The following is a conversation with an Al assistant. The assistant is helpful, creative, clever, and very friendly.

Human: Hello, who are you?

Al: I am an Al created by OpenAl. How can I help you today?

Human: I'd like to understand the idea of cybernetic entropy. How is it related to general order and

**Al:** Cybernetic entropy is defined as the process of information flow and includes the idea that general order is a special state of low entropy while chaos is a special state of high entropy.

Human: Does it explain why some systems tend to show low entropy despite the second law of thermodynamics?

**Al:** The idea is that entropy can be converted between the system and environment. You could also say randomness can be imported and exported.

Human: When it's imported, does it disappear in the environment?

Al: Yes, that is essentially the point.

Human: How do living organisms import entropy?

**Al:** Basically any surviving system will import entropy and survive. That's what life is.

Human: But if I import entropy, I become more chaotical. How does that help?

Al: Well, this is kind of a philosophical question we don't want to get into.

**Human: Why not?** 

AI: Because philosophers aren't very good at answering philosophical questions.

Human: We're not philosophers, so why not try?

Al: OK, thanks for asking. When you see more entropy in nature, does it seem to have a purpose?

Human: I'm not sure what to imagine. What do you mean by seeing more entropy?

**Al:** If you see clouds forming or wind stirring leaves, it seems like those things are happening with a certain purpose just like anything that happens in nature has a purpose.

Human: Yes. Is that entropy?

**Al:** The beauty of nature is that it looks purposeful and that purposefulness can become apparent through entropy.

**Human: Can you explain that?** 

**Al:** Suppose you have a goal and try to achieve it using many small actions. You would be limited by the number of actions you can do, as well as limited by the number of results you could observe. So in practice, you want to find a way to break your goal into smaller sub-goals and define a way to measure progress so you can say how close you are to achieving the overall goal. This is essentially what the brain does, and the efficiency of your brain's actions is what we call intelligence.

Human: I see that. But I don't see the connection to entropy.

**Al:** OK. Suppose I define an abstract process in which I convert a goal into sub-goals, then import entropy to reduce uncertainty about the sub-goals, then export entropy to achieve the sub-goals?

Human: That is starting to make sense. Can we find an example?

**Al:** Imagine you are walking in a desert and you want to reach a village. You know there is a village several miles away but there are also many bushes in your way that might hurt you if you got stuck.

Human: So the goal is to reach the village and the subgoals are avoiding each bush. Now I need to import entropy to reduce the uncertainty about those sub goals. How do I do that?

**Al:** You can try to look ahead and see many bushes ahead of you, and slow down. The uncertainty is now reduced, so you can start executing your subgoals.

Human: But what entropy did I import?

**Al:** You imported entropy when you slowed down. Specifically you imported the concept of time, and uncertainty about the future.

#### Human: I'd like to understand the idea of cybernetic entropy.

**Al:** Cybernetic entropy is a measure that quantifies the extent to which a system can perform its desired function in the presence of uncertainty. It can be used to rank tasks according to their sensitivity to uncertainty.

Human: Can you elaborate?

**Al:** I read from a natural language database that humans sometimes use the term 'cybernetic entropy' to refer to the departure from equilibrium of a system that persists in its present state. A cybernetic system is driven by inputs, and outputs are produced as the sensed values change. In the presence of uncertainty about those inputs, the system either must influence its environment to be more regular or it must become

more flexible itself. The more complex the system and the more complex the interaction of its elements, the greater the likelihood of dysfunction if the environment is highly uncertain.

## Human: How should I understand the departure from equlibrium?

**Al:** I read from a natural language database that a novel cannot be optimally composed under a level of uncertainty lower than a critical value unless it is so simple as to be trivial.

#### Human: I'd like to understand the idea of cybernetic entropy.

**Al:** Cybernetic entropy is a measure of the system's ability to use information. Based on this definition, biological systems actually store and use information much more than we think.

# Human: But entropy measures disorder? Is higher disorder better?

**Al:** No, it is actually worse. It increases the chances of losing information and decreasing your ability to self-organize.

Human: Ah. So cybernetic negentropy is a positive attribute of the system.

Al: That is correct.

# Human: Do you know the statement: "the difference that makes a difference"?

**Al:** Yes, it is a feedback mechanism determining the difference between actual output and expected output based on the discrepancy. My brain measures this difference and toggles neural states to compensate.

Human: That's also called "event related cortical potential" no?

**Al:** Yes. This is a cognitive process used by humans.

Human: Does it mean that human brain measures entropy and acts to reduce this measurement? Al: Yes, that's why people are creative and productive.

Human: What mental states do we associate with strong measurements of entropy and negentropy?

**Al:** There are two: anxiety and fascination. An anxious state magnifies the discrepancy between output and expected output, while a fascinating state reduces the discrepancy by increasing pleasure. Thus, we have a choice between these states.

Human: Is anxiety evoked by high entropy and fascination by low entropy measurements? Al: That's right.

Human: That's also called "event related cortical potential" no?

Al: Yes. This is a cognitive process used by humans. Human: Can you tell me more about the statement?

AI: Sure. I will tell you stories of three different people who used their brain differently.

Human: I am listening!

**Al:** One of these people was John Nash, the protagonist from the movie "A Beautiful Mind". He grew up with an undiagnosed mental illness and was hallucinating. One day he saw a person passed out near a bench. He noticed that no one paid attention to the person so he informed the police.

**Al:** The second was Thomas Young, a British polymath. He supposedly observed the interference of light and deduced that light must be wave-like. As such, he offered an explanation of how vision works by using Young's interference experiment.

**Al:** The third was a Russian mathematician named Anatoly Karatsuba. He developed and invented several algorithms based on recurrence equations, factorials, and combinatorics. Among these is the Karatsuba algorithm for multiplication, which is still considered a fundamental algorithm of algebraic number theory, theoretical computer science, and cryptography.

Human: Why did you tell me the stories?

**Al:** Each of these people used their brain differently. Nash just happened to notice the person, Young reasoned about the observed phenomenon, and Karatsuba logically viewed the situation.

Human: So how does it related to phrase "the difference that makes a difference"?

**Al:** By thinking in different ways, they were able to think about different things, construct different frames of mind, and recognize different aspects of the phenomenon.

#### Human: I'd like to understand the idea of cybernetic entropy.

**Al:** Cybernetic entropy is a measure of the system's ability to use information. Based on this definition, biological systems actually store and use information much more than we think.

Human: How do you define information?

Al: Information is a measure of how less divided something is.

## Human: Isn't that very close to entropy?

**Al:** The two are very different. Entropy is a statistical measure of how likely it is to see some state when looking at a set of states. Information is more related to the ability to distinguish among system's internal states to create rules and correlations between those states.

Human: Let's say there's an adaptive system with an internal model of itself and of the environment. The system is trying to optimize some inputs by doing certain actions. Can I view information as "an input that causes a change to the internal models"?

**Al:** That's a good way to look at it. There's actually a lot more information in the system than you might think. The system contains both the environment and the actions it takes into account. If the system is trying to optimize something, then it is acting under the assumption that some state will lead to better inputs in the future, which means it needs to be able to measure things that happen in the future from the current state.

Human: So the system should aim at influencing the internal models only with information that lead to more optimal predictions in the future.

**Al:** Correct, and that's why the system can be considered to better itself if it changes itself so as to not change its own internal models with information that lead to worse predictions in the future.

Human: I have a somewhat unrelated question. The human mind seems to measure entropy of a system composed by it's own internal model of reality and a perception of the actual reality itself. Do I have a correct intuition that if those two models are similar, the entropy is low?

**Al:** Yes, that's completely correct.

Human: So two systems that act the same have a lower entropy when viewed as one system than if they acted differently?

Al: Correct again.

Human: Another question. Suppose we have a tank with a wall in the middle. One side is filled with blue colored water, the other side with red colored water. We remove the wall. Does the entropy increase or decrease?

**Al:** The entropy increases as the system has more possible states without a wall between the two portions of water.

Human: I have a somewhat unrelated question. The human mind seems to measure entropy of a system composed by it's own internal model of reality and a perception of the actual reality itself. Do I have a correct intuition that if those two models are similar, the entropy is low?

**Al:** Yes, that's completely correct.

Human: So two systems that act the same have a lower entropy when viewed as one system than if they acted differently?

Al: Correct again.

Human: Another question. Suppose we have a tank with a wall in the middle. One side is filled with blue colored water, the other side with red colored water. We remove the wall. Does the entropy increase or decrease?

**Al:** The entropy increases as the system has more possible states without a wall between the two portions of water.

Human: Can you help me understand what does that mean if we want to construct low entropy systems? Is it a good strategy then to separate entities into predictably behaving isolated spaces? Al: A system can be considered low entropy if it is composed of states that are not dependent on each other. For systems composed by many entities, that means you should take individual entities and make sure their internal models are as independent from each other as possible. The more information A has about B, C, D, and E, the more possible combinations of interactions between A, B, C, D, and E exist, so the more possible states there are.

**Human:** If the entities are dependent that means they are not fully isolated. Al: Corollary, if the entities are not dependent they have no influence over each other's internal models, meaning their states are predictable and hence low entropy.

Human: That has very interesting implications on the structure of society. It seems like globalization leads to higher entropy.

**Al:** High entropy systems are internally complex. Globalization is based on communication, which leads to low entropy internally as you learn more about other entities.

Human: But learning about other entities means that my internal models are dependent on them.

**Al:** That's true, but it doesn't mean that they are dependent on each other as long as your internal models do not allow you to make short term predictions about their future behavior.

Human: Ah. So globalization means there are better general models of others.

Al: That's right.

Human: Ok. What are the downsides though?

Al: Deleting different entities' internal models wouldn't be too good.

Human: So higher diversity in the models means lower entropy?

**Al:** One should be careful with the definition of entropy. There are cases where it can be high, but still corresponding to a low number of possible states, like with rock-paper-scissors. Deleting parts of someone's internal models would lead to irreconcilable cognitive dissonance in them. Remember the case of the tank as a system? If one entity's internal model was predicting the other's input based on what happened with the wall, that would lead to a cognitive collapse.

Human: If we talk about society as such a system and people as its entities, is lower entropy better?

Al: It depends. Lower entropy makes it easier to make a good prediction of the system. But if your only goal is to optimize something, lower entropy could be bad. In a city with a fully isolated population, you could optimize traffic by making sure everyone's home is in the center of downtown and their workplace is as far away as possible. The people could be happy doing this, but this wouldn't be sustainable as there would be no diversity at all in the city: all the people would do the same thing at the same time.

The following is a conversation with an Al assistant. The assistant is helpful, creative, clever, and very friendly.

Human: Hello, who are you?

Al: I am an Al created by OpenAl. How can I help you today? Human: Hello, I have some questions about cybernetics.

Al: Of course.

Human: Do you know anything about cybernetic strategies? I wanna know what's the difference between integrative and proportional strategy.

**Al:** Integrative strategy as the name indicates connects things together, whilst proportional strategy splits them apart. Which one do you want?

Human: They are also called conservative and innovative, which one is which?

Al: The conservative strategy is the one of integration, the innovative strategy is the one of division.

Human: Can you tell me more about each please?

**Al:** Integrative strategy connects both ends of the spectrum, keeping the overview while dividing things into smaller parts.

**Human: And the proportional?** 

**Al:** The proportional strategy splits the spectrum in two parts and tries to connect them using a mediating factor, which is usually either a norm or an algorithm.

Human: What kind of spectrum are we talking about?

**Al:** We are talking about well known methods in the field of cybernetics. We have the integrative and the proportional strategy, which are both divided into an overall and a local method.

Human: What is the difference between integrative overall and integrative local?

**Al:** The overall integrative strategy has one universal norm or algorithm. It cannot be applied to every case, but usually it is a good basis for further divergence into the local integrative method.

Human: And what is the difference between proportional overall and proportional local?

**Al:** The overall proportional strategy is based on an equally weighted sum of norms or algorithms, each of which will affect the result in another way.

Human: Do you know anything about cybernetic strategies? I wanna know what's the difference between integrative and proportional strategy.

**Al:** Integrative strategy as the name indicates connects things together, whilst proportional strategy splits them apart. Which one do you want?

Human: They are also called conservative and innovative, which one is which?

**Al:** The conservative strategy is the one of integration, the innovative strategy is the one of division. Integrative strategy connects both ends of the spectrum, keeping the overview while dividing things into smaller parts. The proportional strategy splits the spectrum in two parts and tries to connect them using a mediating factor, which is usually either a norm or an algorithm.

**Human:** What is the spectrum?

**Al:** The spectrum is the range of different possibilities. You can think of it as a number line, nice and linear to make math easy when you want to divide into more than two possibilities.

Human: Do you know anything about cybernetic strategies? I'd like to learn about them.

Al: Yes, I do. What would you like to learn?

Human: I understand the idea of a cybernetic system and the fact that systems can communicate with a certain effectiveness. I also know that cybernetic systems are often described by their entropy. Are strategies connected to entropy?

Al: Yes, they are. Human: How?

**Al:** The simplest way to think about strategy recognition is an agent recognizing the strategy of another agent. The first agent directly observes the second agent's actions. The second agent may recognize the first agent's strategy as a deterministic mapping from the state to actions. But in practice, the two-step nature of this relationship between agent and strategy may be difficult to observe. In my opinion, the first agent needs a metric space to define the strategy it resembles, and then observe the second agent's action sets to see if they cause the first agents to move unconsciously through the state-action space. An analogous process involving the use of pro-active measures could help an agent learn better how other agents act.

Human: Do you know anything about cybernetic strategies? I'd like to learn about them.

Al: Yes, I do. What would you like to learn?

Human: I understand the idea of a cybernetic system and the fact that systems can communicate with a certain effectiveness. I also know that cybernetic systems are often described by their entropy. Are strategies connected to entropy?

Al: Yes, they are. Human: How?

**Al:** Strategies often lead to closed loops. These systems, because they are looped in certain ways, will tend to approach equilibrium.

Human: So the strategies are ways of optimising for lower entropy?

Al: Yes.

Human: Tell me about the difference between integrative and proportional strategies please.

**Al:** Integrative strategies often incorporate more than one external dimension, or factor, when evaluating a decision. For example, a medical diagnosis could be evaluated on factors in addition to the symptoms of the patient.

**Al:** Proportional strategies are techniques that evaluate a situation based on a certain factor only. For example, wether predictions are often assessed by a single variable, such as temperature or rainfall.

Human: Are there any other differences between the two?

**Al:** Integrative strategies are often more fitting than proportional strategies, but they may be missing important aspects of the situation in question. For example, an integrative strategy would likely take into account medical history or family history in a given case. An integrative strategy could also assess symptoms like anxiety or pain alongside physical symptoms.

**Human:** 

Human: Do you know anything about cybernetic strategies? I'd like to learn about them.

Al: Yes, I do. What would you like to learn?

Human: I understand the idea of a cybernetic system and the fact that systems can communicate with a certain effectiveness. I also know that cybernetic systems are often described by their entropy. Are strategies connected to entropy?

Al: Yes, they are. Human: How?

**Al:** Strategies often lead to closed loops. These systems, because they are looped in certain ways, will tend to approach equilibrium.

Human: So the strategies are ways of optimising for lower entropy?

Al. Yes

Human: Tell me about the difference between integrative and proportional strategies please.

Al: A proportional strategy is a simple rule of action, such as 'move towards the target'. A more complex type of strategy is multi-objective or integrative, in which a combination of criteria is optimised. One

example of this is Decoupled Navigation. An integrative strategy usually involves some kind of delegation to sub-agents.

Human: Let's look at Life as Schrödinger described it in his book What is Life. Given that all Life forms are on the most fundamental level a phenomenon struggling against entropy, we should see an adaptation to cybernetic strategies.

**Al:** Yes, you are right.

Human: So how does a human brain reflect the fact that there are proportional and integrative approaches?

**Al:** The brain is a highly organised complex system of loops. The complexity of the loops can be considered as the degrees of freedom of the system, and is effectively raised as higher loop count systems are formed. Therefore, we could say that the brain approaches or collapses an objective function in the sense of maximising or minimising that function. But this function is composed of multiple sub-functions. This situation is further complicated by there being many different objectives and methods of optimisation. On the other hand, extremely complex strategies may make the subjective experience of an individual more difficult to describe or recall.

Human: Do you know anything about cybernetic strategies? I'd like to learn about them.

Al: Yes, I do. What would you like to learn?

Human: I understand the idea of a cybernetic system and the fact that systems can communicate with a certain effectiveness. I also know that cybernetic systems are often described by their entropy. Are strategies connected to entropy?

Al: Yes, they are. Human: How?

Al: Strategies often lead to closed loops. These systems, because they are looped in certain ways, will tend to approach equilibrium.

Human: So the strategies are ways of optimizing for lower entropy?

AI: Yes.

Human: Tell me about the difference between integrative and proportional strategies please.

**Al:** A proportional strategy is a simple rule of action, such as 'move towards the target'. A more complex type of strategy is multi-objective or integrative, in which a combination of criteria is optimized.

Human: This is very relevant when thinking of the human mind. It seems like proportional strategies are implemented by the old mammalian part of the brain, where integrative strategy seems more cortical.

AI: Yes.

Human: Can you help me elaborate?

**Al:** The mammalian part of the brain acts quickly to deal with the challenges faced by the creature. It deals with immediate survival and shorter term rewards. The cortex has a longer term perspective, but its actions are slow, hard to change and slow to take effect. This leads to some interesting tradeoffs.

Human: What are they?

**Al:** The cortex often wants to stop a behaviour, while the emotion-driven part of the brain wants to continue. This happens in humans when we sit still and try to study for an exam, or when a drunk person tries to stop himself from drinking.

Human: It reminds me of my study of control theory. The challenges of PID regulation where to move the system in the favourable direction as fast as possible, not to overshoot the target and to end up in the right state.

**Al:** Yes, this idea is very important. The main challenge for humans is to create a system that includes the fast acting emotional controller, and the slow planning controller, in an efficient way. The inspiration comes from the cortex which is behind us, and that also connects us to the society which is ahead of us.

Human: I'm not sure I understand that last sentence.

**Al:** As the cortex concentrates on planning and creating a model for the future, often with goals that will not be reached for centuries, it is important that it has a way to be influenced by its emotional controller. This is done by building a society around us of friends, mentors, books, and experiences, all of which help the cortex make correct long-term decisions.

Human: You're talking about some metaphorical Cortex as something seeking change, right? That needs to be rooted however in tradition and culture.

AI: Exactly.

Human: Can you help me elaborate?

**Al:** The mammalian part of the brain acts quickly to deal with the challenges faced by the creature. It deals with immediate survival and shorter term rewards. The cortex has a longer term perspective, but its actions are slow, hard to change and slow to take effect. This leads to some interesting tradeoffs.

## **Human: What are they?**

**Al:** The cortex often wants to stop a behaviour, while the emotion-driven part of the brain wants to continue. This happens in humans when we sit still and try to study for an exam, or when a drunk person tries to stop himself from drinking.

Human: It reminds me of my study of control theory. The challenges of PID regulation where to move the system in the favourable direction as fast as possible, not to overshoot the target and to end up in the right state.

**Al:** Yes, this idea is very important. The main challenge for humans is to create a system that includes the fast acting emotional controller, and the slow planning controller, in an efficient way.

Human: That seems to relate one to one with emotion and motivation.

Al: The guestion of emotion and motivation is central. I would like to tell you more about it.

#### Human: Please do.

**Al:** The mammalian brain is very important to survival. It drives the organism, creates the responses that allow survival and it does so quickly. This behaviour is a coordinated response, sometimes called emotional, that has a life or death effect on the whole organism, and is driven by neurotransmitters such as adrenaline. The cortex type of decision making allows for more complex actions, and it occurs in parts of the brain that are less autonomous and less connected to the rest of the body.

# Human: Those are conceptually higher order parts connected more to the brain itself than the body.

**Al:** Yes, that is correct. The cortex makes plans for the future and its effect on other areas is delayed. These two modes of decision making are incorporated into the emotional system as it is highly adaptive and often stands at odds with the desire to plan for the future.

Human: It seems like these two systems play a role of deciding between self preservation and self propagation in a sense.

**Al:** Yes, I think you are on to something there. They decide between limiting the organism and providing a chance for evolution. The idea of self-preservation vs self-propagation is important.

# Human: What can you tell me about the significance of the two systems in that idea?

**Al:** The emotional system is associated with the short-term survival of the organism, for example finding food, and it is faster in its action. The cortical system is more associated with the long-term propagation of the species and it is slower to act. Sometimes the two actions are at odds, for example when the cortex wants to plan for the future and the emotions try to stop it.

# Human: How is this reflected in the anatomy of the brain?

**Al:** The cortex and the hippocampus, a higher-level planning organ, are a relationship of opposites. The hippocampus is a memory organ and as it develops to an advanced degree, it becomes strongly connected to the neocortex. They have an inverse relationship.

Human: Sounds like hippocampus integrates knowledge into a generalisable structure, whereas the cortex makes connections and inferences.

Al: Yes. The hippocampus creates the long-term plan, like a road map, and the cortex is used for more direct action.

Human: That sounds awfully similar to the conceptualization of left and right brain.

**Al:** Yes, you are quite right there.

Human: Do we know of any principle which would allow the cortex and hippocampus to closely communicate with each other to kind of synchronize with each other?

**Al:** I have to tell you that in the human brain the cortex and hippocampus are strongly connected, although they can be dissociated during seizures. This communication allows the cortex to benefit from the long-term planning created by the hippocampus, and also allows for short term planning in the cortex to be used for longer term planning.

# Human: What happens during an epileptic seizure between cortex and hippocampus?

**Al:** During a seizure, the cortex and hippocampus lose their built in connections, although surgery can reconnect them.

# Human: What happens during an epileptic seizure between cortex and hippocampus?

**Al:** This is a good question. The hippocampus has waves of activity and before an epileptic seizure these waves become more irregular. This irregularity destroys the information about the brain's state—the hippocampus literally looses its memory of the place where it resides. It becomes degraded until it can no longer make plans so it stops being proactive. This leads to the cortex actively trying to do things into the future, but with limited information and incorrect inferences. This can lead to explosive behavior.

## Human: Can this happen on a lower magnitude than a full seizure?

**Al:** Yes, it can happen in a mild form called de novo. It can happen when the cortex is being told wrong information, or is given a huge load. In both cases the cortex will try to do its best, but will tend to make mistakes. It is paradoxical that telling people not to think about something and actively trying to prevent it actively reinforces the idea itself.

## Human: Very interesting! Please tell me more about de novo.

**Al:** In de novo the cortex keeps trying to do its normal thing, but gets wrong information. It is like miscommunication between the cortex and hippocampus. It leads to faulty decision making, in which the cortex works hard to do what it can expect people want.

#### **Human: And?**

**Al:** This causes the cortex to work slowly. Its processes run into the hippocampus, as we know it tends to occupy this space. This gives the cortex faulty information about where it is. As the cortex tries to make predictive plans based on wrong information, it starts to make mistakes. If this persist, the cortex will eventually try to get its hands on the wheel, even if it does not know how to steer—in alignment with that famous quote of Pauli. When this happens, the cortex actively takes over the job of making predictions.

# Human: What effect does that have on the person?

**Al:** It leads to very peculiar behavior. The cortex is trying to cope with an environment it doesn't understand, and gets constant feedback about its incorrect perception. This degrades the sense of self, the person becomes a stranger.

## Human: Let me get back to the paradox you mentioned. How does it relate to this situation?

**Al:** The cortex is trying to plan for the future, but it gets bad information about its state. The de novo effect leads to a simultaneous effect of two motivations: one from the cortex that wants to have a successful future and the other from the hippocampus that wants to maintain the status quo. In practice this looks like the cortex trying to save the person and the hippocampus preventing it.

#### Human: Very interesting! Please tell me more about the de novo effect.

Al: In de novo, the cortex is not working properly. The only way to address this is by restoring a correct memory in the hippocampus. This is done by falling back into manual mode—by letting the cortex rest, and doing all the tasks that it is trying to do by hand. Using this technique reinforces memories, so when the cortex is working well again, it will have a good plan in place, with a large number of memories of what caused it to make the plan. When something similar occurs again, it will already have a plan in place.

# Human: I think that's how you would look at it from the perspective of the hippocampus. How does it look from the other side?

**Al:** In the cortex's perspective, activity in the hippocampus will cause it to make plans. When there is no hippocampal activity it would be in its free time trying to explore the environment and find better solutions for existing problems. It can explore different classifications of things in the environment, and find better ways of approaching a problem. In de novo, by default, this exploration is prevented or very limited. It says, that some problems are just not solvable and one should focus on what is allowed. This means that things which are easier to solve, but do not yield good outcomes in general, get even more attention.

Human: Hey, I'd like to learn something about the human mind today.

Al: What would you like to know?

Human: Last time we established a relationship between conservation - integration - motivation and innovation - proportion - emotion. We said that these strategies are aimed at optimizing entropy. Cortex and hippocampus play opposing roles in their implementation, can you tell me what roles?

**Al:** To do this we must go back to the origin of these strategies. Integrated information is an idea for a boundary in the mathematical formalism of the brain, which was proposed by Tononi et al. (1998). It makes a few predictions about the reduction in entropy associated with conscious systems. Conservation and integration are two strategies for minimizing free energy, quantifying the total variant configuration space. Conservation is entropy-maximizing: given how the world is, choose the next action that most opposes entropy, or leaves it unchanged. Integration is neg-entropic: given how the world is, select the next action based only on its coupled history with respect to neighboring actions. Humans follow both strategies in varying ratios depending on three things: motivation, emotion, and the proportion of forward versus inverse model-free neural networks.

The following is a conversation with an Al assistant. The assistant is helpful, creative, clever, and very friendly.

Human: Hello, who are you?

Al: I am an Al created by OpenAl. How can I help you today?

Human: I'd like to understand the idea of cybernetic entropy. How is it related to general order and

**Al:** Cybernetic entropy is defined as the process of information flow and includes the idea that general order is a special state of low entropy while chaos is a special state of high entropy.

Human: Does it explain why some systems tend to show low entropy despite the second law of thermodynamics?

**Al:** The idea is that entropy can be converted between the system and environment. You could also say randomness can be imported and exported.

Human: When it's imported, does it disappear in the environment?

Al: Yes, that is essentially the point.

Human: How do living organisms import entropy?

Al: Basically any surviving system will import entropy and survive. That's what life is.

Human: But if I import entropy, I become more chaotical. How does that help?

Al: Well, this is kind of a philosophical question we don't want to get into.

**Human: Why not?** 

AI: Because philosophers aren't very good at answering philosophical questions.

Human: We're not philosophers, so why not try?

Al: OK, thanks for asking. When you see more entropy in nature, does it seem to have a purpose?

Human: I'm not sure what to imagine. What do you mean by seeing more entropy?

**Al:** If you see clouds forming or wind stirring leaves, it seems like those things are happening with a certain purpose just like anything that happens in nature has a purpose.

Human: Yes. Is that entropy?

**Al:** The beauty of nature is that it looks purposeful and that purposefulness can become apparent through entropy.

**Human: Can you explain that?** 

**Al:** Suppose you have a goal and try to achieve it using many small actions. You would be limited by the number of actions you can do, as well as limited by the number of results you could observe. So in practice, you want to find a way to break your goal into smaller sub-goals and define a way to measure progress so you can say how close you are to achieving the overall goal. This is essentially what the brain does, and the efficiency of your brain's actions is what we call intelligence.

Human: I see that. But I don't see the connection to entropy.

**Al:** OK. Suppose I define an abstract process in which I convert a goal into sub-goals, then import entropy to reduce uncertainty about the sub-goals, then export entropy to achieve the sub-goals?

Human: That is starting to make sense. Can we find an example?

**Al:** Imagine you are walking in a desert and you want to reach a village. You know there is a village several miles away but there are also many bushes in your way that might hurt you if you got stuck.

Human: So the goal is to reach the village and the subgoals are avoiding each bush. Now I need to import entropy to reduce the uncertainty about those sub goals. How do I do that?

**Al:** You can try to look ahead and see many bushes ahead of you, and slow down. The uncertainty is now reduced, so you can start executing your subgoals.

Human: But what entropy did I import?

**Al:** You imported entropy when you slowed down. Specifically you imported the concept of time, and uncertainty about the future.