Proof. Suppose for contradiction that there existed a computable function $f(n)$ such that $BB(n) \leq f(n)$ for all n. We can use this to contradict the halting theorem, as follows. First observe that once the Busy Beaver function can be upper bounded by a computable function, then for any n, the run time of any halting program of length at most n can also be upper bounded by a computable function. This is because if a program of length n halts in finite time, then a trivial modification of that program (of length larger than n, but by a computable factor) is capable of outputting the run time of that program (by keeping track of a suitable "clock" variable, for instance). Applying the upper bound for Busy Beaver to that increased length, one obtains the bound on run time.

Now, to determine whether a given program S halts in finite time or not, one simply simulates (runs) that program for time up to the computable upper bound of the possible running time of S, given by the length of S. If the program has not halted by then, then it never will. This provides a program P obeying the hypotheses of the halting theorem, a contradiction. □

Remark 1.15.28. A variant of the argument shows that $BB(n)$ grows faster than any computable function: thus if f is computable, then $BB(n) > f(n)$ for all sufficiently large n. We leave the proof of this result as an exercise to the reader.

Remark 1.15.29. Sadly, the most important unsolved problem in complexity theory, namely the *$P \neq NP$ problem*, does not seem to be susceptible to the "no self-defeating object" argument, basically because such arguments tend to be *relativisable*, whereas the $P \neq NP$ problem is not; see Section 1.9 for more discussion. On the other hand, one has the curious feature that many proposed *proofs* that $P \neq NP$ appear to be self-defeating; this is most strikingly captured in the celebrated work of Razborov and Rudich [**RaRu1997**], who showed (very roughly speaking) that any sufficiently "natural" proof that $P \neq NP$ could be used to disprove the existence of an object closely related to the belief that $P \neq NP$, namely the existence of pseudo-random number generators. (I am told, though, that diagonalisation arguments can be used to prove some other inclusions or noninclusions in complexity theory that are not subject to the relativisation barrier, though I do not know the details.)

1.15.4. Game theory. Another basic example of the "no self-defeating object" argument arises from game theory, namely the *strategy stealing argument.* Consider for instance a generalised version of naughts and crosses (tic-tac-toe), in which two players take turns placing naughts and crosses on some game board (not necessarily square, and not necessarily two-dimensional), with the naughts player going first, until a certain pattern of all naughts or