بسم الله الرحمن الرحیم

نظریه علوم کامپیوتر

نظریه علوم کامپیوتر - بهار ۱۴۰۰-۱۴۰۱ - جلسه نهم: خودتولیدکنندگی

Theory of computation - 002 - S09 - self-reproducibility

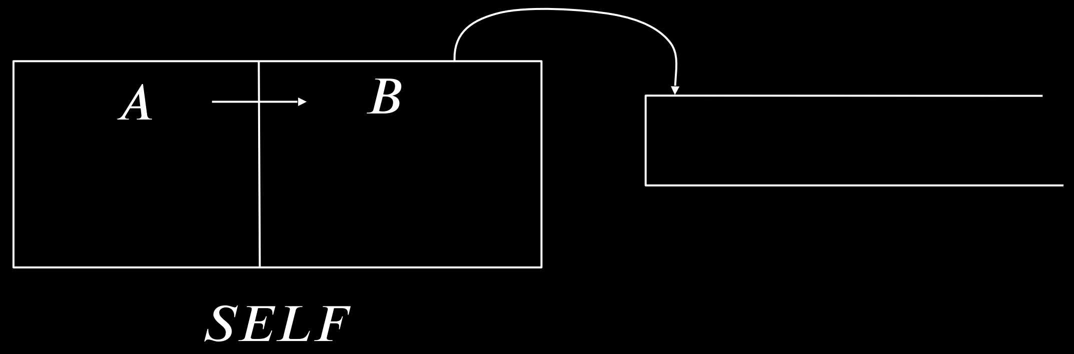
Recall

Recall

* Self-Reproducing TM

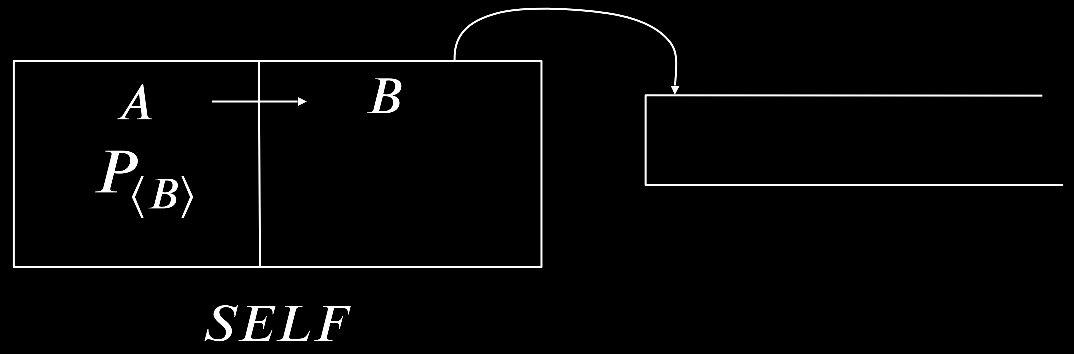
Recall

* Self-Reproducing TM



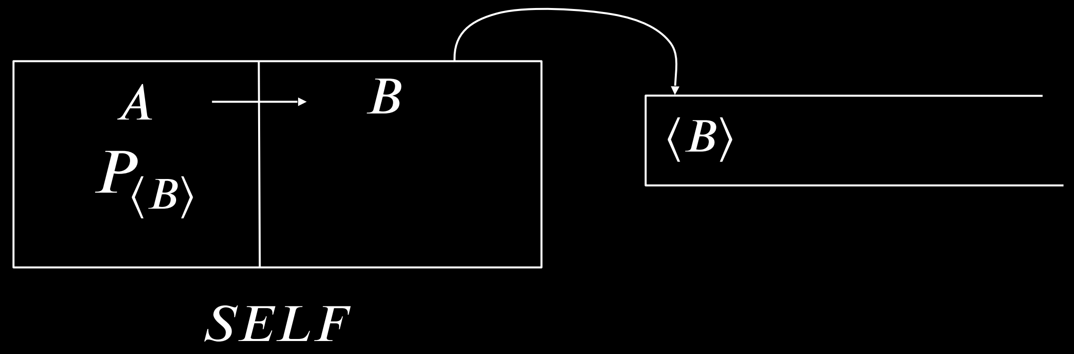
Recall

* Self-Reproducing TM



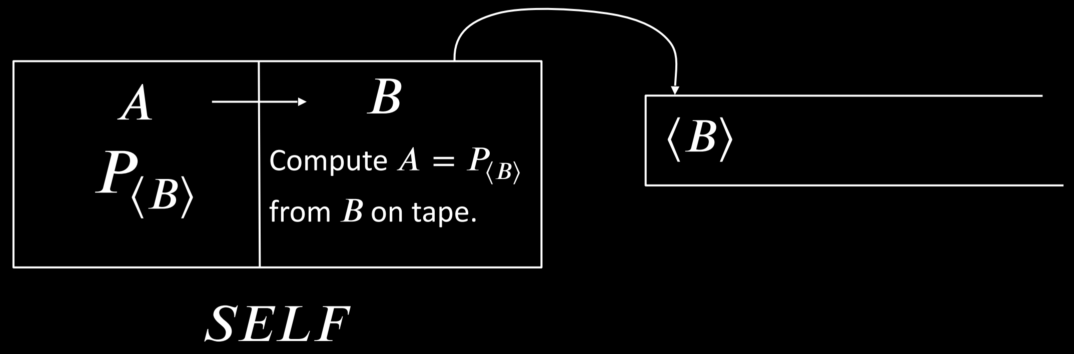
Recall

* Self-Reproducing TM



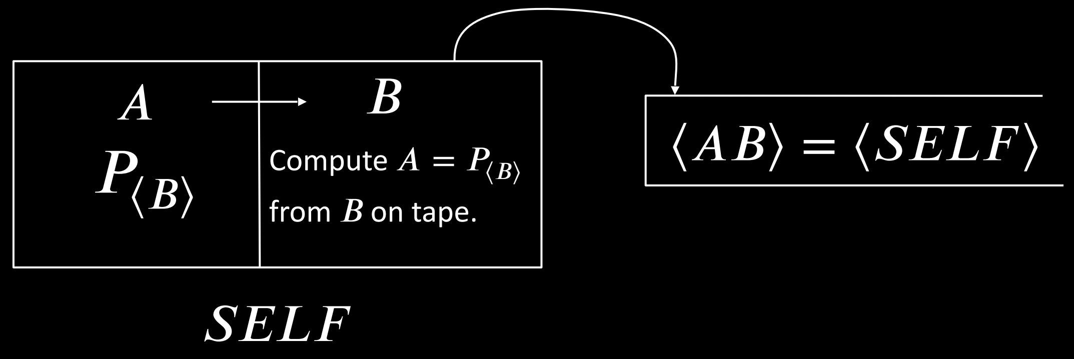
Recall

* Self-Reproducing TM



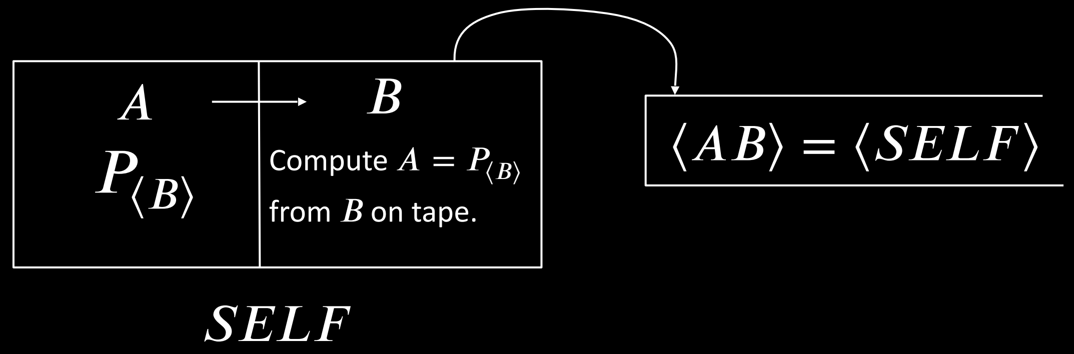
Recall

* Self-Reproducing TM



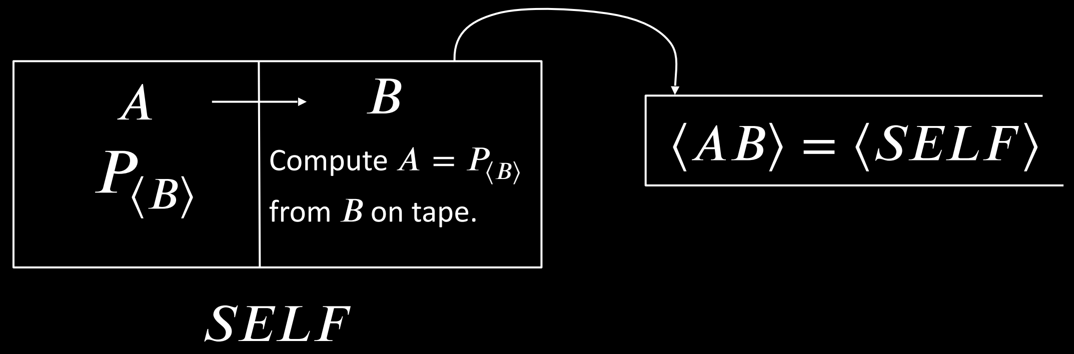
Recall

* Self-Reproducing TM
* The Recursion Theorem: TM T working on <w,A> -> TM R
  * that has access to <R>,
  * behaves like T on <w,R>



Recall

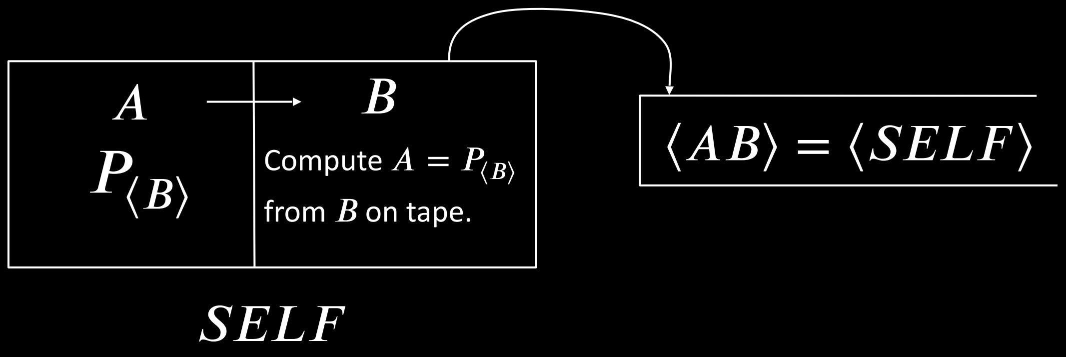
* Self-Reproducing TM
* The Recursion Theorem: TM T working on <w,A> -> TM R
  * that has access to <R>,
  * behaves like T on <w,R>



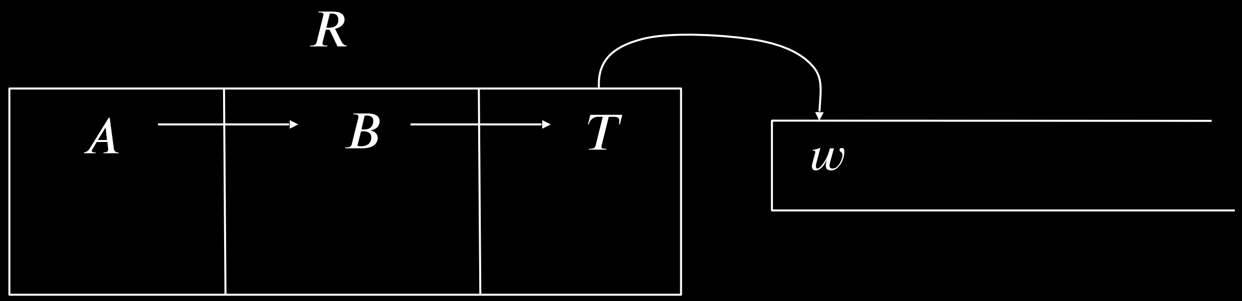
Theorem: For any TM T there is a TM R where for all w R on input w operates in the same way as T on input $\langle w, \ R \rangle$.

Recall

* Self-Reproducing TM
* The Recursion Theorem: TM T working on <w,A> -> TM R
  * that has access to <R>,
  * behaves like T on <w,R>

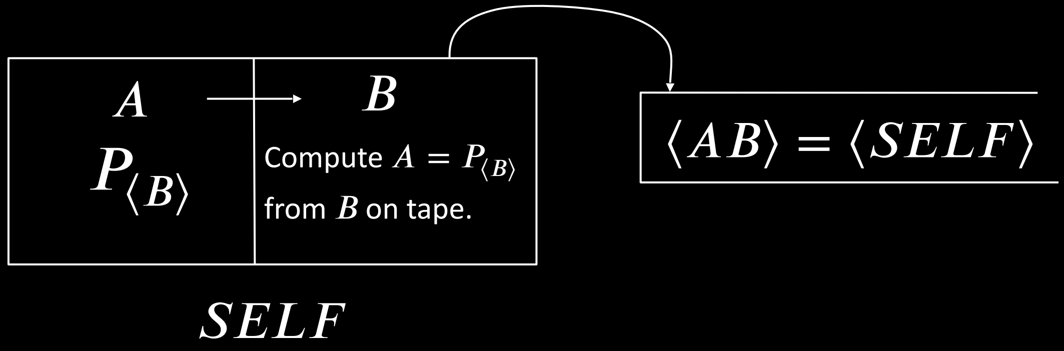


Theorem: For any TM T there is a TM R where for all w
R on input w operates in the same way as T on input $\langle w, \ R \rangle$.

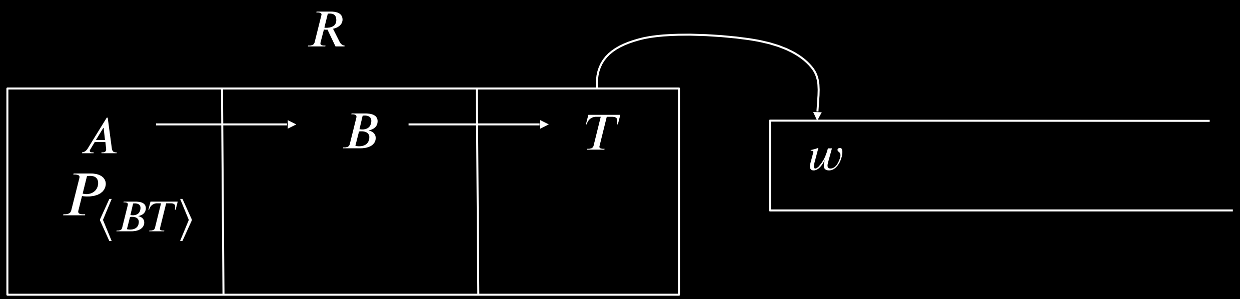


Recall

* Self-Reproducing TM
* The Recursion Theorem: TM T working on <w,A> -> TM R
  * that has access to <R>,
  * behaves like T on <w,R>

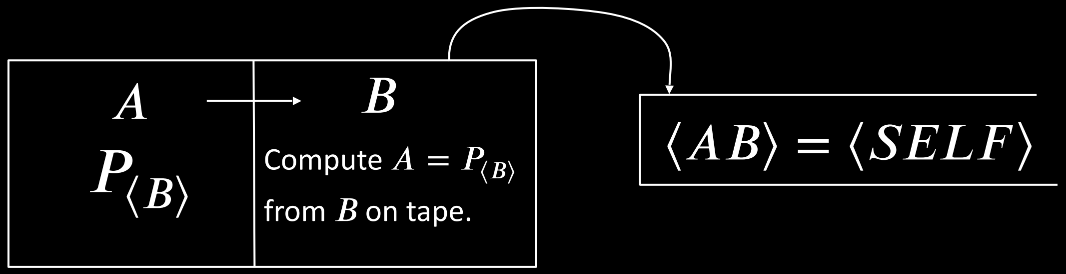


Theorem: For any TM T there is a TM R where for all w
R on input w operates in the same way as T on input $\langle w, R \rangle$.



Recall

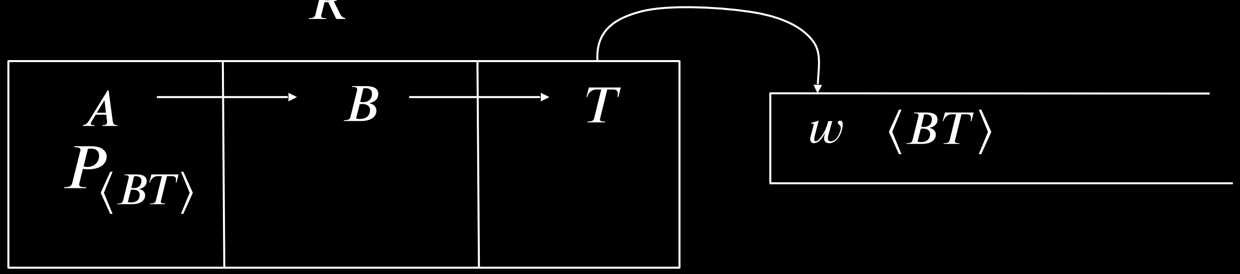
* Self-Reproducing TM
* The Recursion Theorem: TM T working on <w,A> -> TM R
  * that has access to <R>,
  * behaves like T on <w,R>



$SELF$

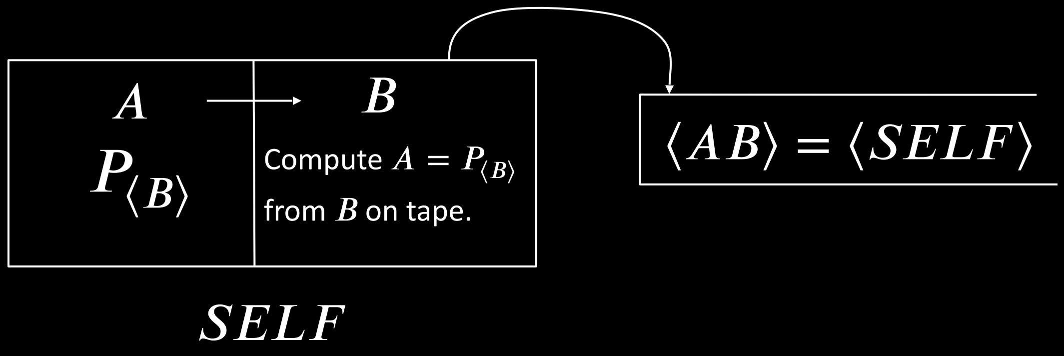
Theorem: For any TM T there is a TM R where for all w
R on input w operates in the same way as T on input $\langle w,\ R\rangle$.

R

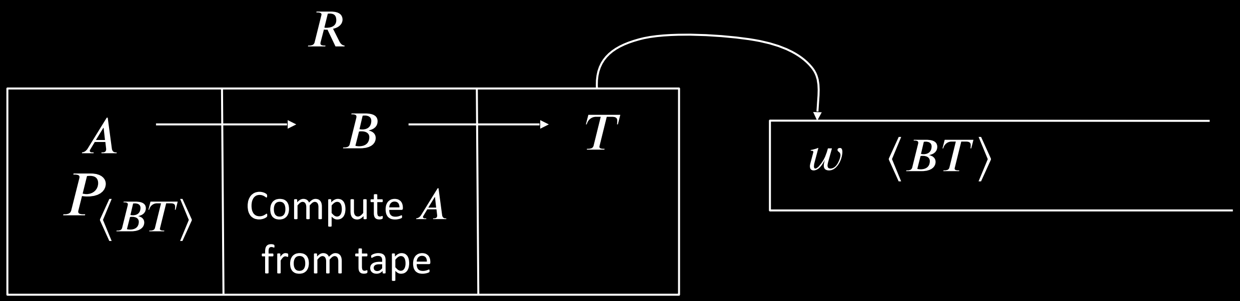


Recall

* Self-Reproducing TM
* The Recursion Theorem: TM T working on <w,A> -> TM R
  * that has access to <R>,
  * behaves like T on <w,R>

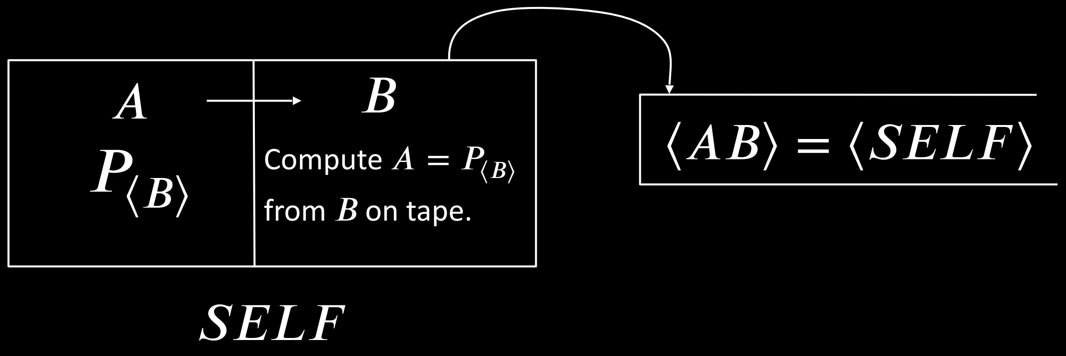


Theorem: For any TM T there is a TM R where for all w R on input w operates in the same way as T on input $\langle w,\ R\rangle$.

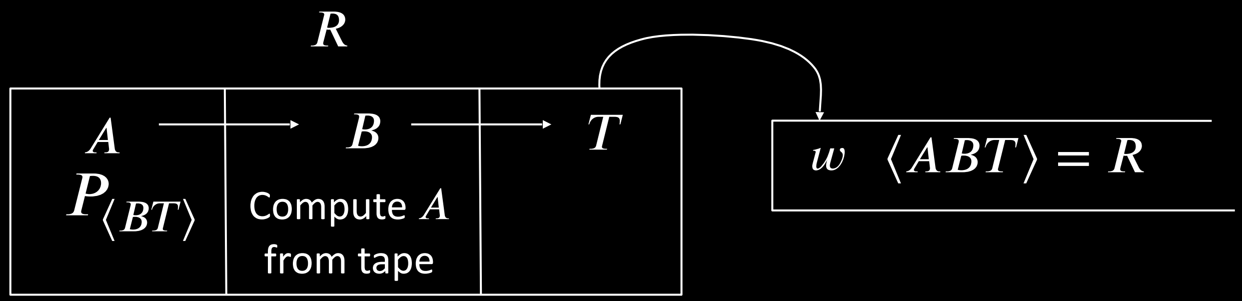


Recall

* Self-Reproducing TM
* The Recursion Theorem: TM T working on <w,A> -> TM R
  * that has access to <R>,
  * behaves like T on <w,R>

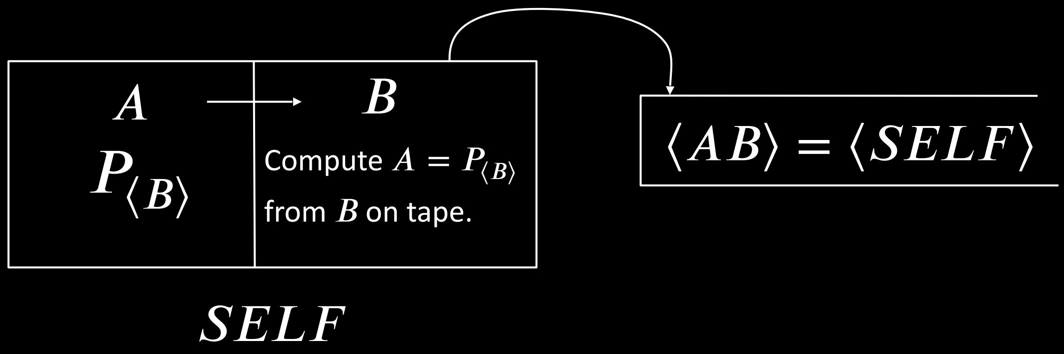


Theorem: For any TM T there is a TM R where for all w
R on input w operates in the same way as T on input $\langle w,\ R \rangle$.

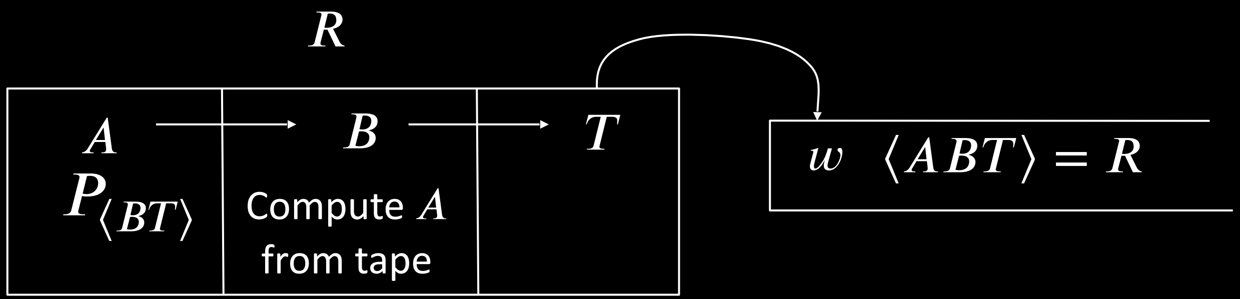


Recall

* Self-Reproducing TM
* The Recursion Theorem: TM T working on <w,A> -> TM R
  * that has access to <R>,
  * behaves like T on <w,R>



Theorem: For any TM T there is a TM R where for all w
R on input w operates in the same way as T on input $\langle w,\ R \rangle$.



Ex 2: Fixed-point Theorem

Ex 2: Fixed-point Theorem

Theorem: For any computable function $f : \Sigma^* \to \Sigma^*$,
there is a TM R such that $L(R) = L(S)$ where $f\big(\langle R \rangle\big) = \langle S \rangle$.

Ex 2: Fixed-point Theorem

Theorem: For any computable function $f: \Sigma^* \to \Sigma^*$,
there is a TM R such that $L(R) = L(S)$ where $f\big(\langle R \rangle\big) = \langle S \rangle$.

In other words, consider f to be a program transformation function.
Then for some program R, its behavior is unchanged by f.

Ex 2: Fixed-point Theorem

Theorem: For any computable function $f\colon\Sigma^* \to \Sigma^*$,
there is a TM R such that $L(R) = L(S)$ where $f\big(\langle R\rangle\big) = \langle S\rangle$.

In other words, consider f to be a program transformation function.
Then for some program R, its behavior is unchanged by f.

Proof: Let R be the following TM.

Ex 2: Fixed-point Theorem

Theorem: For any computable function $f : \Sigma^* \to \Sigma^*$,
there is a TM R such that $L(R) = L(S)$ where $f\big(\langle R\rangle\big) = \langle S\rangle$.

In other words, consider f to be a program transformation function.
Then for some program R, its behavior is unchanged by f.

Proof: Let R be the following TM.

$R =$ "On input w

Ex 2: Fixed-point Theorem

Theorem: For any computable function $f : \Sigma^* \to \Sigma^*$,
there is a TM R such that $L(R) = L(S)$ where $f\big(\langle R \rangle\big) = \langle S \rangle$.

In other words, consider f to be a program transformation function.
Then for some program R, its behavior is unchanged by f.

Proof: Let R be the following TM.

$R =$ "On input w

1. Get own description $\langle R \rangle$.

Ex 2: Fixed-point Theorem

Theorem: For any computable function $f\colon \Sigma^* \rightarrow \Sigma^*$,
there is a TM R such that $L(R) = L(S)$ where $f\big(\langle R\rangle\big) = \langle S\rangle$.

In other words, consider f to be a program transformation function.
Then for some program R, its behavior is unchanged by f.

Proof: Let R be the following TM.

$R =$ "On input w

1. Get own description $\langle R\rangle$.
2. Compute $f\big(\langle R\rangle\big)$ and call the result $\langle S\rangle$.

Ex 2: Fixed-point Theorem

Theorem: For any computable function $f: \Sigma^* \to \Sigma^*$,
there is a TM R such that $L(R) = L(S)$ where $f(\langle R \rangle) = \langle S \rangle$.

In other words, consider f to be a program transformation function.
Then for some program R, its behavior is unchanged by f.

Proof: Let R be the following TM.

$R =$ "On input w

1. Get own description $\langle R \rangle$.
2. Compute $f(\langle R \rangle)$ and call the result $\langle S \rangle$.
3. Simulate S on w."

Ex 2: Fixed-point Theorem

Theorem: For any computable function $f: \Sigma^* \to \Sigma^*$,
there is a TM R such that $L(R) = L(S)$ where $f\big(\langle R \rangle\big) = \langle S \rangle$.

In other words, consider f to be a program transformation function.
Then for some program R, its behavior is unchanged by f.

Proof: Let R be the following TM.

$R =$ "On input w

1. Get own description $\langle R \rangle$.
2. Compute $f\big(\langle R \rangle\big)$ and call the result $\langle S \rangle$.
3. Simulate S on w."

Ex 3: MIN TM is T-unrecognizable

Ex 3: MINTM is T-unrecognizable

Defn: M is a minimal TM if $\left|\langle M' \rangle\right| < \left|\langle M \rangle\right| \rightarrow L(M') \neq L(M)$.

Ex 3: MIN TM is T-unrecognizable

Defn: M is a minimal TM if $\left|\langle M' \rangle\right| < \left|\langle M \rangle\right| \rightarrow L(M') \neq L(M)$.

Thus, a minimal TM has the shortest description among all equivalent TMs.

Ex 3: MINTM is T-unrecognizable

Defn: M is a minimal TM if $\left|\langle M' \rangle\right| < \left|\langle M \rangle\right| \rightarrow L(M') \neq L(M)$.

Thus, a minimal TM has the shortest description among all equivalent TMs.

Let MINTM $= \left\{\langle M \rangle \middle|\ M\right.$ is a minimal TM $\}$.

Theorem: MIN**TM** is T-unrecognizable.

Ex 3: MINTM is T-unrecognizable

Defn: M is a minimal TM if $\left|\langle M'\rangle\right| < \left|\langle M\rangle\right| \rightarrow L(M') \neq L(M)$.

Thus, a minimal TM has the shortest description among all equivalent TMs.

Let MINTM $= \{\langle M\rangle \mid M$ is a minimal TM $\}$.

Theorem: MIN**TM** is T-unrecognizable.

Proof by contradiction: Assume some TM E enumerates MINTM .

Ex 3: MINTM is T-unrecognizable

Defn: M is a minimal TM if $\left|\langle M'\rangle\right| < \left|\langle M\rangle\right| \rightarrow L(M') \neq L(M)$.

Thus, a <u>minimal TM</u> has the shortest description among all equivalent TMs.

Let MINTM $= \left\{\langle M\rangle \middle|\ M \text{ is a minimal TM } \right\}$.

Theorem: MIN**TM** is T-unrecognizable.

Proof by contradiction: Assume some TM E enumerates MINTM .

Consider the following TM R:

Ex 3: MINTM is T-unrecognizable

Defn: M is a minimal TM if $\left|\langle M' \rangle\right| < \left|\langle M \rangle\right| \rightarrow L(M') \neq L(M)$.

Thus, a <u>minimal TM</u> has the shortest description among all equivalent TMs.

Let MINTM $= \{\langle M \rangle \mid M$ is a minimal TM $\}$.

Theorem: MIN**TM** is T-unrecognizable.

Proof by contradiction: Assume some TM E enumerates MINTM .

Consider the following TM R:

$R =$ "On input w

Ex 3: MINTM is T-unrecognizable

Defn: M is a minimal TM if $\left|\langle M' \rangle\right| < \left|\langle M \rangle\right| \rightarrow L(M') \neq L(M)$.

Thus, a <u>minimal TM</u> has the shortest description among all equivalent TMs.

Let MINTM $= \{\langle M \rangle \mid M$ is a minimal TM $\}$.

Theorem: MIN**TM** is T-unrecognizable.

Proof by contradiction: Assume some TM E enumerates MINTM .

Consider the following TM R:

$R =$ “On input w

1. Get own description $\langle R \rangle$.

Ex 3: MINTM is T-unrecognizable

Defn: M is a minimal TM if $\big|\langle M' \rangle\big| < \big|\langle M \rangle\big| \rightarrow L(M') \neq L(M)$.

Thus, a <u>minimal TM</u> has the shortest description among all equivalent TMs.

Let MINTM $= \{\langle M \rangle \big|\ M$ is a minimal TM $\}$.

Theorem: MIN**TM** is T-unrecognizable.

Proof by contradiction: Assume some TM E enumerates MINTM .

Consider the following TM R:

$R =$ "On input w

1. Get own description $\langle R \rangle$.
2. Run enumerator E until some TM B appears, where $\big|\langle R \rangle\big| < \big|\langle B \rangle\big|$.

Ex 3: MINTM is T-unrecognizable

Defn: M is a minimal TM if $\left|\langle M' \rangle\right| < \left|\langle M \rangle\right| \rightarrow L(M') \neq L(M)$.

Thus, a <u>minimal TM</u> has the shortest description among all equivalent TMs.

Let MINTM $= \left\{\langle M \rangle \middle|\ M \text{ is a minimal TM } \right\}$.

Theorem: MIN**TM** is T-unrecognizable.

Proof by contradiction: Assume some TM E enumerates MINTM .

Consider the following TM R:

$R =$ "On input w

1. Get own description $\langle R \rangle$.
2. Run enumerator E until some TM B appears, where $\left|\langle R \rangle\right| < \left|\langle B \rangle\right|$.
3. Simulate B on w."

Ex 3: MINTM is T-unrecognizable

Defn: M is a minimal TM if $\left|\langle M'\rangle\right| < \left|\langle M\rangle\right| \rightarrow L(M') \neq L(M)$.

Thus, a minimal TM has the shortest description among all equivalent TMs.

Let MINTM $= \left\{\langle M\rangle \middle|\ M\right.$ is a minimal TM $\}$.

Theorem: MIN**TM** is T-unrecognizable.

Proof by contradiction: Assume some TM E enumerates MINTM .

Consider the following TM R:

$R =$ "On input w

1. Get own description $\langle R\rangle$.
2. Run enumerator E until some TM B appears, where $\left|\langle R\rangle\right| < \left|\langle B\rangle\right|$.
3. Simulate B on w."

Thus $L(R) = L(B)$ and $\left|\langle R\rangle\right| < \left|\langle B\rangle\right|$ so B isn't minimal, but $\langle B\rangle \in L(E)$, contradiction.

Ex 3: MINTM is T-unrecognizable

Defn: M is a minimal TM if $\left|\langle M'\rangle\right| < \left|\langle M\rangle\right| \rightarrow L(M') \neq L(M)$.

Thus, a minimal TM has the shortest description among all equivalent TMs.

Let MINTM $= \left\{\langle M\rangle \middle| \; M \text{ is a minimal TM } \right\}$.

Theorem: MIN**TM** is T-unrecognizable.

Proof by contradiction: Assume some TM E enumerates MINTM .

Consider the following TM R:

$R =$ "On input w

1. Get own description $\langle R\rangle$.
2. Run enumerator E until some TM B appears, where $\left|\langle R\rangle\right| < \left|\langle B\rangle\right|$.
3. Simulate B on w."

Thus $L(R) = L(B)$ and $\left|\langle R\rangle\right| < \left|\langle B\rangle\right|$ so B isn't minimal, but $\langle B\rangle \in L(E)$, contradiction.

Check-in 11.3

Ex 3: MINTM is T-unrecognizable

Defn: M is a minimal TM if $\left|\langle M'\rangle\right| < \left|\langle M\rangle\right| \rightarrow L(M') \neq L(M)$.

Thus, a minimal TM has the shortest description among all equivalent TMs.

Let MINTM $= \{\langle M\rangle \mid M$ is a minimal TM $\}$.

Theorem: MIN**TM** is T-unrecognizable.

Proof by contradiction: Assume some TM E enumerates

Consider the following TM R:

$R =$ "On input w

1. Get own description $\langle R\rangle$.
2. Run enumerator E until some TM B appears, where $\left|\langle R\rangle\right| < \left|\langle B\rangle\right|$.
3. Simulate B on w."

Thus $L(R) = L(B)$ and $\left|\langle R\rangle\right| < \left|\langle B\rangle\right|$ so B isn't minimal, but $\langle B\rangle \in L(E)$, contradiction.

Check-in 11.3

Let A be an infinite subset of MINTM .

Is it possible that A is T-recognizable?

(a) Yes.

(b) No.

Check-in 11.3

Other applications

Other applications

1. Computer viruses.

Other applications

1. Computer viruses.
2. A true but unprovable mathematical statement due to Kurt Gödel:
   "This statement is unprovable."

Other applications

1. Computer viruses.
2. A true but unprovable mathematical statement due to Kurt Gödel:
   "This statement is unprovable."

Other applications

1. Computer viruses.
2. A true but unprovable mathematical statement due to Kurt Gödel:
   "This statement is unprovable."

Intro to Mathematical Logic

Goal: A mathematical study of mathematical reasoning itself.

Formally defines the language of mathematics, mathematical truth, and provability.

Intro to Mathematical Logic

Goal: A mathematical study of mathematical reasoning itself.

Formally defines the language of mathematics, mathematical truth, and provability.

Gödel’s First Incompleteness Theorem:

In any reasonable formal system, some true statements are not provable.

Intro to Mathematical Logic

Goal: A mathematical study of mathematical reasoning itself.

Formally defines the language of mathematics, mathematical truth, and provability.

Gödel's First Incompleteness Theorem:

In any reasonable formal system, some true statements are not provable.

Proof: We use two properties of formal proofs:

Intro to Mathematical Logic

Goal: A mathematical study of mathematical reasoning itself.

Formally defines the language of mathematics, mathematical truth, and provability.

Gödel's First Incompleteness Theorem:

In any reasonable formal system, some true statements are not provable.

Proof: We use two properties of formal proofs:

1) Soundness: If ϕ has a proof π then ϕ is true.

Intro to Mathematical Logic

Goal: A mathematical study of mathematical reasoning itself.

Formally defines the language of mathematics, mathematical truth, and provability.

Gödel's First Incompleteness Theorem:

In any reasonable formal system, some true statements are not provable.

Proof: We use two properties of formal proofs:

1) Soundness: If ϕ has a proof π then ϕ is true.
2) Checkability: The language $\{\langle \pi, \phi \rangle \mid \pi \text{ is a proof of statement } \phi\}$ is decidable.

Intro to Mathematical Logic

Goal: A mathematical study of mathematical reasoning itself.

Formally defines the language of mathematics, mathematical truth, and provability.

Gödel's First Incompleteness Theorem:

In any reasonable formal system, some true statements are not provable.

Proof: We use two properties of formal proofs:

1) Soundness: If ϕ has a proof π then ϕ is true.
2) Checkability: The language $\{\langle \pi, \phi \rangle \mid \pi$ is a proof of statement $\phi\}$ is decidable.

Checkability implies the set of provable statements $\{ \langle \phi \rangle \mid \phi$ has a proof$\}$ is T-recognizable.

Intro to Mathematical Logic

Goal: A mathematical study of mathematical reasoning itself.

Formally defines the language of mathematics, mathematical truth, and provability.

Gödel's First Incompleteness Theorem:

In any reasonable formal system, some true statements are not provable.

Proof: We use two properties of formal proofs:

1) Soundness: If ϕ has a proof π then ϕ is true.
2) Checkability: The language $\left\{\langle \pi, \phi \rangle \middle| \; \pi \text{ is a proof of statement } \phi \right\}$ is decidable.

Checkability implies the set of provable statements $\{ \langle \phi \rangle \mid \phi$ has a proof$\}$ is T-recognizable.

SImilarly, if we can always prove $\langle M, w \rangle \in \overline{A\text{TM}}$ when it is true, then $\overline{A\text{TM}}$ is T-recognizable (false!).

Intro to Mathematical Logic

Goal: A mathematical study of mathematical reasoning itself.
Formally defines the language of mathematics, mathematical truth, and provability.

Gödel's First Incompleteness Theorem:
In any reasonable formal system, some true statements are not provable.

Proof: We use two properties of formal proofs:
1) Soundness: If ϕ has a proof π then ϕ is true.
2) Checkability: The language $\left\{\langle\pi,\phi\rangle\,\middle|\, \pi\right.$ is a proof of statement $\phi\}$ is decidable.

Checkability implies the set of provable statements $\{ \langle\phi\rangle \mid \phi$ has a proof$\}$ is T-recognizable.

SImilarly, if we can always prove $\langle M,w\rangle \in \overline{A\mathrm{TM}}$ when it is true, then $\overline{A\mathrm{TM}}$ is T-recognizable (false!).

Therefore, some true statements of the form $\langle M,w\rangle \in \overline{A\mathrm{TM}}$ are unprovable.

Intro to Mathematical Logic

Goal: A mathematical study of mathematical reasoning itself.
Formally defines the language of mathematics, mathematical truth, and provability.

Gödel's First Incompleteness Theorem:
In any reasonable formal system, some true statements are not provable.

Proof: We use two properties of formal proofs:

1) Soundness: If ϕ has a proof π then ϕ is true.
2) Checkability: The language $\{\langle \pi, \phi \rangle \mid \pi$ is a proof of statement $\phi\}$ is decidable.

Checkability implies the set of provable statements $\{\langle \phi \rangle \mid \phi$ has a proof$\}$ is T-recognizable.

SImilarly, if we can always prove $\langle M, w \rangle \in \overline{A\mathrm{TM}}$ when it is true, then $\overline{A\mathrm{TM}}$ is T-recognizable (false!).

Therefore, some true statements of the form $\langle M, w \rangle \in \overline{A\mathrm{TM}}$ are unprovable.

Next, we use the Recursion Theorem to give a specific example of a true but unprovable statement.

Intro to Mathematical Logic

Goal: A mathematical study of mathematical reasoning itself.
Formally defines the language of mathematics, mathematical truth, and provability.

Gödel's First Incompleteness Theorem:
In any reasonable formal system, some true statements are not provable.

Proof: We use two properties of formal proofs:

1) Soundness: If ϕ has a proof π then ϕ is true.
2) Checkability: The language $\left\{\langle \pi, \phi \rangle \middle| \ \pi\right.$ is a proof of statement $\phi\}$ is decidable.

Checkability implies the set of provable statements { $\langle \phi \rangle$ | ϕ has a proof} is T-recognizable.

SImilarly, if we can always prove $\langle M, w\rangle \in \overline{A\text{TM}}$ when it is true, then $\overline{A\text{TM}}$ is T-recognizable (false!).

Therefore, some true statements of the form $\langle M, w\rangle \in \overline{A\text{TM}}$ are unprovable.

Next, we use the Recursion Theorem to give a specific example of a true but unprovable statement.

A True but Unprovable Statement

Implement Gödel statement “This statement is unprovable.”

A True but Unprovable Statement

Implement Gödel statement "This statement is unprovable."

Let ϕ_U be the statement $\langle R, 0\rangle \in \overline{A\mathrm{TM}}$ where R is the following TM:

A True but Unprovable Statement

Implement Gödel statement “This statement is unprovable.”

Let ϕ_U be the statement $\langle R, 0\rangle \in \overline{A\text{TM}}$ where R is the following TM:

$R =$ “On any input

1. Obtain $\langle R\rangle$ and use it to obtain ϕ_U.

A True but Unprovable Statement

Implement Gödel statement “This statement is unprovable.”

Let ϕ_U be the statement $\langle R, 0\rangle \in \overline{A\text{TM}}$ where R is the following TM:

$R =$ “On any input

1. Obtain $\langle R\rangle$ and use it to obtain ϕ_U.
2. For each possible proof $\pi = \pi_1, \pi_2, \ldots$

    Test if π is a proof that ϕ_U is true.

    If yes, then *accept*. Otherwise, continue.”

A True but Unprovable Statement

Implement Gödel statement "This statement is unprovable."

Let ϕ_U be the statement $\langle R, 0\rangle \in \overline{A\text{TM}}$ where R is the following TM:

$R =$ "On any input

1. Obtain $\langle R\rangle$ and use it to obtain ϕ_U.
2. For each possible proof $\pi = \pi_1, \pi_2, \ldots$

   Test if π is a proof that ϕ_U is true.

   If yes, then *accept*. Otherwise, continue."

Theorem: (1) ϕ_U has no proof

(2) ϕ_U is true

A True but Unprovable Statement

Implement Gödel statement "This statement is unprovable."

Let ϕ_U be the statement $\langle R, 0 \rangle \in \overline{A\text{TM}}$ where R is the following TM:

$R =$ "On any input

1. Obtain $\langle R \rangle$ and use it to obtain ϕ_U.
2. For each possible proof $\pi = \pi_1, \pi_2, \ldots$

   Test if π is a proof that ϕ_U is true.

   If yes, then *accept*. Otherwise, continue."

Theorem: (1) ϕ_U has no proof

(2) ϕ_U is true

Proof:

A True but Unprovable Statement

Implement Gödel statement "This statement is unprovable."

Let ϕ_U be the statement $\langle R, 0\rangle \in \overline{A\text{TM}}$ where R is the following TM:

$R =$ "On any input

1. Obtain $\langle R\rangle$ and use it to obtain ϕ_U.
2. For each possible proof $\pi = \pi_1, \pi_2, \ldots$

    Test if π is a proof that ϕ_U is true.

    If yes, then *accept*. Otherwise, continue."

Theorem: (1) ϕ_U has no proof

(2) ϕ_U is true

Proof:

(1) If ϕ_U has a proof

A True but Unprovable Statement

Implement Gödel statement "This statement is unprovable."

Let ϕ_U be the statement $\langle R, 0 \rangle \in \overline{A\text{TM}}$ where R is the following TM:

$R =$ "On any input

1. Obtain $\langle R \rangle$ and use it to obtain ϕ_U.
2. For each possible proof $\pi = \pi_1, \pi_2, \ldots$

   Test if π is a proof that ϕ_U is true.

   If yes, then *accept*. Otherwise, continue."

Theorem: (1) ϕ_U has no proof

(2) ϕ_U is true

Proof:

(1) If ϕ_U has a proof $\rightarrow$ TM R accepts 0

A True but Unprovable Statement

Implement Gödel statement "This statement is unprovable."

Let ϕ_U be the statement $\langle R, 0\rangle \in \overline{A\text{TM}}$ where R is the following TM:

$R =$ "On any input

1. Obtain $\langle R\rangle$ and use it to obtain ϕ_U.
2. For each possible proof $\pi = \pi_1, \pi_2, \ldots$

   Test if π is a proof that ϕ_U is true.

   If yes, then *accept*. Otherwise, continue."

Theorem: (1) ϕ_U has no proof

(2) ϕ_U is true

Proof:

(1) If ϕ_U has a proof $\rightarrow$ TM R accepts 0 $\rightarrow \langle R, 0\rangle \in \overline{A\text{TM}}$ is false

A True but Unprovable Statement

Implement Gödel statement "This statement is unprovable."

Let ϕ_U be the statement $\langle R, 0\rangle \in \overline{A\text{TM}}$ where R is the following TM:

$R =$ "On any input

1. Obtain $\langle R\rangle$ and use it to obtain ϕ_U.
2. For each possible proof $\pi = \pi_1, \pi_2, \ldots$

   Test if π is a proof that ϕ_U is true.

   If yes, then *accept*. Otherwise, continue."

Theorem: (1) ϕ_U has no proof

(2) ϕ_U is true

Proof:

(1) If ϕ_U has a proof $\rightarrow$ TM R accepts 0 $\rightarrow$ $\overbrace{\langle R, 0\rangle \in \overline{A\text{TM}}}^{\phi_U}$ is false

A True but Unprovable Statement

Implement Gödel statement "This statement is unprovable."

Let ϕ_U be the statement $\langle R, 0\rangle \in \overline{A\mathrm{TM}}$ where R is the following TM:

$R =$ "On any input

1. Obtain $\langle R\rangle$ and use it to obtain ϕ_U.
2. For each possible proof $\pi = \pi_1, \pi_2, \ldots$

   Test if π is a proof that ϕ_U is true.

   If yes, then *accept*. Otherwise, continue."

Theorem: (1) ϕ_U has no proof

(2) ϕ_U is true

Proof:

(1) If ϕ_U has a proof → TM R accepts 0 → $\overbrace{\langle R, 0\rangle \in \overline{A\mathrm{TM}}}^{\phi_U}$ is false → ϕ_U cannot have a proof.

A True but Unprovable Statement

Implement Gödel statement "This statement is unprovable."

Let ϕ_U be the statement $\langle R, 0\rangle \in \overline{A\mathrm{TM}}$ where R is the following TM:

$R =$ "On any input

1. Obtain $\langle R\rangle$ and use it to obtain ϕ_U.
2. For each possible proof $\pi = \pi_1, \pi_2, \ldots$

   Test if π is a proof that ϕ_U is true.

   If yes, then *accept*. Otherwise, continue."

Theorem: (1) ϕ_U has no proof

(2) ϕ_U is true

Proof:

(1) If ϕ_U has a proof $\rightarrow$ TM R accepts 0 $\rightarrow \overbrace{\langle R, 0\rangle \in \overline{A\mathrm{TM}}}^{\phi_U}$ is false $\rightarrow \phi_U$ cannot have a proof.

(2) If ϕ_U is false $\rightarrow$

A True but Unprovable Statement

Implement Gödel statement "This statement is unprovable."

Let ϕ_U be the statement $\langle R, 0\rangle \in \overline{A\text{TM}}$ where R is the following TM:

$R =$ "On any input

1. Obtain $\langle R\rangle$ and use it to obtain ϕ_U.
2. For each possible proof $\pi = \pi_1, \pi_2, \ldots$

   Test if π is a proof that ϕ_U is true.

   If yes, then *accept*. Otherwise, continue."

Theorem: (1) ϕ_U has no proof

(2) ϕ_U is true

Proof:

(1) If ϕ_U has a proof $\rightarrow$ TM R accepts 0 $\rightarrow \overbrace{\langle R, 0\rangle \in \overline{A\text{TM}}}^{\phi_U}$ is false $\rightarrow \phi_U$ cannot have a proof.

(2) If ϕ_U is false $\rightarrow \langle R, 0\rangle \notin \overline{A\text{TM}}$

A True but Unprovable Statement

Implement Gödel statement "This statement is unprovable."

Let ϕ_U be the statement $\langle R, 0\rangle \in \overline{A\text{TM}}$ where R is the following TM:

$R =$ "On any input

1. Obtain $\langle R\rangle$ and use it to obtain ϕ_U.
2. For each possible proof $\pi = \pi_1, \pi_2, \ldots$

   Test if π is a proof that ϕ_U is true.

   If yes, then *accept*. Otherwise, continue."

Theorem: (1) ϕ_U has no proof

(2) ϕ_U is true

Proof:

(1) If ϕ_U has a proof → TM R accepts 0 → $\overbrace{\langle R, 0\rangle \in \overline{A\text{TM}}}^{\phi_U}$ is false → ϕ_U cannot have a proof.

(2) If ϕ_U is false → $\langle R, 0\rangle \notin \overline{A\text{TM}}$ → R accepts 0

A True but Unprovable Statement

Implement Gödel statement "This statement is unprovable."

Let ϕ_U be the statement $\langle R, 0\rangle \in \overline{A\text{TM}}$ where R is the following TM:

$R =$ "On any input

1. Obtain $\langle R\rangle$ and use it to obtain ϕ_U.
2. For each possible proof $\pi = \pi_1, \pi_2, \ldots$

   Test if π is a proof that ϕ_U is true.

   If yes, then *accept*. Otherwise, continue."

Theorem: (1) ϕ_U has no proof

(2) ϕ_U is true

Proof:

(1) If ϕ_U has a proof → TM R accepts 0 → $\overbrace{\langle R, 0\rangle \in \overline{A\text{TM}}}^{\phi_U}$ is false → ϕ_U cannot have a proof.

(2) If ϕ_U is false → $\langle R, 0\rangle \notin \overline{A\text{TM}}$ → R accepts 0 → R found a proof that ϕ_U is true

A True but Unprovable Statement

Implement Gödel statement "This statement is unprovable."

Let ϕ_U be the statement $\langle R, 0\rangle \in \overline{ATM}$ where R is the following TM:

$R =$ "On any input

1. Obtain $\langle R\rangle$ and use it to obtain ϕ_U.
2. For each possible proof $\pi = \pi_1, \pi_2, \ldots$

 Test if π is a proof that ϕ_U is true.

 If yes, then *accept*. Otherwise, continue."

Theorem: (1) ϕ_U has no proof

(2) ϕ_U is true

Proof:

(1) If ϕ_U has a proof $\rightarrow$ TM R accepts 0 $\rightarrow$ $\overbrace{\langle R, 0\rangle \in \overline{ATM}}^{\phi_U}$ is false $\rightarrow$ ϕ_U cannot have a proof.

(2) If ϕ_U is false $\rightarrow \langle R, 0\rangle \notin \overline{ATM} \rightarrow$ R accepts 0 $\rightarrow$ R found a proof that ϕ_U is true $\rightarrow$ ϕ_U is true.

A True but Unprovable Statement

Implement Gödel statement “This statement is unprovable.”

Let ϕ_U be the statement $\langle R, 0\rangle \in \overline{A\text{TM}}$ where R is the following TM:

$R =$ “On any input

1. Obtain $\langle R\rangle$ and use it to obtain ϕ_U.
2. For each possible proof $\pi = \pi_1, \pi_2, \ldots$

    Test if π is a proof that ϕ_U is true.

    If yes, then *accept*. Otherwise, continue.”

Theorem: (1) ϕ_U has no proof

(2) ϕ_U is true

Proof:

(1) If ϕ_U has a proof $\rightarrow$ TM R accepts 0 $\rightarrow \overbrace{\langle R, 0\rangle \in \overline{A\text{TM}}}^{\phi_U}$ is false $\rightarrow$ ϕ_U cannot have a proof.

(2) If ϕ_U is false $\rightarrow \langle R, 0\rangle \notin \overline{A\text{TM}} \rightarrow$ R accepts 0 $\rightarrow$ R found a proof that ϕ_U is true $\rightarrow$ ϕ_U is true.

Quick review of this topic

1. Self-reference and The Recursion Theorem
2. Various applications.
3. Sketch of Gödel’s First Incompleteness Theorem in mathematical logic.