Rice's Theorem

Definition: $P : \{0,1\}^* \to \{0,1\}$ is a nontrivial property of enumerable languages if

(1) $P(\langle M \rangle)$ depends only on $L(M)$— if $L(M) = L(M')$ then $P(\langle M \rangle) = P(\langle M' \rangle)$.

(2) P is nontrivial— $\exists M_1$ such that $P(\langle M_1 \rangle) = 1$, and $\exists M_0$ such that $P(\langle M_0 \rangle) = 0$.

If P is a property of enumerable languages then we say that a language $L = L(M)$ has property P iff $P(\langle M \rangle) = 1$.

Theorem [Rice]: If P is a nontrivial property of enumerable languages, then

$$L_P = \{\langle M \rangle : \ P(\langle M \rangle) = 1\}$$

is undecidable. (In fact, L_P is not enumerable if $\emptyset$ has property P; and L_P is not co-enumerable if $\emptyset$ does not have property P.)

Proof: Case 1. Assume $\emptyset$ does *not* have property P, ie., $P(K_\emptyset) = 0$, where $K_\emptyset$ is a TM for which $L(M_\emptyset) = \emptyset$. We show $A_{\text{TM}} \leq_{\text{m}} L_P$. To show this, we must exhibit a Turing computable function f for which $\langle M' \rangle = f(\langle M, w \rangle)$ is a machine accepting a language with property P iff M accepts w. We specify the behavior of M' on input x to be:

> "Run M on w.
>
> If M rejects, *reject*.
>
> Run K_P on x, where K_p is a (fixed) machine for which $P(\langle K_P \rangle) = 1$. (Here we are using that P is nontrivial).
>
> If K_P accepts, *accept*; if K_P rejects, *reject*."

Clearly M' is Turing computable from M and w. Observe that

(1) if M accepts w then $L(M') = L(K_P)$, which is a language with property P.

(2) if M does not accept w, then $L(M') = \emptyset$ which, by assumption, is a language which does not have property P.

Case 2. Assume $\emptyset$ *does* have property P. Proceed as above, reducing $\overline{A_{\text{TM}}}$ to L_P. □