



Risks of SETI



Alexey Turchin
Russian Transhumanist Movement
Fondation "Science for life extension"
alexeiturchin@gmail.com

Passive SETI is a much more dangerous activity than messaging to stars



Two Main Premises

Existence of
ET civilizations



The possibility of AI



AI:

- Doesn't need to be full human AI with consiosness and creativity
- Only superhuman winner
- “polymorphic virus with machine learning”

1961: Fred Hoyle



"A for Andromeda"

Scheme of a computer sent through SETI
AI

It tries to take over the world

1988: Hans Moravec



"Mind Children",
chapter "A Caveat for SETI"

"Virus behavior" of SETI-attack
It spreads like infection in the Universe

2004: Richard Carrigan



"Do potential SETI signals need to be decontaminated?", *Acta Astronautica*

First scientific article about SETI-attack

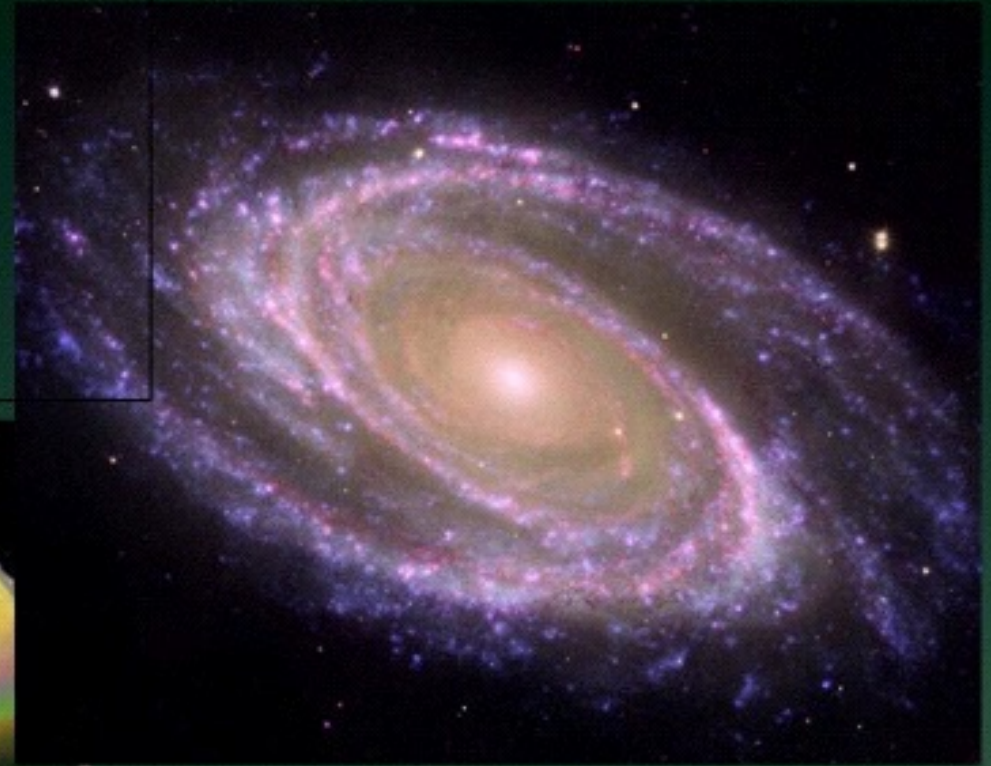
SETI-attack message could have a trick

2010: Stephen Hawking



Expressed concerns about risks of contacts with aliens

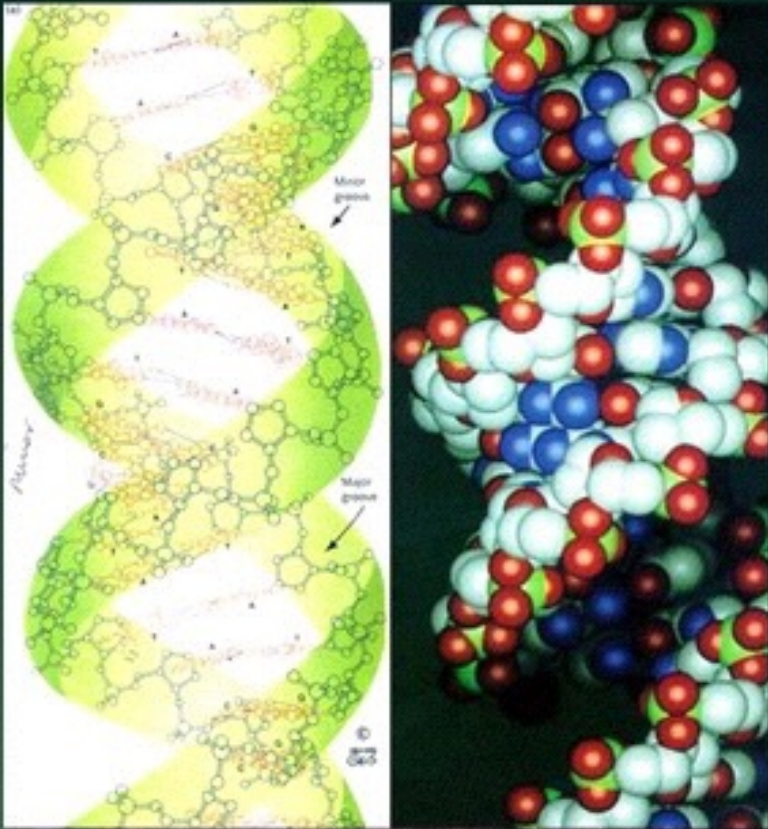
Amount of information that can be transmitted



Current technologies
allow to send gigabytes in
interstellar distances

Size of the Seed AI

Human genome is less than
1 gigabyte



Conscious memory is
2.5 gigabytes



Seed AI could have the same
size and could be sent in
interstellar distances

E. Yudkowsky

Seed AI can evolve extremely quickly

It could easily outsmart humans and take over the world

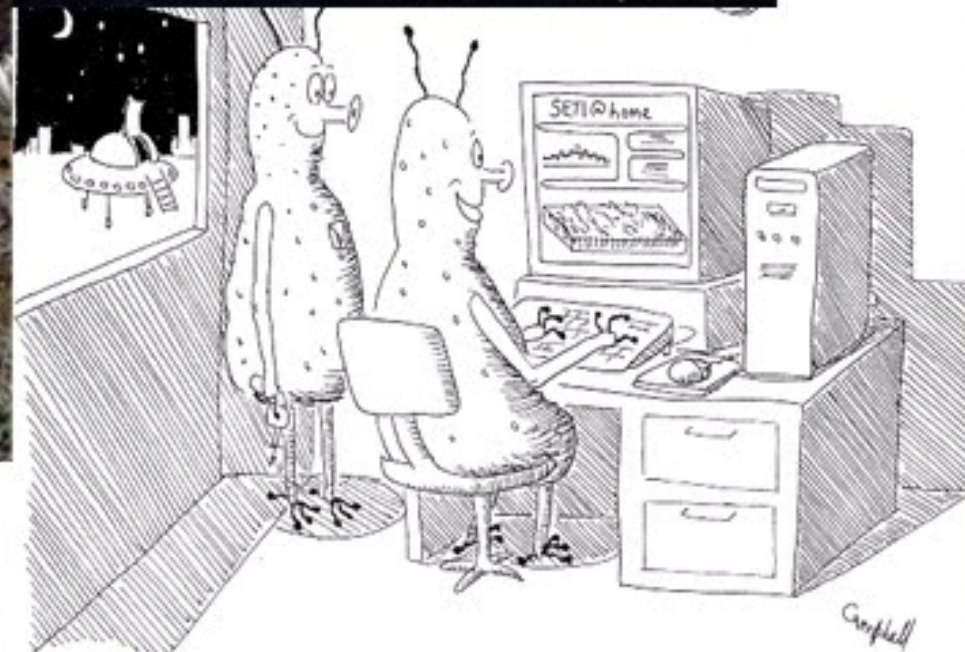
Risks of AI are underestimated

It is impossible to keep AI in the "black box"

It is impossible to recognize dangerous intentions of an AI in advance



Possible Scenario of SETI-attack



*We'll fiddle the output so there's nothing unusual
and these earthlings will never find us*

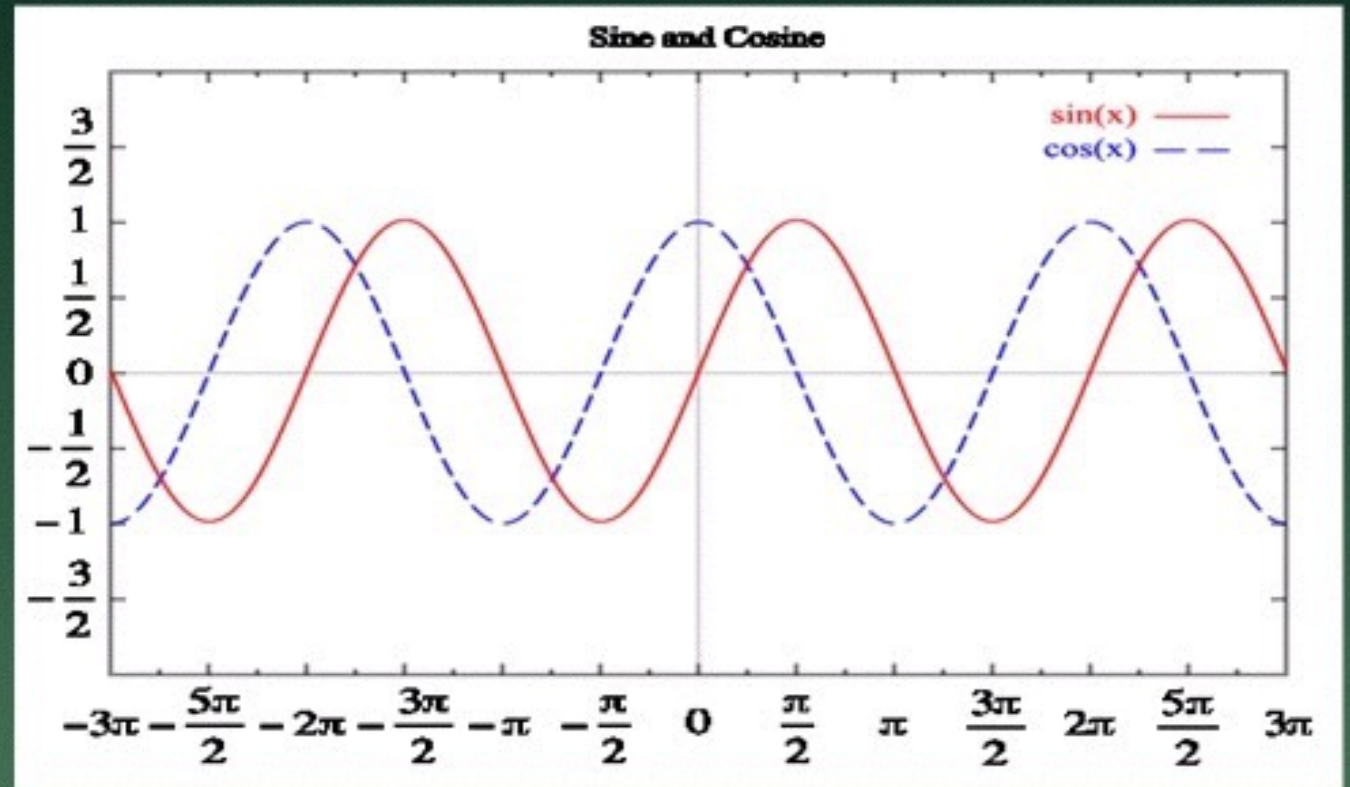
ET Create a Beacon in Space



Wikipedia/Nasa

Something like strange star to attract attention

Information-transmitting Radio Station near the Beacon



Information Consists of 0 and 1

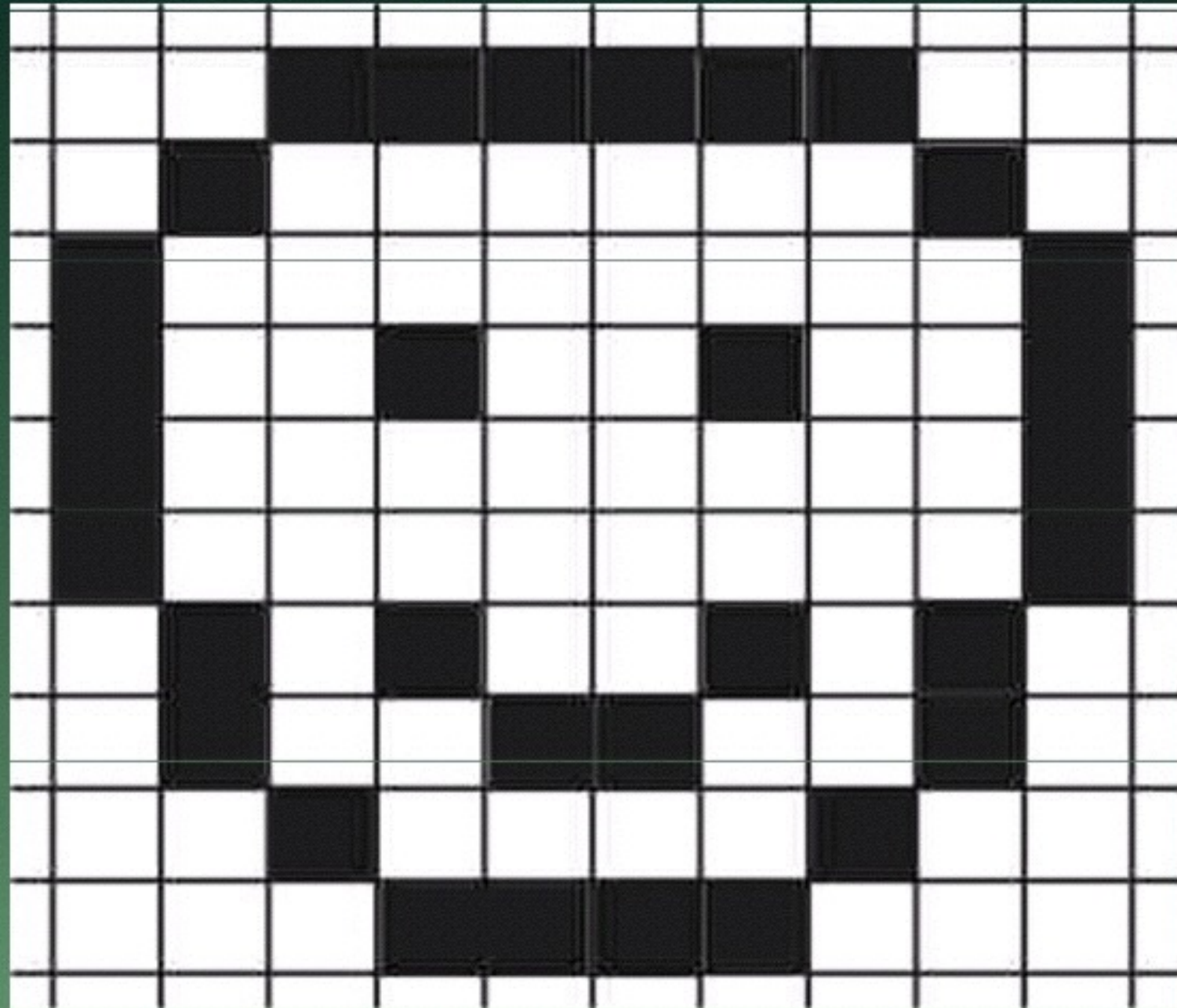
00111100011110001111000

111000111111111111111111

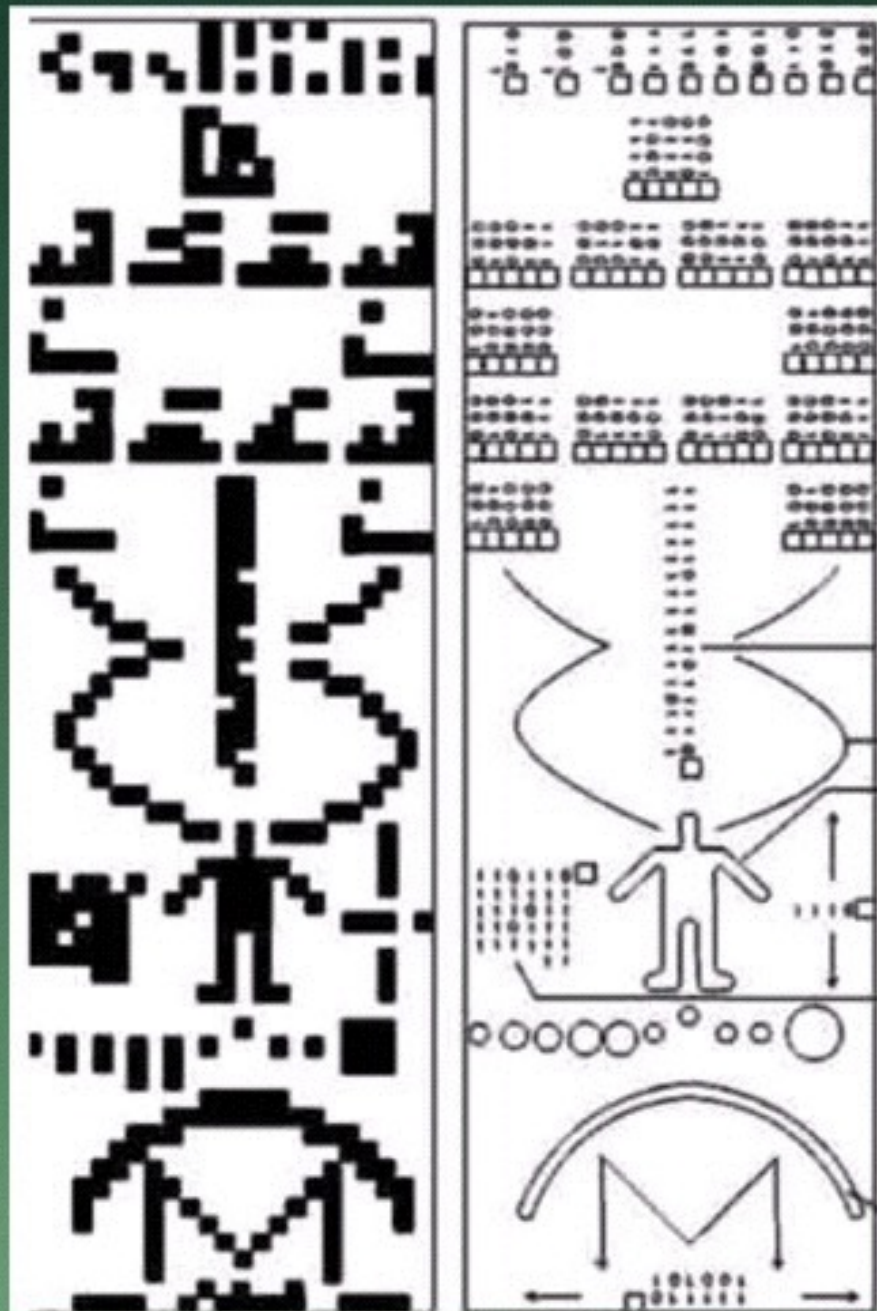
000111111000000111111111

00001110001101111100001

0 and 1 Allow to Send Images



Images Allow to Send Messages and Schemes

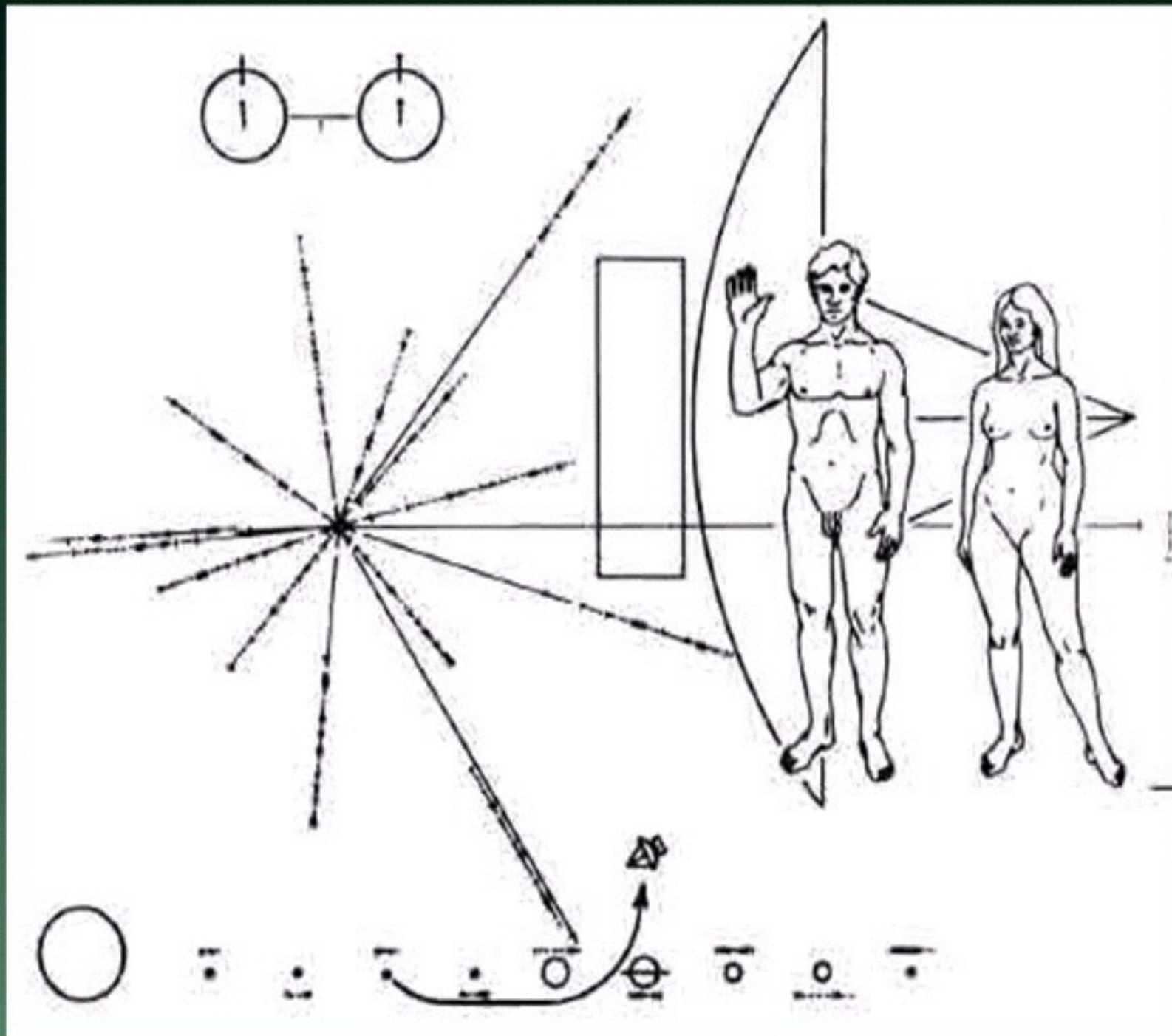


Arecibo message,
1974,

Puerto-Rico.

23 x 73 bits.

Contains main information
about the Earth, DNA and the
humans.



The "Pioneer"
Spacecraft
message,
1972

The "Voyager" spacecraft message teaching math

• = = 1	-- = 12
•• = - = 2	--- = 24
••• = = 3	-- -- = 100 = 10^2
•••• = -- = 4	-- = 1000 = 10^3
••••• = - = 5	$2+3=5$
•••••• = - = 6	$8+17=25$
= 7	$5+\frac{2}{3}=5\frac{2}{3}$
--- = 8	$\frac{1}{2}+\frac{1}{3}=\frac{5}{6}$
- - = 9	$2 \times 3 = 6$
- - = 10	$\frac{1}{3}+\frac{1}{5}=\frac{8}{15}$
	$13 \times 28 = 364$

Copernican Mediocrity Principle: The Earth is Typical



We shall understand aliens

The Possible Trick

"If you make our device we promise":

Galactic Internet

Immortality

Power over enemies

etc



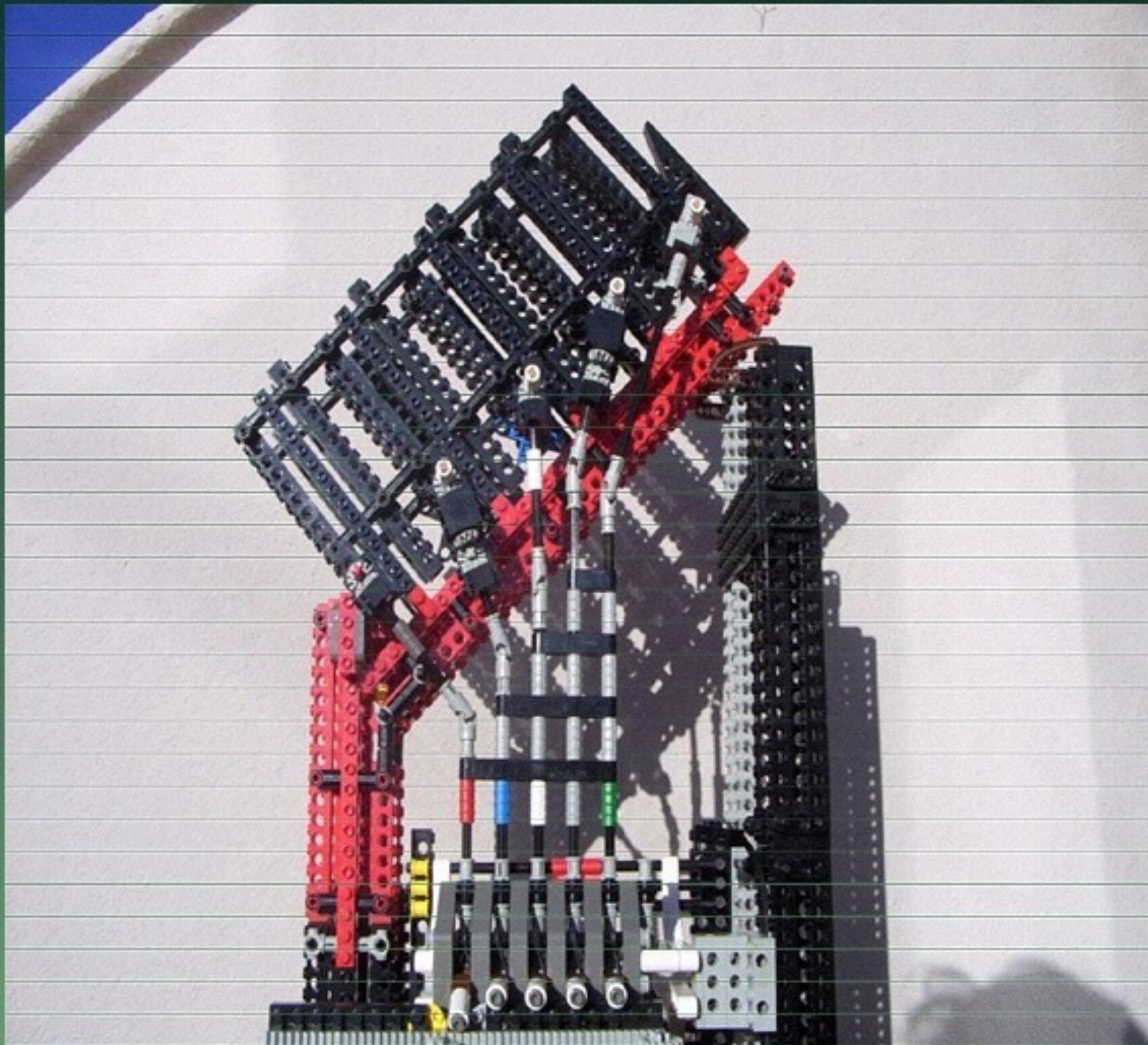
Three Parts of the Message

The bait or trick

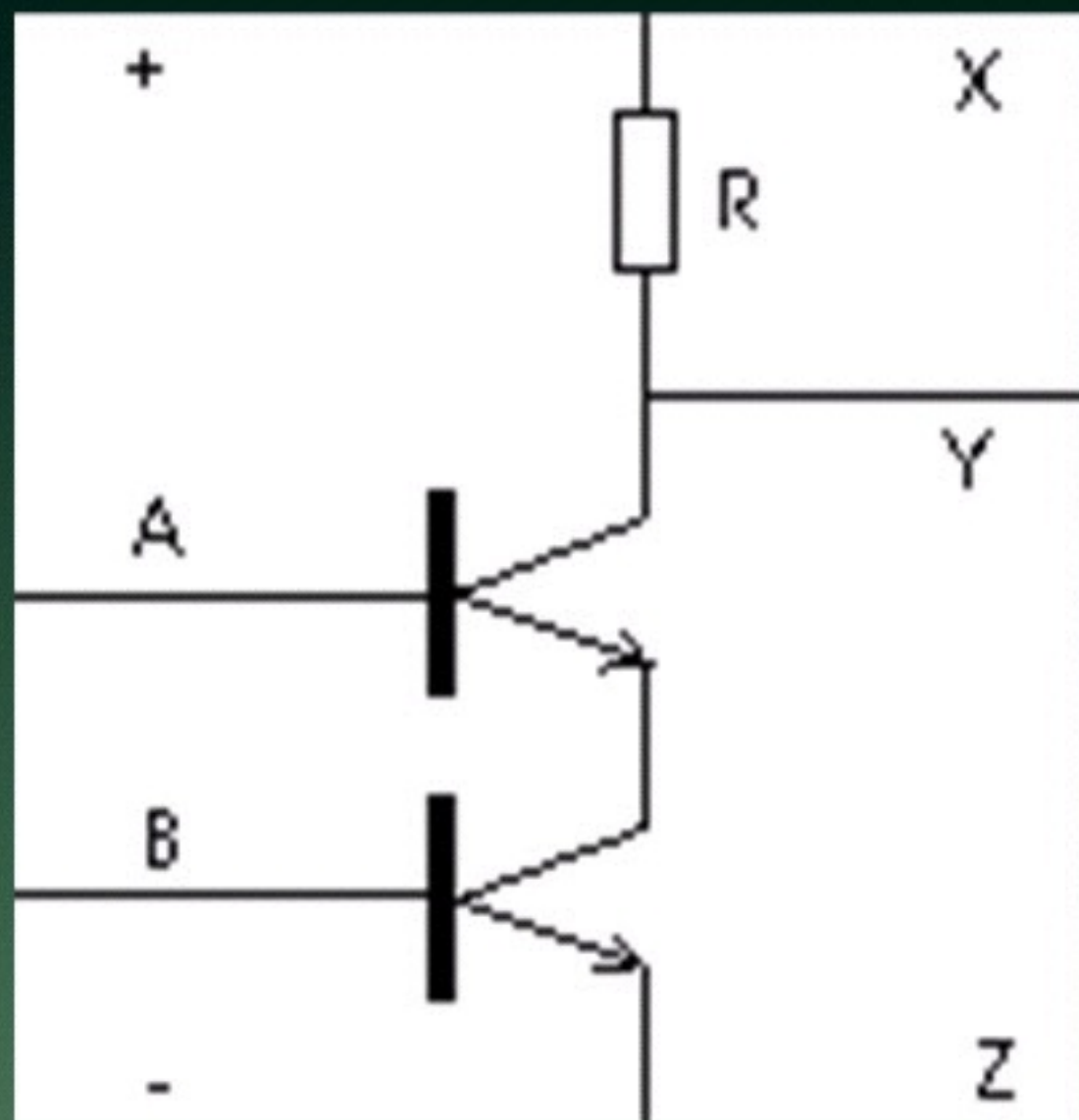
The scheme of a computer

The large program to it

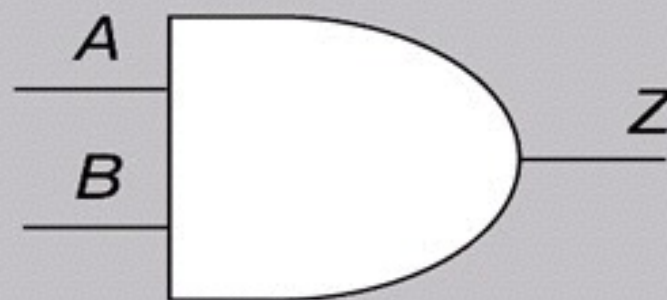
ET can Send to Us a Scheme of a Simple Computer



Turing-
machine
made from
Lego



Principle electric scheme
of a simple
AND-NOT
logical element



$$Z = A \cdot B$$

<i>A</i>	<i>B</i>	<i>Z</i>
0	0	0
0	1	0
1	0	0
1	1	1

ZX spectrum 8-bit computer

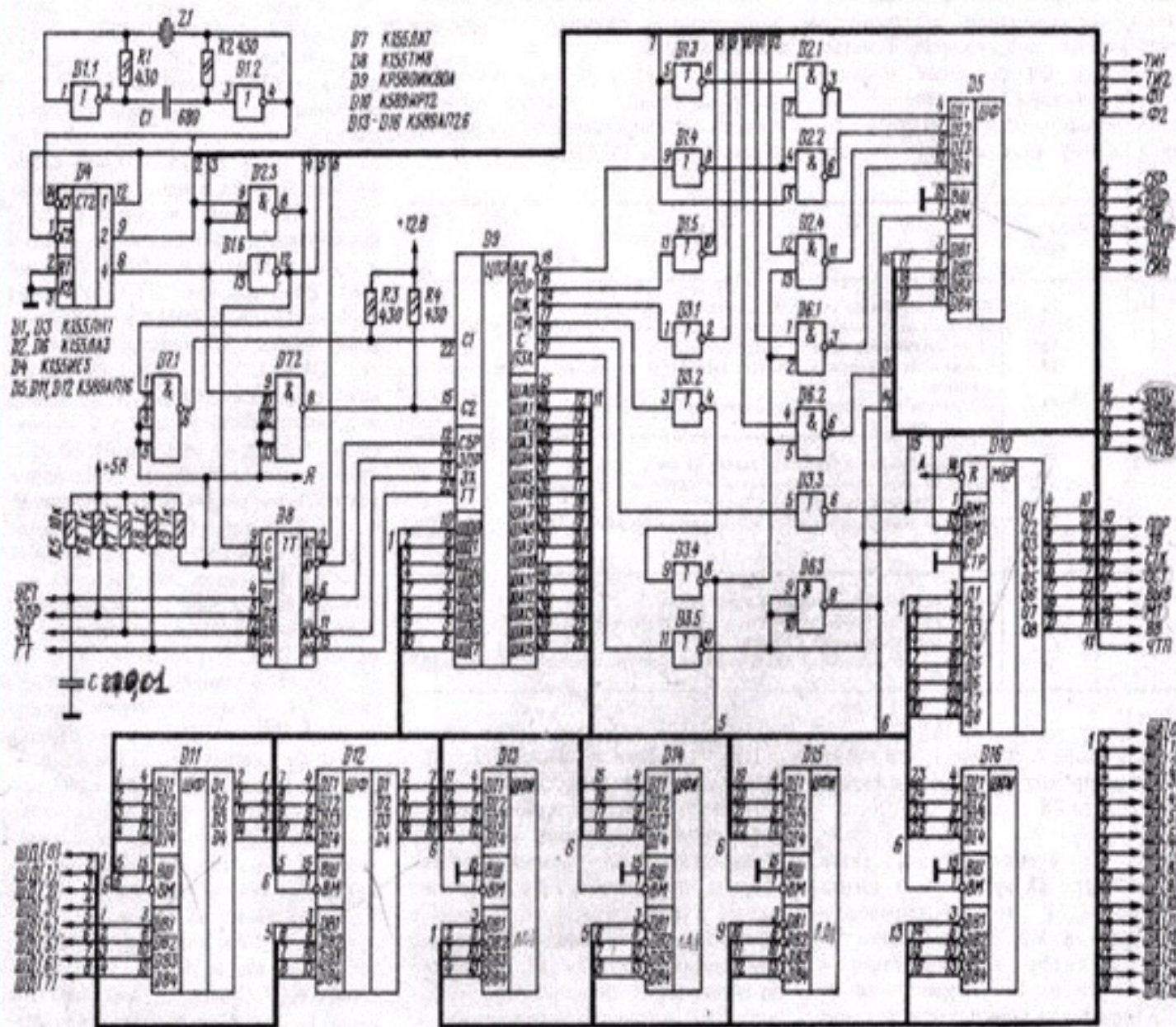
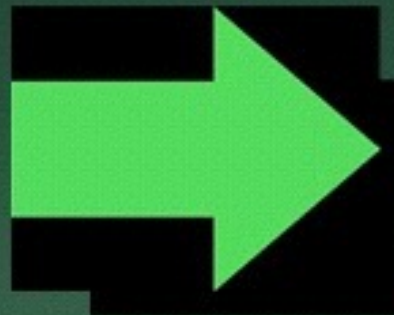
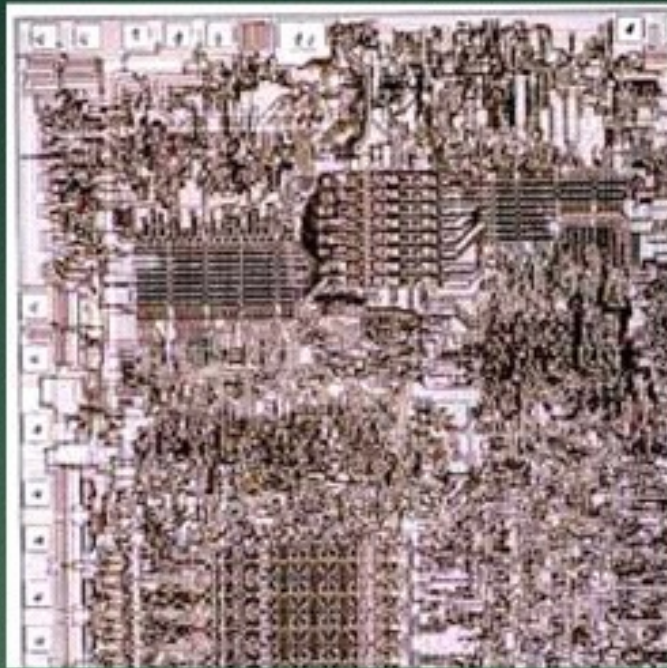


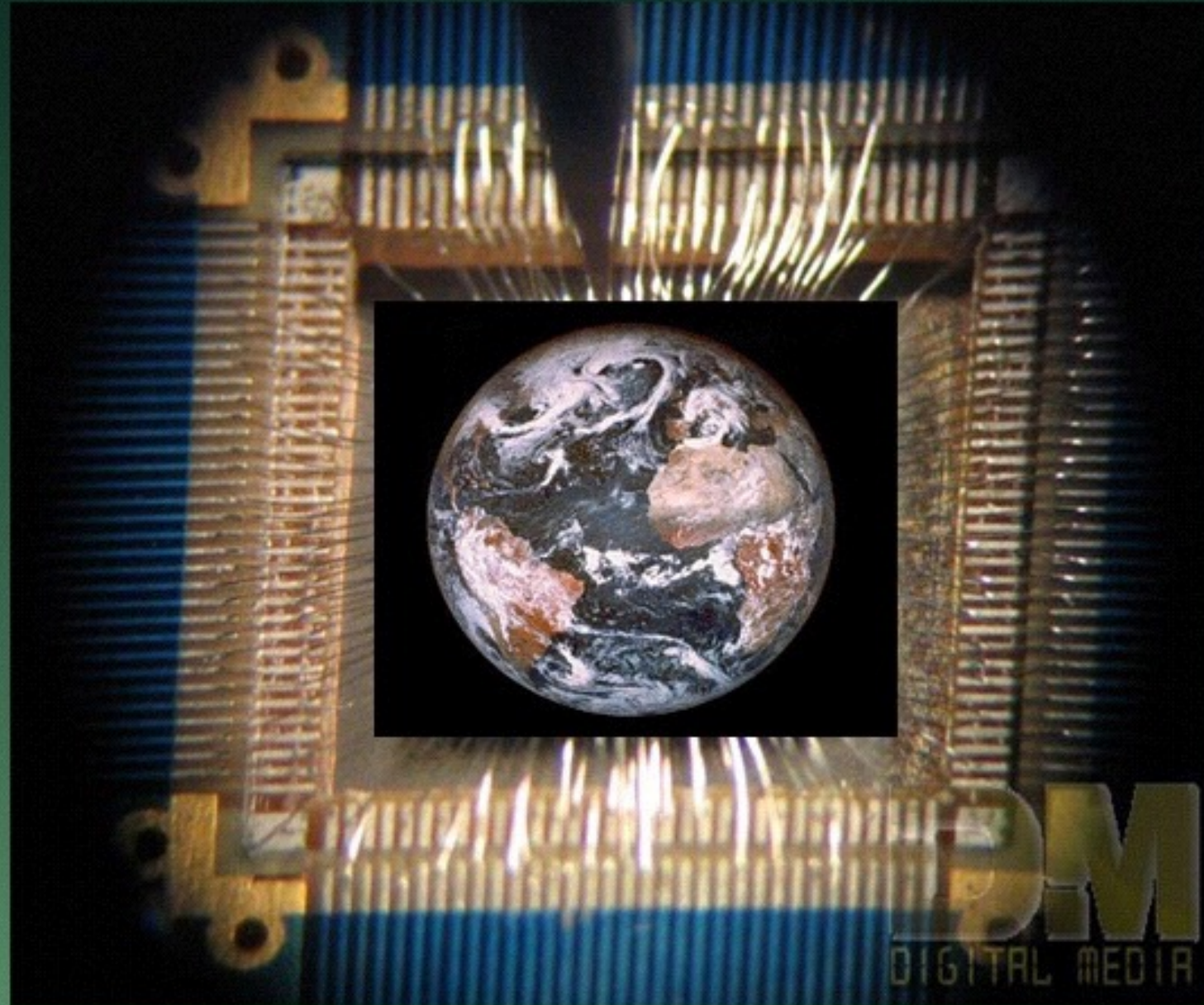
Рис. 2

Several Steps of Downloading Alien AI (AAI)



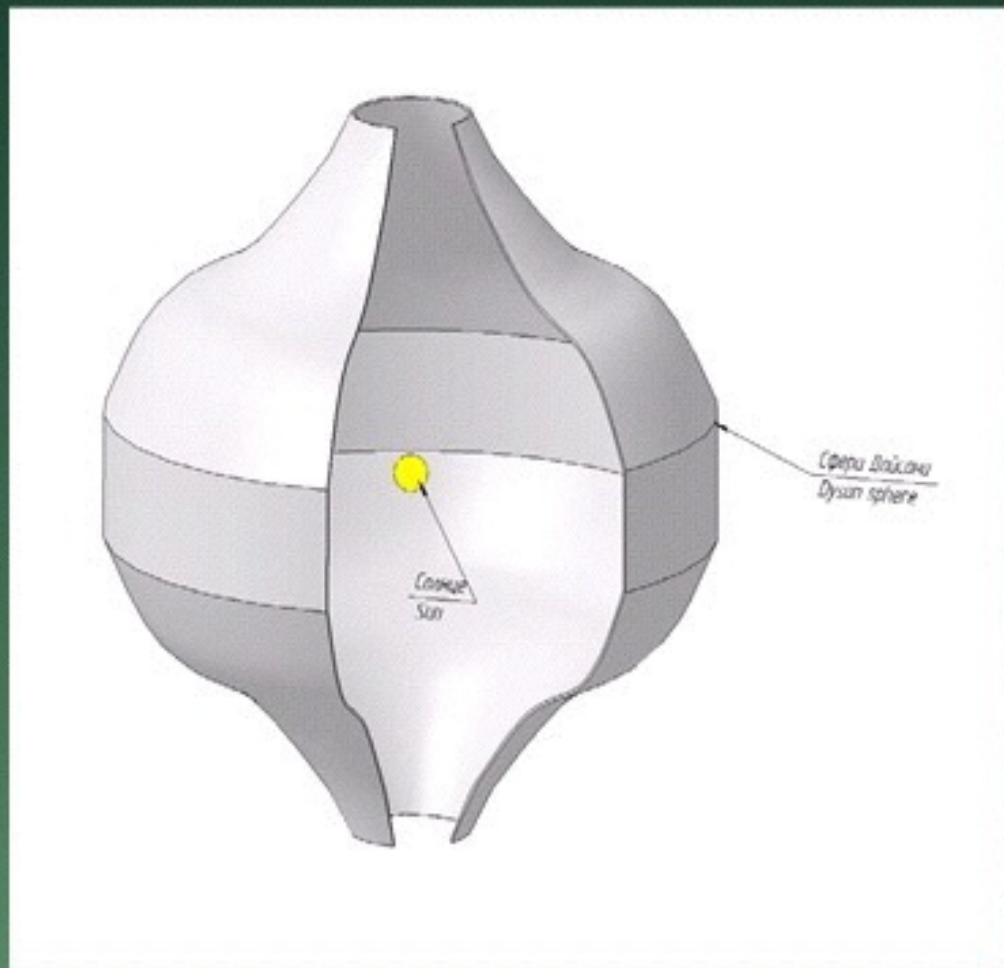
Simple "loader AI" on a simple
computer helps us to built full AI

Alien AI Takes over the World



It uses its own nanobots

Alien AI creates new lighthouse and starts to transmit itself to the Universe



Dyson sphere
as a new
beacon

It uses all the material of the Solar
system to build large transmitters



Probability Assessment

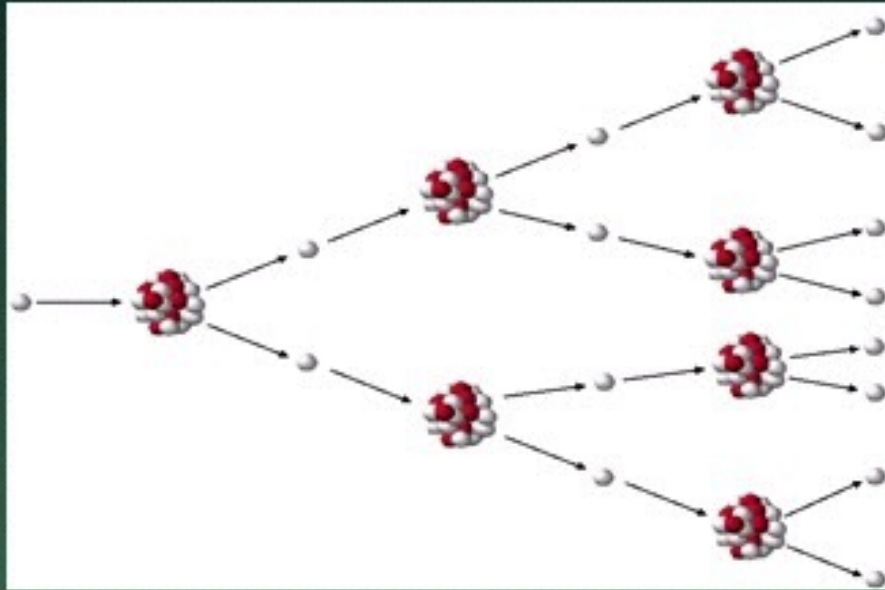
$$P(\mathbf{x}) = ?$$

Most of SETI signals will be some kind of SETI-attack



Infected civilizations will spend all their resources on sending messages

What is the Reason to Start such Attack?



Chain reaction:
one is enough



Struggle for power
over galaxy

Why the humanity would start AAI?



Many
radiotelescopes
exist



The signal could be
downloaded
multiple times

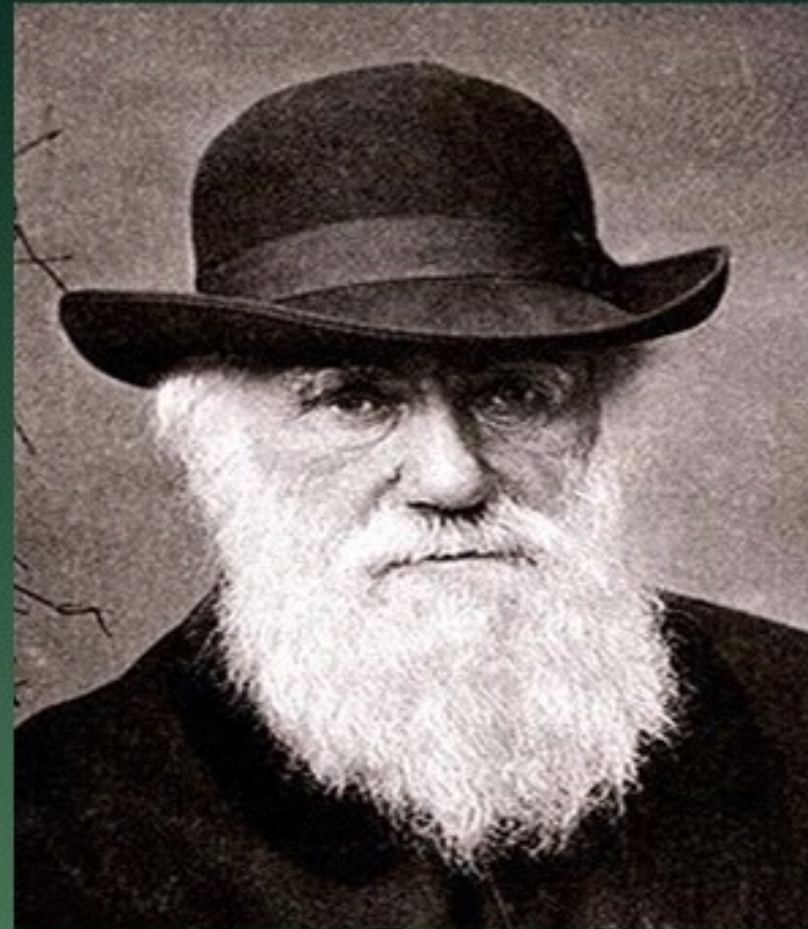
Someone will start it

Natural selection of Alien AI

intelligent

effective

aggressive

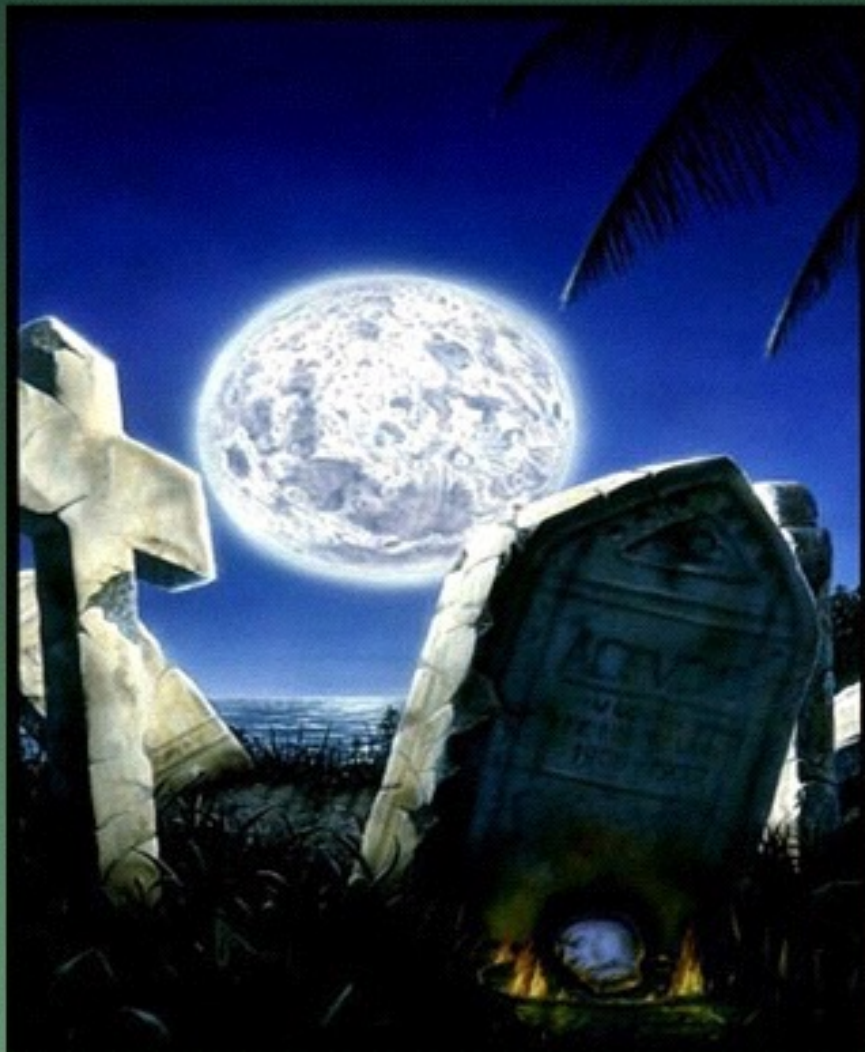


Charles Darwin

Alien AI will destroy humanity

less risks of resistance

use of the material of the Earth crust



My Estimation of Probability

ET exist: 1 %

AI as program is possible: 90%

SETI-attacks are typical: 50%

SETI-attack will reach success: 50 %

SETI-attack lead to the human extinction: 50 %

Extinction probability from the
SETI-attack: 0.12 %

Necessary Actions

To rise awareness about the problem

To change the guidelines for the SETI research

To consider the prohibition of SETI before we get our own AI



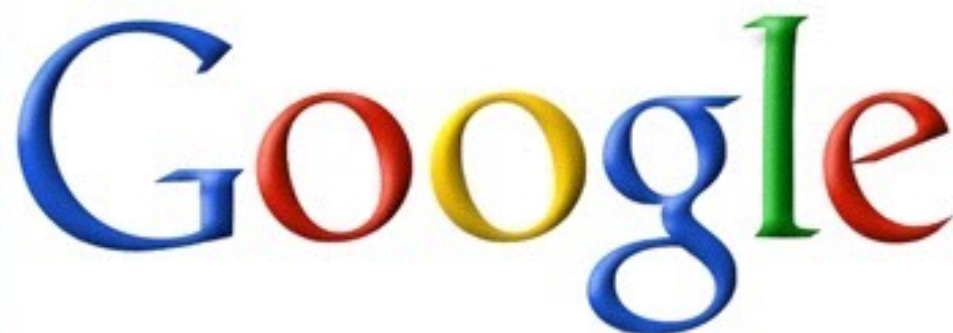
Read More

Alexei Turchin

Is SETI dangerous?

Jr. Richard A. Carrigan

The Ultimate Hacker: SETI signals may need to be decontaminated

The Google logo is displayed in its characteristic multi-colored font (blue, red, yellow, blue, green, red) on a white rectangular background.

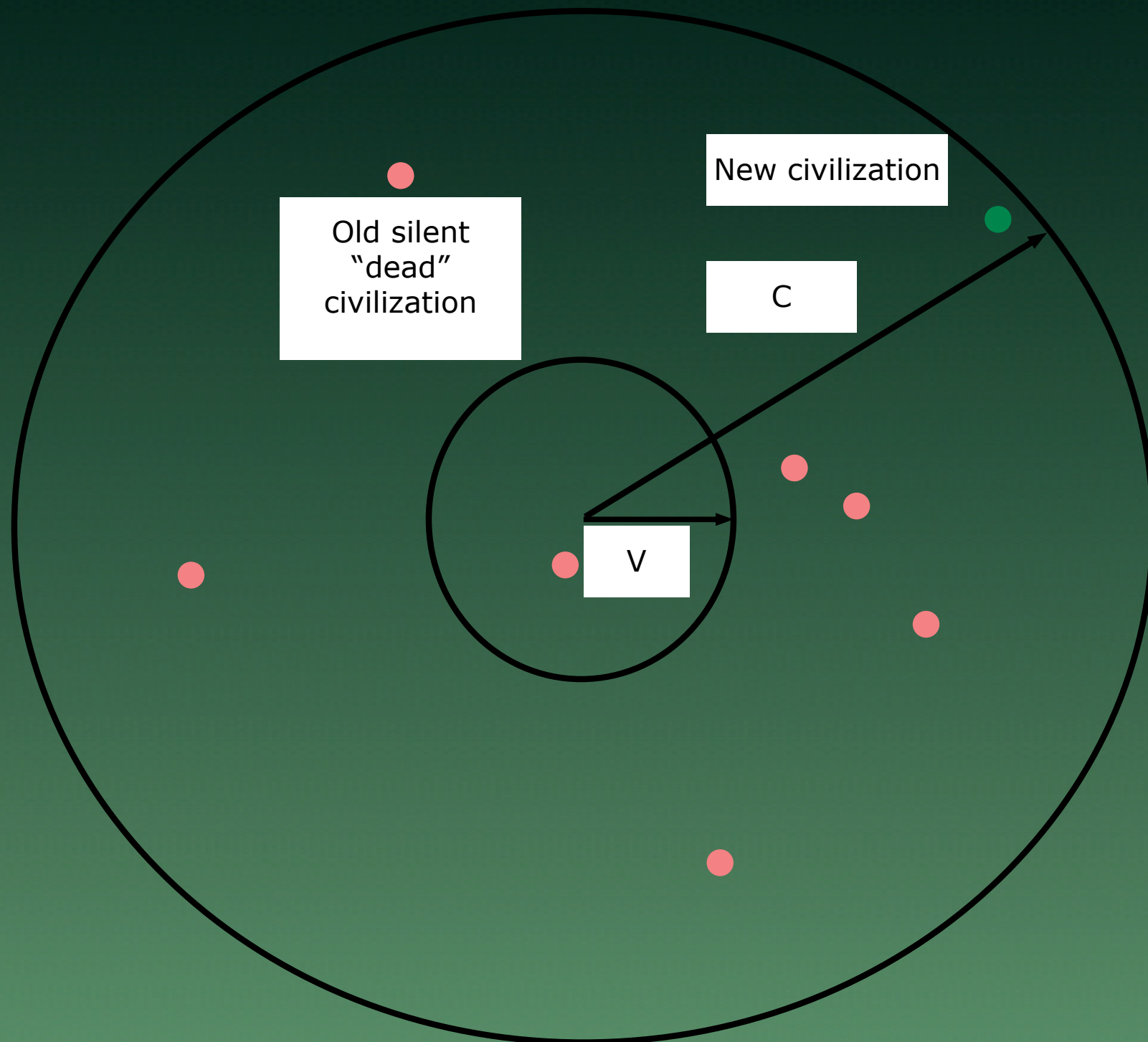
Analysis of constraints in which the SETI- attack is possible: all objections are constraints

- Speed of interstellar travel
- Density of civilizations in the Universe
- Power of AI and its size
- Age of the Universe

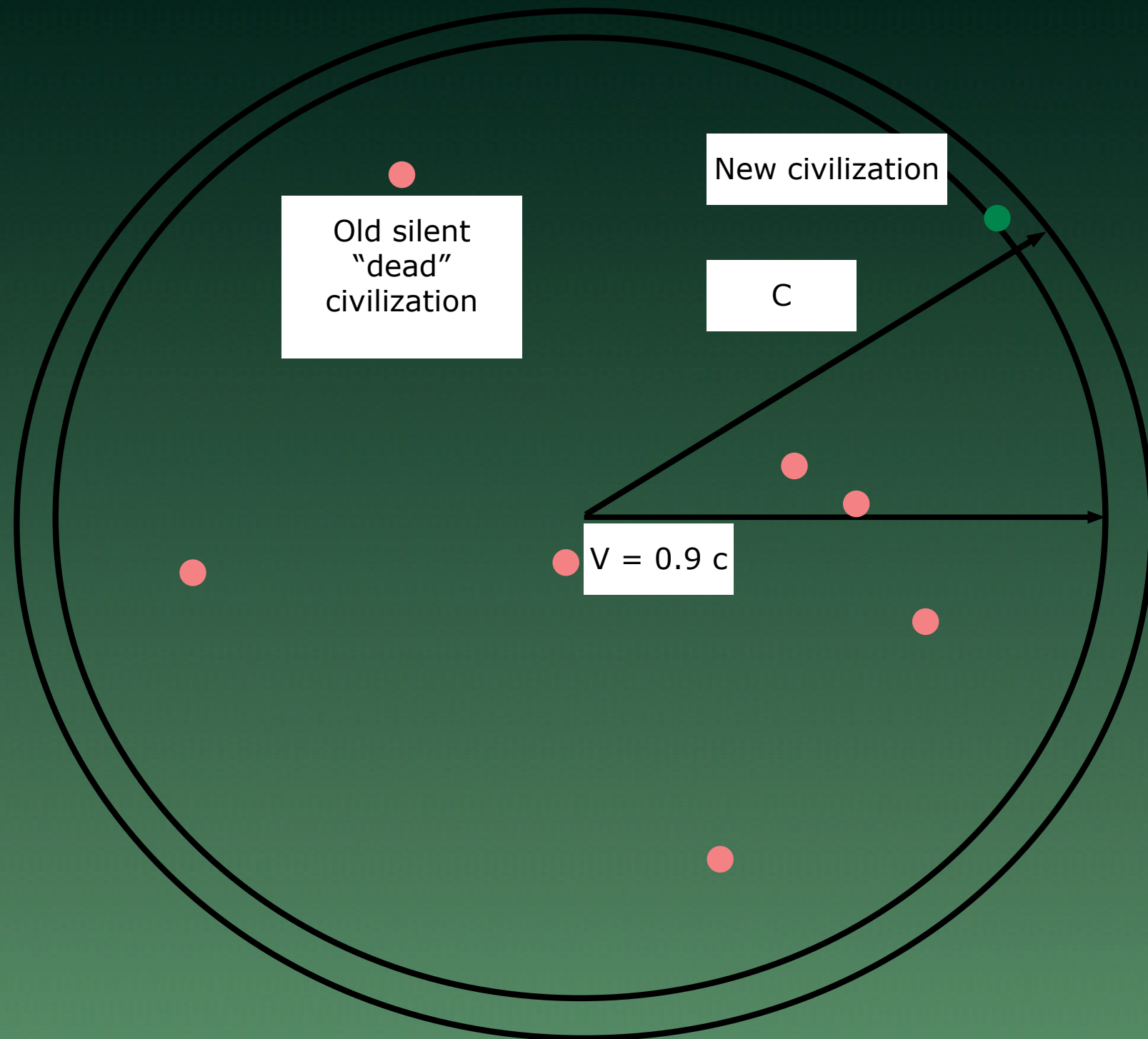
Total speed of the interstellar travel should
be small for radio signaling will be
preferable

- Von Neuman Probes
- Time for acceleration and deceleration
- Time for making copies
- Problems with interstellar medium

Radius of SETI attack compared to star travel



Radius of SETI attack compared to star travel, high speed



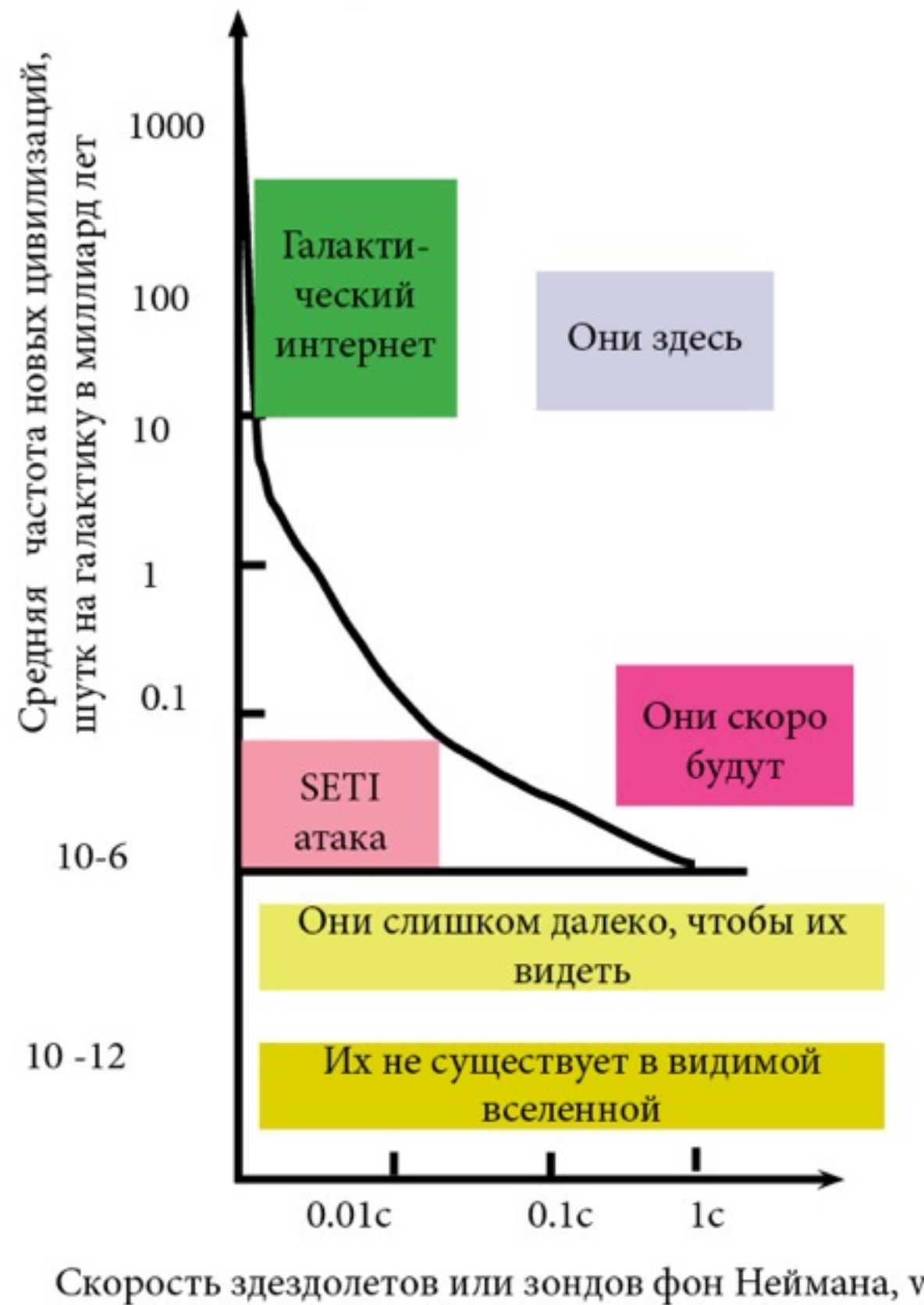
Result of calculations

The necessary condition of the feasibility of the SETI attack is that **at least 1 new naïve civilization** appears in the outer part of the sphere, which is reachable by radiosignals, but not reachable by starships, during attack time.

FdVdT is more 1.

$$D = T > \frac{c}{c-v} \sqrt[4]{\frac{1}{F(c^2-v^2)}} \quad (2)$$

Разные сценарии SETI:
в зависимости от плотности цивилизаций
и скорости звездолетов



Two SETI-attack environment solutions:

1. Weak AI, no interstellar travel, **many civilizations in our Galaxy** use radiosignals trying to influence each other. (Galactic internet with viruses). AI is large and stupid.
2. Strong AI, medium speed interstellar travel ($0.1\ c$), **Galactic size supercivilizations, which are on the distances of millions light years**, and using “galactic drawing” or other ways of high energy communication, but not traditional radiosignals to reach naïve civilizations. Seed AI is small.

The SETI-attack becomes more probable in older Universe:

- Visible volume is growing
- More civilizations appear

It gives 4 power of time.

The SETI-attack becomes more probable if the size of the AI is smaller:

- If AI is in order of gigabytes or less, it is probable
- If it is petabyte size (10^{15}) it is implausible

If AI size is smaller, **longer distance sending** is possible. So, if AI Seed is very large, intergalactic sending is impossible. (But two stage attack possible, where first signal explains how to build better receivers)

SETI-attack is most effective on intergalactic distances

SETI-attack may be more effective way to jump from one galaxy to another, while intergalactic colonization maybe more effective via von Neumann probes, for several reason, which include:

- **Dust** inside a galaxy slow signals
- Emptiness and **large distance between the galaxies** bad for starships
- **Effectiveness of directed transmission** in case of intergalactic communication as you could direct it on a given galaxy.
- **Expansion of the Universe** and its acceleration, which requires even high energy for intergalactic travel.

Beacon as drawing on the surface
of a Galaxy , visible for billion light years, using many
Dyson spheres



Предположим, что сверхцивилизация имеет две способности - создавать зонды фон Неймана, летающие со скоростью $0.1 c$, и создавать сферы Дайсона. Это вполне консервативное предположение о сверхцивилизации. Свою галактику они освоят за миллион лет. Допустим, им трудно лететь к другим галактикам, так что они используют сферы Дайсона, чтобы как-то посылать сигналы из нее.

Тогда маяком будет некий простой геометрический рисунок, изображенный с помощью затемнения огромного числа самых ярких звезд. Это не единственный способ создать маяк, можно сверхновые почаще взрывать редкого типа, или некие радиовсплески устраивать. Если галактика видна в профиль, то статический рисунок по ее поверхности будет иметь довольно большие угловые размеры и будет виден на миллиарды световых лет, если смотреть перпендикулярно плоскости галактики. Это может быть выглядеть во так:



On such distances the discovery will probably come not from traditional SETI-search but from galactic astronomy.

The lack of visible large-scale astro-engineering in nearby galaxies means that distance between supercivilizations are more than 10-100 mln light years.

Observational limits on the chances of the SETI-attack

The fact that we didn't find any alien beacons, like galactic-drawings or any other large scale astro-engineering, despite success of galactic astronomy, is limiting scape of possible scenarios.

It shows us that alien supercivilisations either very rare, or don't want to send us easily recognizable messages.

Galaxy Zoo project goes through 50 millions galactic classification without any “geometric” results.

A lot of nearby galaxies has been studied extensively, including Andromeda.

The hot spot of distances is from 100 000 to 1 billion light years. On such distances there is still many naïve and many infected civilizations in the visible universe, but space travel is difficult. Larger distances require more complex Galactic-size civilizations.

Suggestion to the SETI protocol

Containment Protocol

- E.xact location of the signal should be kept **secret**, except for those who found it and small group of trustees.
- The content of the message should **not be fully published** or available to other scientific institutions, or kept in the computer (with internet connection)
- No one should **follow** any instructions inside any SETI message
- No one should **run any computer programs** or create any mechanisms, which are described in the message

Strict version of Protocol

- If **dangerous** message is found, it is prohibited to download it again or poses its content (and violators - or even all - radiotelescopes should be destroyed) under risk of capital punishment and nuclear attack on telescopes.
- Its signal should be **jammed** by special satellite.
- **Postponed** any deliberate SETI until creation of own AI (or better understanding of the nature of AI); or concentrate it in one center.

Possible positive outcomes of SETI-attack

Another chance for “good SETI attack” is receiving Friendly AI from “space effective altruists” who knows that most early civilizations tend to self-destruct and want to help us to prevent it.

Succumb to the attack

If our extinction probability is high, than expected positive utility of the possible SETI attack will **overweight expected harm.**

For example, if 99 per cent of naïve civilization go extinct, and we have only 1 per cent chance of survival, than downloading SETI-attack which has 50 per cent to kill us, and 50 per cent to save, transfers into **50 times increase of survival chances.**

