2000s, Darwinian biologists (e.g., Francis Collins) were claiming that about 50 percent of the human genome was junk.
- So, if there were n bits of information in the genome's carrying capacity, there were only $.5 \times n$ bits of functional information in the genome.
- On Darwinian grounds, we thus would have a density of .5.
- Dawkins in 2009 put the density at .05: "The greater part (95 per cent in the case of humans) of the genome might as well not be there, for all the difference it makes."

41

ChatGPT on the Dawkins quote

Richard Dawkins made that statement in his book "The Greatest Show on Earth: The Evidence for Evolution," published in 2009. This book presents the scientific evidence supporting the theory of evolution. The quote you mentioned is part of a discussion on "junk DNA" or noncoding regions of the genome. As of my knowledge cut-off in September 2021, the understanding of noncoding DNA has evolved, and it's increasingly acknowledged that these regions can play significant roles in gene regulation and other functions.

42

Information-density intelligence metrics

- The ENCODE project in the second decade of the 2000s has by contrast shown that most of the genome is functional.
- So, if there are n bits of information in the genome's carrying capacity, we now think there are at least $.99 \times n$ bits of functional information in the genome.
- We thus have an information density of close to 1. That is suggestive of intelligent design because designers tend to build things to serve a function.
- Darwinists, on the other hand, expected junk DNA because selection-mutation is sloppy and puts a premium on survival, not on elegance or data compression.

43

Information-density intelligence metrics

- But in fact, information could be more densely packed into genomes/alpha-numeric strings.
- Hetero-palindromes exist in English/DNA, such as the word *stressed*, which read backwards is *desserts*. This indicates an information density of 2.
- Shifted reading frames can be functional. A functional original frame plus two additional functional shifted frames would indicate a possible information density of 3.
- Combining these two approaches could, in principle, yield an information density of 6.
- This is a 6-fold increase in bits, corresponding to a power of 6 decrease in probability, e.g., $n = 3$ bits with an information density of 6 takes an improbability of 1 in 2^3, or 1 in 8, to 1 in $\left(2^3\right)^6 = 2^{18}$, or 1 in roughly 260,000.

44

Information-density intelligence metrics

- What's the intelligence metric here? Two possibilities:
  - The actual density r.
  - $I_r = (r - bd) \times (\#_{\text{bits}})$ (bd = baseline density).
- For Darwinists, r and bd are (or used to be) a lot less than 1.
- ID theorists have traditionally thought r and bd both to be around 1, but now we are finding that r could be a lot more than 1.
- Other places to look for $r > bd$: digital data embedding technologies, steganography, watermarking.
- Multi-dimensional information density — the SATOR-AREPO word square: $bd = 1$, $r = 4$, $\#_{\text{bits}} = 125$ (5 bits per letter), $I_r = 375$.

45

The SATOR-AREPO Word Square

S A T O R
A R E P O
T E N E T
O P E R A
R O T A S

"Arepo the sower holds the wheels at work."

46

Information-density intelligence metrics

- What if organisms/engineered systems have layers of information?
- One layer could be self-interested, to assist the organisms in their life cycle.
- Another layer could be altruistic, helping other organisms but without a benefit to the organism in question, perhaps through some sort of lateral transfer. (This is speculative; it would be utterly non-Darwinian.)
- Another layer could be pedagogical, helping biologist-engineers to discover otherwise hidden features of life. (Compare discoverability in astrophysics as developed by Gonzalez and Richards, where, for instance, solar eclipses are crucial to the advance of astrophysics.).
- What if an "operational manual" were embedded in biology, such as in the genome? (This is speculative and also utterly non-Darwinian.)
- With information density, as soon as it goes over 1, ID seems strongly implicated.

47

D.K.Y. Chiu & T.H. Lui, "Integrated Use of Multiple Interdependent Patterns for Biomolecular Sequence Analysis," *International Journal of Fuzzy Systems*, 4(3) (September 2002): 766–775.

The opening paragraph of this article reads: "Detection of complex specified information is introduced to infer unknown underlying causes for observed patterns [10]. By complex information, it refers to information obtained from observed pattern or patterns that are highly improbable by random chance alone. We evaluate here the complex pattern corresponding to multiple observations of statistical interdependency such that they all deviate significantly from the prior or null hypothesis. Such multiple interdependent patterns when consistently observed can be a powerful indication of common underlying causes. That is, detection of significant multiple interdependent patterns in a consistent way can lead to the discovery of possible new or hidden knowledge." Reference number [10] here is to William Dembski's *The Design Inference*.

48

The Origination Inequality

$$p_{\text{origin}} \leq p_{\text{avail}} \times p_{\text{synch}} \times p_{\text{local}} \times p_{\text{i-c-r}}$$
$$\times\, p_{\text{i-f-c}} \times p_{\text{o-o-a}} \times p_{\text{config}}$$

Note that these probabilities multiply because each next probability is conditional on the previous one.

49

Seven Probabilistic Hurdles That Must Be Overcome in Building a System

- Availability
- Synchronization
- Localization
- Interfering Cross-Reactions
- Interface Compability
- Order of Assembly
- Configuration

50

Availability

Are the parts needed to build a given functioning (perhaps irreducibly complex biochemical) system (perhaps like the bacterial flagellum) even available?

51

Synchronization

Are these parts available at the right time so that they can be incorporated when needed into the system being built (evolved)?

52

Localization

Even with parts that are available at the right time for inclusion in a system being built, can the parts (break free of the systems in which they are currently integrated and) be brought to the "construction site" of the system being built (evolving system)?

53

Interfering Cross-Reactions

Given that the right parts can be brought together at the right time in the right place, how can the wrong parts that would otherwise gum up the works be excluded from the "construction site" of the system being built?

54

Interface Compatibility

Are the parts that are being recruited for inclusion in the system that's being built mutually compatible in the sense of meshing or interfacing tightly/neatly so that, once suitably positioned, the parts work together to form a functioning system? ("Standardization of parts" is important to this point.)

55

Order of Assembly

Even with all and only the right parts reaching the right place at the right time, and even with full interface compatibility, will they be assembled in the right order to form a functioning system?

56

Configuration

Even with all the right parts slated to be assembled in the right order, will they be arranged in the right way to form a functioning system?

Additional probabilistic hurdles

- *Retention probability* p_{reten}, the probability that items available at the right time and in the right place stay at the right place long enough (i.e., are retained) for origination to take place. Place this probability after p_{local} in the origination inequality.
- *Proportionality probability* p_{propor}, the probability that items available at the right time, in the right place, and for long enough occur in the right proportion for origination to take place. Place this probability after p_{reten} in the origination inequality (once retention is factored in).

The Origination Inequality (probabilistic form)

$$p_{\text{origin}} \leq p_{\text{avail}} \times p_{\text{synch}} \times p_{\text{local}} \times p_{\text{i-c-r}}$$

$$\times\, p_{\text{i-f-c}} \times p_{\text{o-o-a}} \times p_{\text{config}}$$

59

The Origination Inequality (information-theoretic form)

$$I_{\text{origin}} \geq I_{\text{avail}} + I_{\text{synch}} + I_{\text{local}} + I_{\text{i-c-r}}$$

$$+\, I_{\text{i-f-c}} + I_{\text{o-o-a}} + I_{\text{config}}$$

"Engineering Construction Inequality" — Each of these terms can be conceived as an intelligence metric.

60

Waiting/Stopping-Time Intelligence Metrics

- How long does it take to successfully conclude a search or finish some task requiring intelligence?
- The shorter the time, the greater the intelligence.
- A waiting time asks how long the time to success.
- Time can be calculated in terms of seconds, hours, etc.
- In computation, technology, and evolution it typically makes more sense to think of time in terms of number of elementary steps to successful conclusion (e.g., FLOPS, assembly steps, mutations, or generations).
- Hössjer, Ola, Günter Bechly, and Ann Gauger, "On the Waiting Time Until Coordinated Mutations Get Fixed in Regulatory Sequences," *Journal of Theoretical Biology* 524(2021) : 110657.

61

Waiting/Stopping-Time Intelligence Metrics

- Let's take N (# of steps) to signify the time to success. The bigger N, the less intelligence; the smaller N, the more intelligence.
- How then to interpret N as an intelligence metric? Suggestion: Treat the geometric distribution with probability p as prototypical for waiting times, extending it by analogy to stochastic processes in general.
- For a geometric distribution, the average waiting time to success is $1/p = E(N)$. If we set $q = 1 - p$, then $q = 1 - \frac{1}{E(N)} = \frac{E(N)-1}{E(N)}$, and it makes sense to define an intelligence metric inspired by the geometric distribution, namely, $I_N = \log_2\left(\frac{1}{q}\right) = \log_2\left(\frac{E(N)}{E(N)-1}\right)$.
- I_N now increases as $E(N)$ decreases. Note that for a divine intelligence that solves everything instantly ($E(N) = 1$), $I_N = \infty$. On the other hand, for a slow intelligence that takes forever ($E(N)$ goes to ∞), I_N will go to zero, indicating zero intelligence.

62

Recap of Intelligence Metrics

- Information measures readily lend themselves to intelligence metrics.
- These metrics can be unrestricted as well as relativized (corresponding to unconditional vs. conditional information).
- Dominating intelligence makes for clearer interpretation of intelligence metrics (the one intelligence is entirely superior to the other in respect of the task in question).
- SC/CSI is the natural intelligence metric when there's a question about who or what the underlying intelligence is (effect-to-cause reasoning).
- Information density measures can be conceived as intelligence metrics, and even conceived as information metrics counting adjusted number of bits.

63

Recap of Intelligence Metrics

- The origination inequality, especially in its information-theoretic form, provides multiple measures of how much intelligence is required to build systems (whether by technological assembly or by an evolutionary process).
- Average waiting-times naturally lend themselves to intelligence metrics (by taking a log of the expected waiting time in analogy to the geometric distribution, which provides the simplest waiting time).
- Percentile intelligence metrics and competitive intelligence metrics were included for completeness, but don't have a natural interpretation.

64

Key Engineering Benefits of Intelligence Metrics

- Intelligence metrics give us a diagnostic tool for where to look for design (in biological as well as engineered systems, which may be the same thing). Think *intelligence threshold*.
- And they give us an assessment tool for determining the degree of intelligence required for design. Think *intelligence test*.

65